ROCKWOOL Foundation Berlin
Institute for the Economy and the Future of Work (RFBerlin)

DISCUSSION PAPER SERIES 190/26

Racial Screening on the Big Screen? Evidence from the Motion Picture Industry

Liang Zhong, Angela Crema, M. Daniele Paserman

www.rfberlin.com JULY 2026

Racial Screening on the Big Screen? Evidence from the Motion Picture Industry

Authors

Liang Zhong, Angela Crema, M. Daniele Paserman

Reference

JEL Codes: J15, L82

Keywords: Discrimination, Identification, Motion Picture Industry, Machine Learning

Recommended Citation: Liang Zhong, Angela Crema, M. Daniele Paserman (2026): Racial Screening on the Big Screen? Evidence from the Motion Picture Industry. RFBerlin Discussion Paper No. 190/26

Access

Disclaimer

Imprint

RFBerlin
ROCKWOOL Foundation Berlin –
Institute for the Economy
and the Future of Work

Gormannstrasse 22, 10119 Berlin
Tel: +49 (0) 151 143 444 67
E-mail: info@rfberlin.com
Web: www.rfberlin.com

Racial Screening on the Big Screen? Evidence from the Motion Picture Industry*

Liang Zhong ⓡ
University of Hong Kong

Angela Crema ⓡ
University of Rochester

M. Daniele Paserman ⓡ
Boston University, NBER, and Rockwool Foundation Berlin

July 2026

Abstract

We develop a model of discrimination that allows us to interpret observed differences in outcomes across groups, conditional on passing a screening test, as taste-based (employer), statistical, or customer discrimination. We apply this framework to investigate the nature of non-white underrepresentation in the US motion picture industry. Leveraging a novel data set with racial identifiers for the cast of 7,000 motion pictures, we show that, conditional on production, non-white movies exhibit higher average revenues and a smaller variance. Our findings can be rationalized in the context of our model if non-white movies are held to higher standards for production.

Keywords: Discrimination, Identification, Motion Picture Industry, Machine Learning
JEL codes: J15, L82

*We thank Matthew Gudgeon, Patrick Kline, Erzo F.P. Luttmer, Elena Manresa, Anna Weber, Kevin Lang, and seminar participants at EEA-ESEM 2022, SOLE 2022, Georgia Tech, the United States Military Academy, the University of Nebraska-Lincoln, the University of Pittsburgh, the Boston University Labor Reading Group, and the New York University Econometrics Student Lunch for helpful suggestions. We also thank Venkata Andra, Tabitha Fortner, Srihaasa Kompella, Antonia Lehnert, Haoyu Liu, Jacob Park, Richard Pilleul, Jacob Pothen, Rishab Sudhir, Jason Wang, and Hongyi Yu for outstanding research assistance. All errors are our own. ⓡ Author names are in random order following Ray ⓡ Robson (2018).

1 Introduction

An employer must decide whether to hire a job applicant. An admission committee must decide whether to admit a candidate to its entering freshman class. A journal editor must decide whether to accept an article for publication. All these settings are characterized by a *decision maker* who must make an in-or-out decision about an *applicant*, having only imperfect information about the applicant's quality. The decision maker may use information about the applicant's race or gender to guide their decision, which may result in discrimination, i.e., the unequal treatment of applicants with otherwise identical characteristics. The econometrician, however, can typically observe only the ex-post outcomes of these decisions: the worker's productivity, the student's grades, or the number of citations received by an article. If we observe differences by race or gender in outcomes, what can we infer about the extent and nature of discrimination by the decision maker?

This paper develops a methodology to study discrimination in environments in which we only observe outcomes for applicants who pass a screening test. Discrimination may arise because of different standards for selection, noisier signals, or differences in the demand for services offered by different groups. Our framework is especially parsimonious, as it allows one to distinguish between these three mechanisms using only the first two moments of the distribution of observed outcomes among screened-in applicants.

We apply our methodology to study the under-representation of nonwhite actors in the U.S. motion picture industry. Hollywood is of intrinsic interest, not only because of its economic relevance.[1] First, movies are an example of a market in which *consumer* perceptions of quality are shaped not only by the underlying product, but also by the identity of the people associated with it. In such major settings—ranging from services to media, entertainment, and advertising-intensive industries—customer-facing or publicly visible individuals play a central role in shaping demand and firm success. Second, there is a widespread perception of bias in the industry: for example, in the 2010s, only 7% of the nominees for the Academy Awards were African Americans, approximately half of their proportion in the population.[2]

[1] In 2024, the gross domestic product of the Motion Picture and Sound Recording Industries amounted to $120 billion, or about 0.45% of total U.S. goods and services (U.S. Bureau of Economic Analysis, 2025).

[2] https://www.washingtonpost.com/news/arts-and-entertainment/wp/2016/02/26/

The implications of such disparities have the potential to go beyond equal opportunity within the entertainment industry and extend to how audiences, and especially younger viewers, form aspirations and identities through representation and role models (Riley, 2024). Does this underrepresentation reflect racial bias? Although industry representatives sometimes invoke audience preferences to justify the lack of diversity,[3] this hypothesis has not been systematically tested against equally plausible alternatives within a unified empirical framework. The scope for this analysis is strengthened by a key empirical advantage: the availability of an accurate productivity measure—box office revenue—which enables identification of the nature of discrimination.

We begin by presenting a model of screening in the presence of discrimination in the context of the movie industry. The model allows us to interpret differences in box-office revenue (*productivity*), conditional on production (*passing the screening test*). In the model, a studio (the *decision maker*) receives an offer to produce a movie or "script." The studio sees the script *type*, i.e., the expected racial composition of the cast—"white" or "non-white"—and receives a noisy signal of the movie's expected box-office revenue. Based on that information, it must choose whether to *screen in* and produce the movie and release it to the public or not. Our model delivers separate identification of the three types of discrimination previewed earlier: a) *customer discrimination*, whereby moviegoers have a preference for white movies over non-white movies; b) *employer* or *taste-based discrimination*, where the studio derives disutility from producing a non-white movie (Becker, 1957); and c) *statistical discrimination*, where the signal conveyed by non-white movies is less informative about the movie's true quality (Aigner and Cain, 1977).

To test the model's predictions, we construct a novel data set with racial identifiers for

these-charts-explain-how-oscars-diversity-is-way-more-complicated-than-you-think/, accessed on October 26, 2021. Recently, the Academy of Motion Picture Arts and Sciences has announced a multitude of diversity-oriented changes, including diversity requirements for movies that wish to be nominated for the Academy Award in the Best Picture category. In this paper, we analyze a time period that precedes the inclusion of such standards.

[3]For example, in defending the casting of white actors for figures of ancient Egyptian or Hebrew origin in the Biblical epic *Exodus: Gods and Kings*, director Ridley Scott stated "I can't mount a film of this budget, where I have to rely on tax rebates in Spain, and say that my lead actor is Mohammad so-and-so from such-and-such...I'm just not going to get it financed."(Source: https://variety.com/2014/film/news/ridley-scott-exodus-gods-and-kings-christian-bale-1201363668/).

the cast of more than 7,000 motion pictures released in the United States between 1997 and 2017. We obtained the data by scraping the popular website IMDb (IMDb, 2018) and combining it with Opus Data (Nash Information Services, LLC, n.d.), a database with extensive information on the movie industry. The racial identifiers are constructed by combining human raters' classifications and a machine learning architecture that integrates a convolutional neural network (CNN) and support vector machine (SVM; Anwar and Islam, 2017).[4]

In our main analysis, we define a movie as "non-white" if two of the four top-billed performers are classified as non-white.[5] We document the following findings. First, the average box-office revenue of non-white movies is substantially *higher* than that of white movies. The raw non-white/white revenue gap is about 91 log points (150%). The inclusion of a standard set of control variables for other movie characteristics and the cast reduces the gap to roughly 38-43 log points (about 46-54%), still large and highly statistically significant. Second, the box office premium of non-white movies is driven primarily by movies in the bottom half of the distribution. Quantile regressions show that the adjusted gap is around 54 log points (about 72%) at the bottom quantiles of the distribution, but the gap at the upper end of the distribution shrinks to about 25 log points (about 28%). These results are robust to different definitions of non-white movies or different dependent variables (e.g., profit margins or profits). Third, we create a measure of the extent to which a movie's box-office revenue overperforms relative to expectations. Following Moretti (2011), we calculate this as the residual in a regression of opening weekend box-office revenue on the number of opening-weekend theaters. We find that relative to white movies, non-white movies substantially overperform relative to expectations.

These results are not consistent with either customer discrimination or statistical discrimination. Instead, we argue that the results are consistent with taste-based discrimination:[6]

[4]We rely on the machine learning algorithm to classify the 8% of actors in our data whose racial classification was an object of disagreement for more than two of our eight (sometimes nine) human raters. We find that the algorithm obtains a classification accuracy of more than 95% in our validation data set, which is considered excellent in the image classification literature. See section 3 for further details.

[5]The non-white category includes mostly African-Americans but may also include Asians, Hispanics, and other ethnicities.

[6]Our identification argument relies on the ordering of the means of the (observed) white and non-white box-office revenue distributions, as well as on the ordering of the variances. Therefore, our results may be interpreted as taste-based discrimination quantitatively dominating any other forms of discrimination that

non-white movies are held to a higher standard, i.e., they are produced only if the expected revenue surpasses a threshold that is higher than the one set for white movies. This pattern may result from either pure taste-based discrimination or a systematic underestimation of the box-office potential of non-white movies.

This paper is situated within a broad literature documenting and exploring discrimination in a variety of settings. This section focuses on the streams of this body of work to which we directly contribute. We refer the reader to several excellent surveys in economics, including Fang and Moro (2011), Lang and Lehmann (2012), Bertrand and Duflo (2017), Lang and Spitzer (2020), and Onuchic (2026) for a broader overview.

A vast body of literature aims to understand the nature of unequal treatments by either distinguishing between statistical and taste-based discrimination in the data (Guryan and Charles, 2013; Lippens et al., 2022); or testing for the presence of one of the two in a specific market or context. These goals have been pursued experimentally (List, 2004; Zussman, 2013; Doleac and Stein, 2013; Agan and Starr, 2018; Cui et al., 2020; Bohren et al., 2023; Gallen and Wasserman, 2023; Chan, 2026), as well as by testing theoretical predictions on secondary data (Altonji and Pierret, 2001; Knowles et al., 2001; Charles and Guryan, 2008). We contribute to this literature by proposing a simple theoretical framework that nests not only employer taste-based and statistical discrimination, but also customer attitudes, and delivers testable predictions for each source of unequal treatment. We see this as a step forward in the study of segregation in interaction-driven markets such as media and entertainment (Kuppuswamy and Younkin, 2020), healthcare (Chan, 2026), retail and other customer-facing occupations (Holzer and Ihlanfeldt, 1998; Carlsson and Rooth, 2007), as well as online marketplaces and gig platforms (Botelho et al., 2025a,b).

The paper is also related to the literature comparing outcomes between groups to detect the presence of taste-based discrimination, often in the context of unequal treatment of minorities in the criminal justice system. The overarching problem at the heart of average comparisons (or average-based outcome tests; Becker 1957) is that of *infra-marginality*, i.e., in the racial setting, differences in averages might mask both unequal treatment for

may be at play.

individuals that are identical but for their race, as well as racial differences in the distributions of unobserved characteristics. Canay et al. (2024) present an extensive discussion on the conditions required for such tests to be valid. The existing literature has dealt with this problem via either random assignments of candidates to decision makers (Arnold et al., 2022) or exploiting the timing of release decisions made by parole boards (Anwar and Fang, 2015); restrictions on relative bias, when the races of both police officers (or employers) and drivers (or employees) are known (Anwar and Fang, 2006); specifying equilibrium models such that marginal and average success rates of police searches coincide (Knowles et al., 2001); or estimating decision thresholds directly by imposing distributional assumptions (Simoiu et al., 2017; Pierson et al., 2018; Pierson, 2020). In our model, the means (and variances) of observed box-office revenues for white and nonwhite films can be mapped to taste-based discrimination by virtue of the functional form assumptions imposed on the underlying outcome distributions. While we introduce the model using normal distributions to match our empirical setting, we show that the identification results extend to a broader class of parameterizations, allowing researchers to apply the framework across a variety of contexts.

In using higher-order moments of the outcome distribution, our test has a similar flavor to those recently proposed by Bharadwaj et al. (2024) and Benson et al. (2024). Bharadwaj et al.'s test is based on the comparison, in the sense of first-order stochastic dominance, between the entire wage distributions of different groups under the implicit assumption that all workers are employed and there is no screening of workers based on expected productivity. Our approach shares Bharadwaj et al.'s insight that studying a non-binary outcome (over a binary outcome, e.g., callback) adds a margin that allows separate identification of different sources of discrimination. However, we explicitly model the effect of different forms of discrimination on the (continuous) *truncated* distribution of outcomes, conditional on production. In parallel work developed independently, Benson et al. propose a model of racial bias in hiring that nests taste-based discrimination, screening discrimination—where employers are better at screening same-race applicants—and complementary production—where workers are more productive when working with managers of the same race.

They achieve identification through testable implications that rely on the mean and variance of white and non-white workers' productivity under managers of different races, which they test within the retail context. While the two papers share similar testable implications, our model departs from Benson et al. in two respects. The first—the ability to accommodate customer-driven disparities—has already been discussed. Second, our framework does not rely on observing the race of the decision maker. We argue that this is an important contribution to study discrimination in contexts where decisions are likely made by groups rather than single individuals (e.g., admission committees, parole boards, lending organizations, grant review panels); or the decision maker in charge might be influenced by other layers of the organization or industry actors (e.g., media and artistic production, health care treatment approvals, charging decisions, regulatory or legal compliance decisions); or, as is probably quite common, the identity of the decision maker is not observed.

The paper also adds to the line of research that documents the presence of racial discrimination in the motion picture industry (Weaver, 2011; Fowdur et al., 2012). Closest to our work is the paper by Kuppuswamy and Younkin (2020), who find that movies with multiple African-American actors enjoy a box office premium. They rule out customer racial tastes as a discrimination mechanism through an experimental approach. We complement their work and confirm their main conclusion using a more comprehensive data set and an analytical framework that distinguishes among different forms of discrimination using only secondary (box office) data. Our application is also related to a broad empirical literature on labor market and recruiting discrimination, which among the most recent contributions include Åslund et al. (2014); Dustmann et al. (2016); Hedegaard and Tyran (2018); Kline et al. (2022), as well as customer discrimination more broadly (Neumark et al. 1996; Bar and Zussman 2017; Combes et al. 2016; Leonard et al. 2010; Chan 2026 in traditional labor market and service settings; Kahn and Sherer 1988; Nardinelli and Simon 1990; Stone and Warren 1999; Burdekin and Idson 1991 in sport contexts.)

The rest of the paper proceeds as follows. Section 2 presents our theoretical model, discusses its empirical implications, and illustrates generalizability to contexts other than the movie industry through examples and critical extensions. Section 3 describes the data

and the process used to classify performers by race. Section 4 presents the main empirical findings and assesses the robustness of the results to different definitions of race or dependent variables. Section 5 presents suggestive evidence of incorrect beliefs on the revenue potential of non-white movies within the industry. Section 6 discusses and concludes.

2 A model of the screening process

We present here our model of screening. The goal is to derive testable implications for customer, taste-based, and statistical discrimination based on the outcomes of applicants who are screened in.[7] We present the theoretical framework using terminology from the movie industry: the *applicant* is a movie or script, which can be either white- or nonwhite-led; the *decision maker* is the studio;[8] *productivity* corresponds to the film's log revenue; and *customers* are moviegoers. In the baseline model, "the studio" denotes a representative decision maker—or equivalently identical studios using the same type-specific screening rule—which lets us focus on selection into production rather than sorting across studios. After deriving the main predictions, we discuss applications beyond movies and extensions with endogenous movie type and sorting across heterogeneous studios.

With this interpretation in place, we now turn to the formal model.

Step 1: Script arrival

There are two types of movies: white movies, denoted by w, and non-white movies, denoted by b. The studio maximizes expected *log* revenue, denoted by π. The studio receives a script and perfectly observes its type t. However, box office revenues are not observed. We assume that ex-ante box-office revenues of a movie of type t follow a log-normal distribution[9]

[7] Although we focus on applications where customer preferences meaningfully influence decisions, similar empirical patterns in other contexts could be interpreted as *first-moment statistical discrimination*, whereby decision makers act on perceived differences in average productivity across groups (see, e.g., Neumark, 2018). While less often discussed in the literature, this channel may be of independent interest in settings where customer-driven preferences are weak or absent.

[8] In the production stage, the ultimate authority for greenlighting a movie rests with the studio, typically exercised through a chain of executives. The studio often delegates operational tasks—such as initiating projects, arranging financing, and overseeing casting and budgeting—to producers, who carry them out within parameters set by the studio.

[9] The log-normal assumption is made for analytical convenience. In Appendix B, we explore alternative distributional assumptions. Most of our results are not sensitive to the specific distributional assumptions.

with type-specific parameters μ_t and $\sigma^2_{\pi t}$:

$$\pi \mid t \sim N(\mu_t, \sigma^2_{\pi t}), \qquad \forall t \in \{w, b\}$$

.

Step 2: Signal and prior updating

Based on the script, the studio updates its prior about the movie's success. Formally, we can think of the studio observing a signal (y) of the movie's box-office revenue. The signal is normally distributed and is well-calibrated, meaning that in expectation, it is equal to the movie's actual (log) box-office revenue, but it is noisy. Critically, we assume that the precision of the signal may differ by movie racial type. Therefore:

$$y \mid \pi, t \sim N(\pi, \sigma^2_{yt}).$$

Given this setup, it is straightforward to calculate the posterior mean of log box-office revenue, conditional on the signal and the movie's type:

$$E(\pi \mid y, t) = \frac{\sigma^2_{\pi t}}{\sigma^2_{\pi t} + \sigma^2_{yt}} y + \frac{\sigma^2_{yt}}{\sigma^2_{\pi t} + \sigma^2_{yt}} \mu_t. \tag{1}$$

Step 3: Production decision

The studio produces a movie and releases it to the public if the expected log box-office revenue, conditional on the movie's type and signal, exceeds a given threshold. This threshold (the *revenue threshold*) is exogenously given. We can think of it as the reservation revenue from a sequential search model, i.e., the value of the revenue that makes the studio indifferent between producing the movie or waiting for a better script.[10] We denote this revenue threshold π_{0t}, making the critical assumption that the threshold is type-specific.

Later in the paper, we highlight which results depend on the log-normal distribution.

[10] We interpret the race-specific revenue threshold as capturing the disutility cost of producing a movie of a given racial type. In our framework, the studio does not explicitly internalize production costs, but an extension in Section 2.2 implicitly allows the studio to target profits and endogenously choose the movie's racial type, taking into account differences in the cost of hiring white versus non-white casts. If hiring a non-white cast is, on average, cheaper than hiring a white cast (Appendix Figure D.2), then a simple comparison of race-specific revenue thresholds will likely *understate* the extent of taste-based discrimination in the market. Section 4.4 provides further discussion and shows that our results are robust to using profits as the dependent variable.

For example, this could result from the studio having a taste for producing movies of a given type.

The movie is produced if

$$E(\pi|y,t) > \pi_{0t}, \tag{2}$$

This is equivalent to saying that the movie is produced only if the signal y exceeds a given threshold (the *signal threshold*). Based on equation (1) and condition (2), it is easy to show that the signal threshold is

$$\bar{y}_t = \pi_{0t} + (\pi_{0t} - \mu_t)\frac{\sigma_{yt}^2}{\sigma_{\pi t}^2}. \tag{3}$$

The signal threshold is type-specific and depends on the revenue threshold, the parameters of the prior distribution, and the precision of the signal.

This threshold, together with the statistical features of the ex-ante distribution of box-office revenue and the distribution of revenue conditional on the signal, determines the ex-post distribution of box-office revenue. The following proposition establishes the comparative statics of the signal threshold with respect to the parameters of the model.

Proposition 1 *The following comparative statics results hold:*

(a) $\bar{y}_t$ *decreases in* μ_t.

(b) $\bar{y}_t$ *increases in* π_{0t}.

(c) If $\pi_{0t} > \mu_t$, $\bar{y}_t$ *increases in* σ_{yt}^2.

(d) If $\pi_{0t} < \mu_t$, $\bar{y}_t$ *decreases in* σ_{yt}^2.

Proof. See Appendix A ■

The first two statements in Proposition 1 are straightforward and intuitive. If the ex-ante expected (log) revenue is higher (a high μ_t), the movie is produced even if the signal is not very good. Similarly, when the revenue threshold (π_{0t}) is high, the signal must be excellent

to produce the movie. The third and fourth items in the Proposition are more involved but are familiar from the literature on statistical discrimination (Aigner and Cain, 1977; Lundberg and Startz, 1983; Neumark, 2012). Intuitively, if the signal is less precise (a high value of σ_{yt}^2) and the studio wants to produce only high-revenue movies, it will have to set a high signal threshold to make sure it only picks the right tail of the revenue distribution (item (c) in Proposition 1); on the other hand, if the studio only wants to cull out very low revenue movies and the signal is uninformative, the threshold must be set at a low value to ensure that only the very worst (i.e., lowest-revenue) movies are weeded out (item (d) in the proposition).

2.1 Predictions for empirical work

Proposition 1 characterizes the signal threshold that determines whether a movie is produced. In practice, we do not observe this threshold. Instead, we observe box-office revenue only for movies that are produced and released. Let

$$s_t^2 \equiv \text{Var}\left(E[\pi \mid y, t] \mid t\right) = \frac{\sigma_{\pi t}^4}{\sigma_{\pi t}^2 + \sigma_{yt}^2}, \qquad x_t \equiv \frac{\pi_{0t} - \mu_t}{s_t},$$

and let $\lambda(x) = \phi(x)/(1-\Phi(x))$ denote the inverse Mills ratio. The mean and variance of log box-office revenue, conditional on production, are:[11]

$$E(\pi \mid y > \bar{y}_t, t) = \mu_t + s_t \lambda(x_t), \tag{4}$$

$$\begin{aligned} \text{Var}(\pi \mid y > \bar{y}_t, t) &= \sigma_{\pi t}^2 + s_t^2 \left[x_t \lambda(x_t) - \lambda(x_t)^2\right] \\ &= s_t^2 \left[1 + \frac{\sigma_{yt}^2}{\sigma_{\pi t}^2} + x_t \lambda(x_t) - \lambda(x_t)^2\right]. \end{aligned} \tag{5}$$

The second line in (5) is often the most transparent expression for the variance; the third line gives the equivalent factored form.

We can then formulate our central proposition, which enables us to predict how different types of discrimination affect box-office revenues of produced white and non-white movies.

[11] These expressions follow from the moments of a truncated bivariate normal distribution; see Rosenbaum (1961).

Proposition 2 *Let* $E_t \equiv E(\pi \mid y > \bar{y}_t, t)$ *and* $V_t \equiv \text{Var}(\pi \mid y > \bar{y}_t, t)$ *be the mean and variance of log box-office revenue conditional on production, as defined in equations (4) and (5). Holding fixed the other parameters of the model, the following comparative statics results hold:*

(a) E_t *and* V_t *increase in* μ_t.

(b) E_t *increases in* π_{0t}, *while* V_t *decreases in* π_{0t}.

(c) E_t *decreases in* σ^2_{yt}, *while* V_t *increases in* σ^2_{yt}.

Proof. See Appendix A. ■

We first focus on the intuition behind the comparative statics of E_t with respect to the parameters. The intuition for the first two results is straightforward: expected revenue conditional on production is higher, the more to the right lies the prior distribution of revenue (result (a)), and the higher is the revenue threshold (result (b)). The third result implies that expected revenue conditional on production increases with signal precision. This result may seem counter-intuitive, as the signal threshold can either increase or decrease with σ^2_{yt} (Proposition 1, results from (c) and (d)). To gain intuition, it is useful to consider the extreme cases of a perfectly informative ($\sigma^2_{yt} = 0$) vis-à-vis a perfectly uninformative signal ($\sigma^2_{yt} \to \infty$). If the signal is perfectly informative, the movie is produced only if the signal (which is exactly equal to box-office revenue) is above the revenue threshold. This implies that expected revenue conditional on production is strictly greater than μ_t because some movies will be below the threshold and are not produced. On the other hand, if the signal is perfectly uninformative, whether a movie exceeds the signal threshold conveys no information about its revenue – the expected revenue conditional on production is, therefore, μ_t.

For the variance results, it is helpful to consider the case of a perfectly informative signal. The distribution of revenue conditional on production is a truncated normal distribution, with the truncation point equal to the revenue threshold π_{0t}. If the whole distribution is shifted to the right and the threshold remains fixed, it is easy to see that the variance also increases (result (a)). If the revenue threshold π_{0t} increases, the truncation point shifts to the right and the distribution variance decreases (result (b)). As for the third result, it is again

useful to consider the two polar cases of a perfectly informative vs. a perfectly uninformative signal: With a perfectly informative signal, the distribution of box-office revenue is a truncated normal distribution, which necessarily has a smaller variance than the untruncated distribution that results from a perfectly uninformative signal.

We can now use Proposition 2 to characterize the mean and variance of observed box-office revenues for white and non-white movies under different types of discrimination.

Case 1: Customer discrimination. Customer discrimination implies that the viewing public has a preference for white movies over non-white ones. In terms of our model, this means that the entire distribution of log box-office revenue for white movies is shifted to the right relative to the distribution for non-white movies, or $\mu_b < \mu_w$.

Then, by result 1, it follows that $E_b < E_w$, and $V_b < V_w$. We can, therefore, state the following prediction:

Prediction 1 *Under customer discrimination, the mean log box-office revenue for non-white movies is lower than for white movies, and the variance of log box-office revenue for non-white movies is lower than for white movies.*

Case 2: Taste-based discrimination. We can think of taste-based discrimination as the studio suffering a utility loss from producing non-white movies. Holding everything else constant, the studio will produce a non-white movie only if the expected log revenue exceeds a higher threshold than the one she sets for white movies to compensate her for the disutility of producing a non-white movie. In this case, $\pi_{0b} > \pi_{0w}$. By result 2, we have that $E_b > E_w$ and $V_b < V_w$. We can, therefore, state Prediction 2:

Prediction 2 *Under taste-based discrimination, the mean log box-office revenue for non-white movies is higher than for white movies. The variance of log box-office revenue for non-white movies is lower than for white movies.*

Case 3: Statistical discrimination. We classify under statistical discrimination the case where the informativeness of the signal for non-white movies is smaller than the one for white movies. We believe this assumption is plausible as historically there have been fewer

movies with non-white characters, and the (mostly white) studios may find it more difficult to evaluate how successful a movie with non-white characters will be. In this case, $\sigma^2_{yb} > \sigma^2_{yw}$. By result 3, we have that $E_b < E_w$ and $V_b > V_w$. We can, therefore, state prediction 3:

Prediction 3 *Under statistical discrimination, the mean log box-office revenue for non-white movies is lower than for white movies, and the variance of log box-office revenue for non-white movies is higher than for white movies.*

Table 1 summarizes our model predictions. In the remainder of the paper, we use the above predictions to assess the extent and nature of discrimination in the motion picture industry. Before turning to the empirical application, we discuss extensions of the model and the applicability of our framework to settings beyond the movie industry.

2.2 Extensions

Functional form. In Appendixes B.1 and B.2, we explore two departures from the normal-normal model. First, we consider a case in which studios care only about the binary outcome "whether a movie is a hit" and decide to produce the script only if the posterior probability exceeds a certain threshold (we dub this the Beta-Binomial model). The comparative statics for the signal threshold in this model match exactly those of the normal-normal model, and so do the testable predictions. Second, we consider the case where the prior distribution of revenue is Pareto rather than log-normal (the Pareto model). The comparative statics for the signal threshold in the Pareto model match exactly those of the normal-normal model for the cases of customer and taste-based discrimination. The predictions for the observed outcomes, however, are somewhat different: a) under taste-based discrimination, both the mean *and* the variance of log revenue conditional on production are predicted to be higher for non-white movies; and b) under statistical discrimination, both the mean and the variance appear to have a U-shaped relationship with the noise of the signal. Nevertheless, under the Pareto model none of the three forms of discrimination can match the observed patterns that mean log revenue is *higher* for non-white movies, while the variance of log-revenue is *lower* for non-white movies (see Section 4.5).

Endogenous type. The model presented in section 2 assumes that the decision maker receives an *application* of a given type. Within the movie context, this is a natural conceptualization for genres in which the storyline is likely to imply the racial background of the characters, such as dramas and comedies.[12] It also holds in contexts outside of the movie industry: in the editorial setting, where journal editors receive completed articles and make an in-or-out decision on publication; in the hiring and admission settings, whenever positions are "posted," and committees receive curricula from external candidates; etc.

There are, however, settings in which it may be more natural to think in terms of a *product* type (rather than an applicant type), and to treat that as an endogenous decision made by the screener. In the movie context, studios may receive action and adventure scripts without explicit specifications for cast characteristics; in the healthcare context—and more broadly in services—supervisors may select employees from an existing pool and determine how to assign them to different shifts or tasks; in consulting, managers may decide how to group associates into teams, and which team to assign to each client.

To connect our framework more directly to the latter category of settings, we present in Appendix B.3 an extension with endogenous movie types. In that model, the studio receives a race-neutral script and decides whether to produce it as a white movie, as a nonwhite movie, or not at all, based on a pair of race-specific quality signals.[13] We show by simulation that this extension delivers testable implications identical to our main model.

Sorting. The single-studio nature of the model in Section 2 is not well-suited to settings where applicants may sort into decision makers. In practice, job applicants choose where to send their resumes, prospective students where to apply, and scholars choose where to submit their articles.

In the movie context, selection may assume different forms: (a) Scriptwriters may submit their projects to studios that are on average more successful with movies of that type—either because of specialized expertise, or because of market segmentation based on customer

[12] In fact, our empirical results are driven by genres in which the assumption of non-race-neutral scripts is more likely to hold (see Table 7).

[13] To simplify the analysis, we assume in the extension that the studio bases its production decision on net returns. This assumption is consistent with the main model, in which gross and net returns yield equivalent implications.

preferences. Within a multi-studio version of our model with studio types B and W, this amounts to assuming, without loss of generality, $\mu_{wW} > \mu_{wB}$ and $\mu_{bW} < \mu_{bB}$ (where μ_{jJ} stands for the mean of the prior distribution for a movie of type j produced by studio J; and similar notation applies to the remaining parameters); (b) Alternatively, scripts might be submitted to studios to take advantage of preference heterogeneity across decision makers: without loss of generality, $\pi_{0wW} < \pi_{0wB}$ and $\pi_{0bW} > \pi_{0bB}$; (c) Finally, sorting may arise from studios' comparative advantage in evaluating a particular movie type, potentially stemming from accumulated learning over time, e.g., $\sigma^2_{ywW} < \sigma^2_{ywB}$ and $\sigma^2_{ybW} > \sigma^2_{ybB}$.

Appendix B.4 characterizes the conditions under which these three sorting mechanisms generate the same qualitative mean-variance predictions as, respectively, customer discrimination, taste-based discrimination, and statistical discrimination in the baseline model. The appendix also shows that the strongest versions of sorting imply full sorting across studio types, a prediction that is not supported in our data.

Inaccurate beliefs. Finally, we consider the case in which the studio may have inaccurate beliefs about the parameters of the model for non-white movies. Appendix B.5 shows that when the studio underestimates the mean of the prior distribution of log revenue or underestimates the precision of the signal, it will apply a systematically higher threshold for producing non-white movies. Thus, even absent explicit animus, inaccurate beliefs can generate outcomes observationally equivalent to taste-based discrimination.

2.3 Beyond movies

We now expand on the generalizability of our findings by discussing other settings in which our framework can be meaningfully applied.

Settings where public image and representation are important. The movie industry is just one example of a setting where racial bias among customers may arise due to homophily and the desire for representation. Sports agents, talent managers, and fashion brand executives may all make recruiting decisions based on how they expect their audiences to perceive memorabilia (Nardinelli and Simon, 1990), endorsements, and campaign imagery—beyond their own preferences and the uncertainty surrounding eventual outcomes.

Services involving interpersonal trust. In professional settings such as hospitals, counseling centers, and law or consulting firms, employers may factor in how patients or clients will perceive prospective employees, as trust often shapes the effectiveness of these relationships. This holds regardless of the source of customer-side bias—whether it stems from pure taste, lack of information, or misinformation (Chan, 2026).

Markets relying on public ratings. The rise of online rating platforms has amplified the salience and consequences of customer perceptions in customer-facing sectors such as sales and hospitality—domains where discrimination has been shown to be more pronounced than in other parts of the economy (Kline et al., 2022; Edelman et al., 2017).

Thus, while our empirical application focuses on the movie industry, the analytical framework can be applied more broadly. We view our approach as particularly useful in settings in which demand for the product or service also depends on the personal characteristics of the provider, so that there is real scope for customer discrimination. In these settings, being able to distinguish between different types of discrimination becomes especially important.

3 Data on movies

We now turn to the empirical section of the paper, where we test the model implications within the movie context. We start by discussing the classification of white and non-white movies in our dataset.

3.1 Facial classification

A key ingredient of our paper is creating a data set with racial identifiers for movie casts. We focus on the four top-billed actors in the movies in our sample. Top-billed actors—those whose names appear first in the film's credits—are typically given greater prominence in movie posters, advertising materials, and trailers. We view this as a primary channel through which viewers form their initial impressions of a movie (Kagan et al., 2024).[14] We

[14]Holding the script fixed, which actors receive top billing may be the outcome of contractual negotiations. When top billing does not reflect the prominence of roles in the film, the movie production process may be better understood through the lens of our model extension with endogenous type (see Appendix B.3). Our results remain robust when controlling for cast star power, a proxy for actors' contractual strength.

use the top four actors as a natural operational definition of the main cast, because movie websites such as IMDb typically present only a few principal cast members in their summary view.

Some recent papers have used machine learning tools to classify images based on skin tone (Adukia et al., 2023; Colella, 2021). We note that these methods are only partially adequate for our purposes. First, we are interested in classifying images of all non-white actors, including those of Asian, Native American, and other ethnicities that are hard to classify based on skin tone alone. Second, even the most accurate machine-learning algorithm will yield some error rate and, most importantly, may not be able to fully capture human perceptions, which are likely the most important dimension for classification in an entertainment context. Therefore, we relied on a team of ten human raters to assign racial identifiers to more than 7000 performer images downloaded from the popular website IMDb (IMDb, 2018).[15]

Each rater was assigned 8 blocks of about 800 performers[16] and was asked to assess whether they thought the person in the image was *White/Caucasian*, *Black/African-American*, *Hispanic*, *Asian*, *Native American/Pacific Islander*, *South Asian*, or *Other*. The option *Unable to Tell* was also made available to the respondents. Raters were specifically instructed not to consult the internet for any information about the performer and to classify the image based on their perception alone. This procedure resulted in between 8 and 9 human ratings for each of the performers in our data set. We assigned to each image the modal classification as long as no more than two raters disagreed on that image's classification.[17] This allowed us to classify about 92% of the performers in our sample as either White (79.8%), Black (9.1%) or Asian (2.6%). For the remaining performers in the sample, we used the machine learning algorithm proposed by Anwar and Islam (2017), described in more detail in Appendix C.

[15]The raters were Boston University undergraduate students recruited in June-July 2023. Seven out of ten were economics majors, four were women, and six were international students. Having young raters is an advantage in this context, as the U.S. moviegoing public skews young (https://www.pewresearch.org/short-reads/2026/03/06/as-the-academy-awards-approach-a-look-at-moviegoing-habits-in-the-united-states/).

[16]One rater completed only four blocks.

[17]We grouped together the White/Caucasian and Hispanic categories, as we realized that it was difficult to accurately distinguish between the two. None of the substantive results in the paper are meaningfully affected if we do not impose this grouping.

3.2 Additional variables

Our analysis is based on a sample of more than 7000 motion pictures released in the United States between 1997 and 2017. Our sample comes from Opus Data (Nash Information Services, LLC, n.d.), a database with detailed information on the movie industry. The data set includes all movies released or re-released since 1997 for which domestic or international theatrical or video revenue data are available. Our main dependent variable, total revenue at the movie level, is complete and reliable from 1997 onward, as it is constructed from weekend and weekly revenue figures, for which OPUS has maintained full coverage since that year. We rely on IMDb for the approximately 5% of observations in our sample for which OPUS revenues are unavailable. We gather aggregate financial data (box office revenue, production budget, opening weekend revenue, etc.) and metadata (genre, production method) for all movies in our sample.

The main variable of interest in our data set is gross domestic box-office revenue, in 2005 dollars.[18] Additional variables include production costs, movie run time, Metacritic score, release date, MPAA rating, number of theaters in which the movie was released, and number of weeks in which the movie was in theaters. Finally, we collected information on the gender and age of the four top-billed performers. We create a variable called "star power," equal to the cumulative box-office revenue of all movies in which each performer appeared up to the release date of the current movie.[19]

[18] Our baseline definition of a movie's revenue includes domestic box-office revenues and excludes international box-office sales as well as DVD and Blu-ray revenues. While the information for revenues other than from domestic theaters is available in OPUS, it is not in IMDB, which is the data source we use for revenues whenever the information in OPUS is missing. Reassuringly, we note that, as we restrict the analysis to non-missing OPUS data, the progressive inclusion of DVD, Blu-ray, and international sales does not qualitatively alter our main results. The results are available upon request. Our definition of box-office revenues also excludes streaming revenues, which are not available in our data. Throughout most of our analysis period, which ends in 2017, streaming accounted for at most a small share of spending on entertainment (see Appendix Figure D.1).

[19] For movies whose top-billed cast has zero prior cumulative box-office revenue, Ln(Star power) is coded as zero.

3.3 Summary statistics

Summary statistics are shown in Table 2. The top panel shows that about 12 percent of the top-billed performers in our sample are non-white. About three-quarters of the movies have zero non-white performers, and about 18 percent have only one non-white performer. Our baseline analysis defines a movie as non-white if at least two of the four top-billed performers are non-white. Based on this definition, about 8 percent of the movies in our sample are non-white. We also assess the robustness of the results to different definitions of non-white movies.

As for the other variables, the distribution of box office revenue is heavily skewed to the right. Therefore, we use its logarithm as the main dependent variable in our baseline analysis. We collapse the "niche" genres into broader categories so that all movies fall into one of five broad genres. For some of the variables, we only have incomplete data: For example, production costs are available for only about 56 percent of the sample,[20] while the Metacritic score is available only for 71 percent of the sample. To maximize sample size, in the empirical analysis, we replace missing values with zeros and add a dummy variable indicating that the variable is missing if the missing value is not central to the analysis.

4 Empirical tests of the model's predictions

4.1 Non-parametric analysis

Figure 1 presents a box-whisker plot of box-office revenue by the number of non-white performers in the movie (out of the four top-billed actors). The mean box-office revenue increases markedly with the number of non-white performers, while the dispersion of the distribution decreases. Also, the 25^{th} percentile (and lower adjacent value) visibly increases with the number of non-white performers. On the other hand, the 75^{th} percentile is quite stable across cast racial compositions, and the upper adjacent value falls slightly. We interpret these patterns as the left tail of the non-white movie distribution being missing, which is consistent

[20] Probit and logit regressions suggest that movies with higher revenue have a significantly (in the statistical sense) higher probability of non-missing cost information, while the conditional difference between white and non-white movies is not statistically distinguishable from zero. The main result is robust to restricting the sample to observations with non-missing cost information: see columns 5 and 6 in Table 3.

with the notion that non-white movies are held to a higher standard for production.

4.2 White-nonwhite gap in average revenue: OLS regressions

We now turn to a more systematic test of our model predictions, starting with an analysis of how the average revenue differs across white and nonwhite movies in our sample. The main regression model is the following:

$$\ln y_{it} = \beta_0 + \beta_1 Nonwhite_{it} + \beta_2 X_{it} + \delta_t + \varepsilon_{it}, \tag{6}$$

where y_{it} denotes domestic box-office revenue, in 2005 U.S. dollars, of movie i released in year t; $Nonwhite_{it}$, the key explanatory variable of interest, is a dummy variable indicating whether at least two of the four top-billed performers are non-white; X_{it} is a vector of additional control variables, including both cast (average age, gender composition, the "star power" variable described previously) and movie (production budget, MPAA rating, Metacritic score, run time, genre dummies) characteristics; δ_t is a year-of-release fixed effect, and ε_{it} is the robust standard error clustered by distributor.[21] Through the lens of our theoretical model, including covariates on the right-hand side of the regression amounts to recognizing that the parameters driving the studio's decision—prior revenue, signal variance, and production threshold—may vary not only with the script's racial type but also with other script characteristics. In this perspective, we want to test how the estimated racial gap changes upon the inclusion of additional observable traits.

The results are presented in Table 3. The first column of the table shows the unadjusted difference in mean log revenue between white and non-white movies without any controls. The mean box-office revenue of non-white movies is almost 2.5 times as high as that of white movies ($\exp(0.914) \approx 2.5$). In column 2, we include controls for other characteristics of the cast (average age and gender composition), and the coefficient remains almost unchanged. In column 3, we add controls for the production budget, a dummy for whether the production budget is missing, and other movie characteristics, including genre, year-of-release, and distributor fixed effects. The coefficient on the non-white indicator drops to 0.379, implying

[21] We collapse all distributors with only one movie in our data set into one single distributor category (Unknown).

that non-white movies earn about 46 percent more than white movies at the box office. In column 4, we add controls for cast star power and movie Metacritic score,[22] and the coefficient remains stable at 0.377 (46% revenue gap) and highly significant. We next restrict the analysis to movies with non-missing data on production costs. In column 5, we replicate column 2, and the coefficient drops from 0.819 to 0.529, which explains in part what drives the decrease of the coefficient from column 2 to column 3. Finally, column 6 replicates column 4 in the restricted sample. The results are mostly unchanged – the coefficient on the non-white indicator rises to 0.433, implying that non-white movies earn on average about 54 percent more than white movies.

These initial results on the differences between white and non-white movies are not consistent with either a model of customer discrimination, where audiences prefer white movies to non-white movies nor a model of statistical discrimination, where the signal conveyed by non-white scripts is less informative about future box-office revenue. Both models predict that white movies should have, on average, higher box-office revenue than non-white movies, in contrast to our findings. Instead, the results are consistent with a model of taste-based discrimination, where non-white movies are held to a higher standard, i.e., they are only produced if the revenue exceeds a higher threshold than the one required of white movies. In what follows, we look at how other features of the distribution differ between white and non-white movies.

4.3 Quantile regressions

The model described in Section 2 derives predictions for not only the mean but also the variance of box-office revenues. In this subsection, we analyze other measures of dispersion, namely the white-nonwhite gap at different percentiles of the revenue distribution. Specifically, we estimate a series of quantile regressions of the following type:

$$Q_\tau(\ln y_{it}|Nonwhite, X) = \gamma_{0\tau} + \gamma_{1\tau} Nonwhite_{it} + \gamma_{2\tau} X_{it} + \delta_t,$$

[22] We include these controls separately as (a) they are not necessarily implied by the script, and (b) might be endogenous to race.

where $Q_\tau(\ln y_{it}|Nonwhite, X)$ denotes the τth conditional quantile of the distribution of log box-office revenue, and $\tau \in \{0.05, 0.10, ...0.95\}$. The main coefficients of interest are the $\gamma_{1\tau}$'s, which measure the gap in conditional quantiles across the white and non-white box-office revenue distributions..

Figure 2 plots the quantile regression coefficients against the quantiles. As was already apparent from the box-whisker plots in Figure 1, from the 20$^{\text{th}}$ quantile onwards there is a clear downward trend in the quantile coefficients: The white-nonwhite gap at the lower quantiles is around 60 log points, while it is only about 20 log points at the upper quantiles. This finding reinforces the interpretation that there is "missing mass" in the left-tail of the non-white revenue distribution, or in other words, that non-white movies at the low end of the distribution of box-office revenue are not produced, while comparable white movies are.

4.4 Robustness

We next investigate the robustness of our results to different definitions of movie type and different dependent variables.

Classification of non-white movies. In Table 4, we consider additional definitions of "non-white" movies. The first column in the table reproduces the results using our baseline classification of non-white movies as those in which at least two of the four top-billed performers are non-white. The first row in the table shows the OLS results from Table 3, while the remaining rows present the quantile regression coefficients at selected quantiles. All specifications include the full set of control variables.

In column 2, we change the definition of non-white movies to include all movies in which at least *one* of the four top-billed performers is non-white. We view this as a noisier indicator of the movie type, as a non-white actor may be cast in a supporting role in a movie that is mainly about white characters and storylines (a form of *tokenism*). Using this definition, the OLS coefficient is substantially reduced (about 23 log points) but still large and highly statistically significant. The pattern of quantile regression coefficients is also clearly downward sloping, with the gap going from about 26 log points at the 25$^{\text{th}}$ to about 19 log points at the 90$^{\text{th}}$ percentile. In column 3, we replace the dummy indicator for non-white movies with the

share of non-whites among the four top-billed performers. The results are quantitatively and qualitatively similar to those of the baseline specification. Finally, in column 4, we classify a movie as non-white only if the top-billed performer is non-white. According to this definition, the average white-nonwhite premium is slightly smaller than in the baseline (46 log points), and the pattern of the quantile regression coefficients is also downward sloping.[23]

On the whole, Table 4 shows that the main conclusions regarding the white-nonwhite premium and the nature of discrimination in the industry are not sensitive to the exact definition of non-white movies.

Choice of the dependent variable. In all the analyses so far, we have looked at the logarithm of box-office revenue as the primary dependent variable of interest. The main reason for this choice is that box-office revenue is readily available for almost all movies, and it has been traditionally used as the primary metric for assessing the commercial success of a movie. However, studios also consider the expected cost of a movie when making production decisions. While we have addressed this in part by including production costs as an explanatory variable in Table 3, one may also want to work with profits directly. In the Opus data set, we observe a movie's production budget for about 56% of all movies so that we can calculate various measures of profit.[24] We report the results of this analysis in Table 5. The sample includes only those movies for which we observe the production budget. All specifications include the full set of control variables.

In column 1, we use the logarithm of the gross profit margin as a dependent variable, defined as the ratio of domestic box-office revenue to the production budget. The results are broadly consistent with those in the previous sections: Non-white movies have on average a substantially higher profit margin, and the white-nonwhite gap becomes smaller as we move from the low to the high end of the distribution.

[23] Our main result is robust to restricting the sample to movies with cast popularity (see section 3 for a definition of "star power") below the median, ruling out that the non-white premium that we find is driven by "superstar" non-white movies exclusively.

[24] Our measure of profits should only be viewed as a coarse estimate. First, the production budget does not represent the entirety of a movie's production costs, which typically also include marketing costs. Marketing costs are rarely disclosed. Second, the studio typically does not collect all of the box-office revenue, as theaters also receive a cut depending on bilateral negotiations as well as other factors such as the length of time that the movie has been in theaters. Third, cost sharing – a common practice in the movie industry (Weinstein, 1998) – is likely to reduce the extent to which studios internalize costs in their decision making.

In column 2, we focus on the total profit, calculated simply as the difference between box-office revenue and the production budget. It is still the case that the average non-white movie earns a higher profit than the average white movie (by about $8.9 million), holding other characteristics fixed. However, we no longer observe a clear declining pattern in the white-nonwhite gap as we move from lower to upper quantiles in the profit distribution. In fact, the gap appears to be fairly stable (at least in the statistical sense) at all quantiles of the distribution. This could be partly due to the shape of the profit distribution, which tends to be quite right-skewed. We confirm this in column 3, where we use the level of box-office revenue (rather than the logarithm) as the dependent variable. We find a positive but imprecisely estimated premium favoring non-white movies, and now the pattern of quantile regression coefficients shows that the gap becomes larger as we move from the low to the high end of the distribution. We note that, given the substantial right skewness in the revenue distribution, the predictions regarding the variance of box-office revenues *in levels* conditional on production derived from a model that assumes normal distributions no longer hold necessarily.

4.5 The white-nonwhite gap in residual variance

An alternative approach to verify our dispersion predictions is to explore how the residual variance differs across white and non-white movies. Borrowing from the heteroskedasticity literature, we posit that the squared residuals from the OLS regression in equation 6 have the form:

$$u_{it}^2 = \exp(Z_{it}'\alpha),$$

where the vector Z_{it} contains a subset of the variables included in the main regression (potentially, all of them); We then estimate regressions of $\ln \hat{u}_{it}^2$ on the racial indicator and additional control variables. The results are reported in Table 6.

In column 1, the residual variance is assumed to depend only on the racial indicator. Consistent with the results of the box-whisker plot and quantile regressions, we find that non-white movies have a substantially lower residual variance than white movies. In columns 2 and 3, we progressively add additional controls to the variance regression. The results are

essentially unchanged – the residual variance of non-white movies is lower than that of white movies.

In the remaining three columns, we experiment with different definitions of non-white movies. The coefficients on the racial variable in the residual regressions are somewhat smaller in absolute value, but still highly statistically significant.[25]

Overall, our results are consistent with what our theoretical model defines as taste-based discrimination, i.e., non-white movies being held to a higher standard, which results in a higher mean and lower variance of box-office revenue for the produced non-white movies.

4.6 Heterogeneity Analysis

In Table 7, we explore the heterogeneity of our results along a number of different dimensions.[26] First, we look at whether our results are driven by movies produced and distributed by specific segments of the industry. One concern is that our results may capture differences between movies produced by the major studios (the so-called "Big-Six")[27] vs. those produced by smaller studios. It could be that the smaller box-office revenue of white movies reflects the fact that these are often produced by small independent studios, while non-white movies are passed over by these studios altogether. Columns 1 and 2 of the table, however, show that this is not the case: the non-white revenue premium is present among movies distributed by both types of studios.

We next look at differences across genres (columns 3-5 of the table). The non-white premium is more pronounced among comedies and dramas, where the script is more likely to convey information about the racial composition of the cast. By contrast, the non-white premium is small and not statistically significant in action/adventure movies.

[25] The coefficient on the racial variable remains negative and statistically significant when the residual is constructed from a regression where the outcome variable is the log of the profit margin (as in column 1 of Table 5), but it is imprecisely estimated when the residual is constructed from regressions where the dependent variables are profits and revenues in levels (as in columns 2 and 3 of Table 5.) Results are available upon request.

[26] Descriptive statistics for the variables used in the heterogeneity analysis are provided in Appendix D.

[27] These Big-Six studios are: Warner Bros., Paramount Pictures, Walt Disney, Sony / Columbia Pictures, Universal Studios, and 20th Century Fox. These six studios accounted for almost 90% of the US/Canadian market as of 2007. In 2020, Disney acquired 20th Century Fox, and the group is now commonly referred to as the "Big Five". The Big-Six control is included in our regression analysis as a control.

Columns 6 and 7 examine heterogeneity by time period. We look separately at movies produced before and after 2007, the median year in our sample. If taste-based discrimination declines over time, either because of a change in attitudes or because of a change in the competitive landscape, we would expect the non-white premium to shrink. There is some evidence in support of these hypotheses: The non-white premium is 56 log points before 2007 and 31 log points from 2007 onward, even though the difference is not statistically significant at conventional levels ($p = 0.121$).

Finally, in columns 8 and 9, we look at whether the results differ by the gender composition of the cast. We define "female" movies as those in which (strictly) more than 50% of the leading actors are women. The non-white premium is slightly larger among female movies, suggesting that non-white movies must pass an even higher threshold if the cast is predominantly female.

5 Alternative explanation: is the industry surprised?

The empirical results so far suggest that non-white movies are held to higher production standards than white movies. A candidate interpretation of these patterns is that studios dislike producing non-white movies and face a disutility cost every time they produce one. As a result, the expected revenue for producing non-white movies needs to be higher than the expected revenue for producing white movies (in the context of the model, $\pi_{0b} > \pi_{0w}$).

An alternative, non-mutually exclusive, interpretation is that the industry systematically underestimates the revenue potential of non-white movies relative to white movies.[28,29] In other words, expected box-office revenue for non-white movies conditional on the signal is $E[\pi_b \mid y_b]$, but studios perceive it to be $\hat{E}[\pi_b \mid y_b] = E[\pi_b \mid y_b] - e_b$, with $e_b > 0$. This explanation would yield similar predictions to the ones derived from taste-based discrimination, even if the nature of discrimination in the industry is quite different.

Our model intrinsically cannot identify taste-based disutility costs and biased beliefs

[28] Davidson (2012) argues that the film industry is known for having a hard time forecasting movies' success, as well as analyzing past results.

[29] See Chan (2026); Bohren et al. (2023); Esponda et al. (2022); Bordalo et al. (2016); Fong and Luttmer (2011) for evidence of inaccurate beliefs in other contexts.

separately (see Appendix B.5 for details). Nevertheless, we can make some progress on this front by exploiting the decision that distributors make on the number of theaters in which the movie is displayed on the opening weekend. We argue that this is a proxy of the market's expectation of the movie's potential *after* production, as distributors' decision-making is less likely to be affected by taste-based or statistical discrimination: While studios "sign" a movie as a creation of theirs and create a permanent bond with the film, distributors are more likely to make choices based on purely profit-maximizing considerations once the movie has been produced. Moreover, statistical discrimination should also be of relatively less importance at the distribution stage, because distributors also observe the ex-post quality of the movie rather than just the script.

We conjecture that distributors choose the number of theaters based on expected customer demand. If non-white movies are displayed in fewer theaters than white movies, this indicates that distributors expect relatively smaller revenue from the non-white movies. Therefore, if non-white movies have the same level of customer demand but are displayed in fewer theaters, we conclude that distributors underestimate their revenue potential.

Using data on the number of screens in which a movie is shown, we test the hypothesis that the industry systematically underestimates the revenue potential of non-white movies. Specifically, we first regress first-weekend box office revenues on the number of theaters in which movies are projected over the first weekend upon their release. Following Moretti (2011), we interpret this as a proxy for the industry expectation of a movie's box-office revenue. The residuals from this regression can then be viewed as a measure of the industry's underestimation or overestimation of a movie's revenue potential. If the non-white mean residual is significantly larger than the white mean residual, this suggests that the industry systematically underestimates non-white movies' revenue potential relative to white movies.

We start by running a simple bivariate regression of log first-weekend revenues on the log number of theaters. The R-squared of this regression is 0.89 (column 1 of Table 8), and it remains relatively stable as further controls are added (columns 2 and 3). The number of theaters is, hence, a good predictor of first-weekend revenues.

We then test whether the residuals obtained from the regressions in the top panel of the

table differ on average between non-white and white movies. The results are presented in the bottom panel of the table. We find that the mean residual for non-white movies is positive across specifications, while the mean residual for white movies is close to zero. That is, the industry underestimates the first-weekend success of non-white movies relative to white movies. The difference between the white and the non-white residual is always statistically significant. We conclude that our results might be at least partly explained by a systematic underestimation of non-white movies' box-office potential within the industry.

6 Conclusion

This paper presents a framework for detecting the extent and nature of discrimination in contexts in which decision-makers screen applicants. The econometrician can only observe the outcomes of applicants who successfully pass the screening process. The framework nests several leading theories of discrimination and derives a rich set of testable empirical predictions.

We apply these tests in the context of racial representation in the U.S. motion picture industry. We show that non-white movies earn a box-office premium. The gap is particularly pronounced at low quantiles of the distribution, suggesting that non-white movies with low box-office potential are never produced; in other words, non-white movies are held to a higher standard in the production decision. In the context of our model, this evidence is consistent with taste-based discrimination, i.e., studios suffering a utility loss from producing non-white movies. The evidence is also consistent with studios and distributors having inaccurate beliefs and systematically underestimating the revenue potential of non-white movies. On the other hand, the evidence is not consistent with simple customer discrimination against non-white movies, nor with a statistical discrimination story in which the signal sent by non-white movies is less precise.

These results may appear puzzling to the extent that they hint at lost profits and relatively slow learning in the industry for non-white movies' potential. While our results indicate that the non-white revenue premium has declined from 56 log points before 2007 to 31 log points from 2007 onward, the point estimate for the latter period is far from

zero in an economically significant way. Some of the *a priori* plausible explanations do not seem applicable to our setting: it is unlikely that learning is hindered by the non-white premium being too small to be detected or consequential, or that too few non-white movies are produced. The fact that, conditional on production, non-white movies are relatively more successful rules out customers' attitudes and pre-market discrimination as leading explanations (Becker, 1957). We argue that other industry-specific forces might be at play, including the high concentration of the motion picture industry, with the "Big Six" studios typically accounting for more than 80% of the industry's total market share. In non-competitive industries, firms may have more latitude to indulge their discriminatory taste. The introduction and growth of streaming services and smartphone applications in the 2010s appear to have increased the amount of competition in the industry (Kuehn and Lampe, 2023), which is consistent with the declining non-white premium in the latter part of our sample. There are also documented challenges and uncertainty that industry actors face in predicting movies' revenues and profitability, even when data are available (Lash and Zhao, 2016).[30] We leave the investigation of this puzzle to future research, along with the analysis of the consequences of the recent diversity-promoting rules that the Academy of Motion Picture Arts and Sciences set for those aspiring to best-picture qualifications (Sperling, 2020a,b).

While the specific application in this paper looked at the motion picture industry, our model can be readily applied to other contexts in which decision makers can use group identifiers to screen applicants, and one can observe the outcome or productivity of successful applicants: the output of workers hired for a particular job, the academic performance of students admitted to a freshman class, or the number of citations accumulated by a published journal article. These other contexts are promising avenues for future research.

[30]For discussions and examples in the public press, see Yahr (2016); Gladwell (2006); Thompson (2013); Snee (2016).

References

Adukia, Anjali, Eble, Alex, Harrison, Elizabeth, Runesha, Haidong B., and Szasz, Tamas (2023). "What We Teach About Race and Gender: Representation in Images and Text of Children's Books," *Quarterly Journal of Economics*, 138 (4), 2225–2285.

Agan, Amanda and Starr, Sonja (2018). "Ban the box, criminal records, and racial discrimination: A field experiment," *The Quarterly Journal of Economics*, 133 (1), 191–235.

Aigner, Dennis J. and Cain, Glen G. (1977). "Statistical theories of discrimination in labor markets," *Industrial and Labor Relations Review*, 30 (2), 175–187.

Altonji, Joseph G. and Pierret, Charles R. (2001). "Employer learning and statistical discrimination," *The Quarterly Journal of Economics*, 116 (1), 313–350.

Anwar, Imran and Islam, Nazrul Ul (2017). "Learned features are better for ethnicity classification," *Cybernetics and Information Technologies*, 17 (3), 152–164.

Anwar, Shamena and Fang, Hanming (2006). "An alternative test of racial prejudice in motor vehicle searches: Theory and evidence," *American Economic Review*, 96 (1), 127–151.

——— (2015). "Testing for racial prejudice in the parole board release process: Theory and evidence," *The Journal of Legal Studies*, 44 (1), 1–37.

Arnold, David, Dobbie, Will, and Hull, Peter (2022). "Measuring racial discrimination in bail decisions," *American Economic Review*, 112 (9), 2992–3038.

Åslund, Olof, Hensvik, Lena, and Skans, Oskar Nordström (2014). "Seeking Similarity: How Immigrants and Natives Manage in the Labor Market," *Journal of Labor Economics*, 32, 405–441.

Bar, Revital and Zussman, Asaf (2017). "Customer Discrimination: Evidence from Israel," *Journal of Labor Economics*, 35 (4), 1031–1059.

Baricz, Árpád (2008). "Mills' ratio: Monotonicity patterns and functional inequalities," *Journal of Mathematical Analysis and Applications*, 340 (2), 1362–1370.

Becker, Gary S. (1957). *The Economics of Discrimination*: University of Chicago Press.

Benson, Alan, Board, Simon, and ter Vehn, Moritz Meyer (2024). "Discrimination in hiring: Evidence from retail sales," *Review of Economic Studies*, 91 (4), 1956–1987.

Bertrand, Marianne and Duflo, Esther (2017). "Field experiments on discrimination," in *Handbook of Economic Field Experiments*, 1, 309–393.

Bharadwaj, Prashant, Deb, Rajat, and Renou, Ludovic (2024). "Statistical discrimination and the distribution of wages."

Bohren, J. Aislinn, Haggag, Kareem, Imas, Alex, and Pope, Devin G. (2023). "Inaccurate statistical discrimination: An identification problem," *Review of Economics and Statistics*, 1–45.

Bordalo, Pedro, Coffman, Katherine, Gennaioli, Nicola, and Shleifer, Andrei (2016). "Stereotypes," *The Quarterly Journal of Economics*, 131 (4), 1753–1794, 10.1093/qje/qjw029.

Botelho, Tristan L., DeCelles, Katherine A., Humes, Demetrius, and Jun, Sora (2025a). "Research: How Gig Platforms Can Mitigate Racial Bias in Ratings," *Harvard Business Review*, https://hbr.org/2025/03/research-how-gig-platforms-can-mitigate-racial-bias-in-ratings, Published March 14, 2025.

Botelho, Tristan L., Jun, Sora, Humes, Demetrius, and DeCelles, Katherine A. (2025b). "Scale dichotomization reduces customer racial discrimination and income inequality," *Nature*, 639, 395–403, 10.1038/s41586-024-07052-6, Open access.

Burdekin, Richard CK and Idson, Todd L (1991). "Customer Preferences, Attendance and the Racial Structure of Professional Basketball Teams," *Applied Economics*, 23 (1), 179–186.

Canay, Ivan A., Mogstad, Magne, and Mountjoy, Jack (2024). "On the use of outcome tests for detecting bias in decision making," *The Review of Economic Studies*, 91 (4), 2135–2167.

Carlsson, Magnus and Rooth, Dan-Olof (2007). "Evidence of ethnic discrimination in the Swedish labor market using experimental data," *Labour Economics*, 14 (4), 716–729.

Chan, Alex (2026). "Customer Discrimination and Quality Signals," Conditionally Accepted.

Charles, Kerwin Kofi and Guryan, Jonathan (2008). "Prejudice and wages: an empirical assessment of Becker's The Economics of Discrimination," *Journal of Political Economy*, 116 (5), 773–809.

Colella, Francesco (2021). "Who Benefits from Support? The Heterogeneous Effects of Supporters on Athletes' Performance by Skin Colour," Mimeo., Université de Lausanne.

Combes, Pierre-Philippe, Decreuse, Bruno, Laouenan, Morgane, and Trannoy, Alain (2016). "Customer Discrimination and Employment Outcomes: Theory and Evidence from the French Labor Market," *Journal of Labor Economics*, 34 (1), 107–160.

Cui, Ruomeng, Li, Jun, and Zhang, Dennis J (2020). "Reducing Discrimination with Reviews in the Sharing Economy: Evidence from Field Experiments on Airbnb," *Management Science*, 66 (3), 1071–1094, 10.1287/mnsc.2018.3253.

Davidson, Adam (2012). "How Does the Film Industry Actually Make Money?" *The New York Times Magazine*, https://www.nytimes.com/2012/07/01/magazine/how-does-the-film-industry-actually-make-money.html.

Doleac, Jennifer L. and Stein, Luke C. (2013). "The visible hand: Race and online market outcomes," *The Economic Journal*, 123 (572), F469–F492.

Dustmann, Christian, Glitz, Albrecht, and Schönberg, Uta (2016). "Referral-Based Job Search Networks," *Review of Economic Studies*, 83, 514–546.

Edelman, Benjamin, Luca, Michael, and Svirsky, Dan (2017). "Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment," *American Economic Journal: Applied Economics*, 9 (2), 1–22, 10.1257/app.20160213.

Esponda, Ignacio, Oprea, Ryan, and Yuksel, Sevgi (2022). "Discrimination Without Reason: Biases in Statistical Discrimination."

Fang, Hanming and Moro, Andrea (2011). "Theories of statistical discrimination and affirmative action: A survey," in *Handbook of Social Economics*, 1, 133–200.

Fong, Christina M and Luttmer, Erzo FP (2011). "Do Fairness and Race Matter in Generosity? Evidence from a Nationally Representative Charity Experiment," *Journal of Public Economics*, 95 (5-6), 372–394.

Fowdur, Linley, Kadiyali, Vrinda, and Prince, Jeffrey (2012). "Racial bias in expert quality assessment: A study of newspaper movie reviews," *Journal of Economic Behavior & Organization*, 84 (1), 292–307.

Gallen, Yana and Wasserman, Melanie (2023). "Does Information Affect Homophily?" *Journal of Public Economics*, 222, 104876, 10.1016/j.jpubeco.2023.104876.

Gladwell, Malcolm (2006). "The Formula," October 16, https://www.newyorker.com/magazine/2006/10/16/the-formula, Accessed: 2024-11-10.

Guryan, Jonathan and Charles, Kerwin K. (2013). "Taste-based or statistical discrimination: the economics of discrimination returns to its roots," *The Economic Journal*, 123 (572), F417–F432.

Hedegaard, Morten S. and Tyran, Jean-Robert (2018). "The Price of Prejudice," *American Economic Journal: Applied Economics*, 10, 40–63.

Holzer, Harry J. and Ihlanfeldt, Keith R. (1998). "Customer discrimination and employment outcomes for minority workers," *The Quarterly Journal of Economics*, 113 (3), 835–867.

IMDb (2018). "Internet Movie Database," https://www.imdb.com/, Data scraped by the authors between 2017-2018.

Kagan, Dima, Shapira, Bracha, Puzis, Rami, and Baram-Tsabari, Ayelet (2024). "Ethnic representation analysis of commercial movie posters," *Humanities and Social Sciences Communications*, 11 (1), 1–12, 10.1057/s41599-023-02040-y.

Kahn, Lawrence M and Sherer, Peter D (1988). "Racial Differences in Professional Basketball Players' Compensation," *Journal of Labor Economics*, 6 (1), 40–61.

Kline, Patrick, Rose, Evan K., and Walters, Christopher R. (2022). "Systemic Discrimination Among Large US Employers," *The Quarterly Journal of Economics*, 137 (4), 1963–2036.

Knowles, John, Persico, Nicola, and Todd, Petra (2001). "Racial bias in motor vehicle searches: Theory and evidence," *Journal of Political Economy*, 109 (1), 203–229.

Kuehn, Jörg and Lampe, Ryan (2023). "Competition and Product Composition: Evidence from Hollywood," *International Journal of Industrial Organization*, 91, 102981, 10.1016/j.ijindorg.2023.102981.

Kuppuswamy, Venkat and Younkin, Peter (2020). "Testing the Theory of Consumer Discrimination as an Explanation for the Lack of Minority Hiring in Hollywood Films," *Management Science*, 66 (3), 1227–1247.

Lang, Kevin and Lehmann, Jee-Yeon K. (2012). "Racial discrimination in the labor market: Theory and empirics," *Journal of Economic Literature*, 50 (4), 959–1006.

Lang, Kevin and Spitzer, Ariella K. (2020). "Race discrimination: An economic perspective," *Journal of Economic Perspectives*, 34 (2), 68–89.

Lash, Michael T. and Zhao, Kang (2016). "Early Predictions of Movie Success," *Journal of Management Information Systems*, 33 (3), 874–903, https://www.jstor.org/stable/10.2307/26614064.

Leonard, Jonathan S, Levine, David I, and Giuliano, Laura (2010). "Customer Discrimination," *The Review of Economics and Statistics*, 92 (3), 670–678.

Lippens, Louis, Baert, Stijn, Ghekiere, Astrid, Verhaeghe, Pieter-Paul, and Derous, Eva (2022). "Is labour market discrimination against ethnic minorities better explained by taste or statistics? A systematic review of the empirical evidence," *Journal of Ethnic and Migration Studies*, 48 (17), 4243–4276.

List, John A. (2004). "The nature and extent of discrimination in the marketplace: Evidence from the field," *The Quarterly Journal of Economics*, 119 (1), 49–89.

Lundberg, Shelly J. and Startz, Richard (1983). "Private Discrimination and Social Intervention in Competitive Labor Markets," *American Economic Review*, 73 (3), 340–347.

Moretti, Enrico (2011). "Social learning and peer effects in consumption: Evidence from movie sales," *The Review of Economic Studies*, 78 (1), 356–393.

Nardinelli, Clark and Simon, Curtis (1990). "Customer Racial Discrimination in the Market for Memorabilia: The Case of Baseball," *The Quarterly Journal of Economics*, 105 (3), 575–595.

Nash Information Services, LLC (n.d.). "OpusData," https://www.opusdata.com/, Proprietary database. Data provided to the authors under license.

Neumark, David (2012). "Detecting Discrimination in Audit and Correspondence Studies," *Journal of Human Resources*, 47 (4), 1128–1157.

——— (2018). "Experimental research on labor market discrimination," *Journal of Economic Literature*, 56 (3), 799–866.

Neumark, David, Bank, Roy J, and Van Nort, Kyle D (1996). "Sex Discrimination in Restaurant Hiring: An Audit Study," *The Quarterly Journal of Economics*, 111 (3), 915–941.

Onuchic, Paula (2026). "Recent contributions to theories of discrimination," *JPE: Micro*, Conditionally Accepted.

Pierson, Emma (2020). "Assessing racial inequality in COVID-19 testing with Bayesian threshold tests," *arXiv preprint*.

Pierson, Emma, Corbett-Davies, Sam, and Goel, Sharad (2018). "Fast threshold tests for detecting discrimination," in *International conference on artificial intelligence and statistics*, 96–105: PMLR.

Riley, Emma (2024). "Role Models in Movies: The Impact of Queen of Katwe on Students' Educational Attainment," *The Review of Economics and Statistics*, 106 (2), 334–351.

Rosenbaum, S. (1961). "Moments of a Truncated Bivariate Normal Distribution," *Journal of the Royal Statistical Society. Series B (Methodological)*, 23 (2), 405–408.

Simoiu, Claudia, Corbett-Davies, Sam, and Goel, Sharad (2017). "The problem of infra-marginality in outcome tests for discrimination," *arXiv preprint*.

Snee, Tom (2016). "Predicting Box Office: Boffo or Bomb?," February, https://now.uiowa.edu/news/2016/02/predicting-box-office-boffo-or-bomb, Accessed: 2024-11-10.

Sperling, Nicole (2020a). "The Oscars Will Add a Diversity Requirement for Eligibility," *The New York Times*, https://www.nytimes.com/2020/09/08/movies/oscars-diversity-rules-best-picture.html, Accessed: 2024-11-10.

——— (2020b). "Academy Sets New Diversity Requirements for Oscars Best Picture Eligibility," *The New York Times*, https://www.nytimes.com/2020/09/09/movies/oscars-best-picture-diversity.html, Accessed: 2024-11-10.

Stone, Eric W and Warren, Ronald S (1999). "Customer Discrimination in Professional Basketball: Evidence from the Trading-Card Market," *Applied Economics*, 31 (6), 679–685.

Thompson, Ben (2013). "Solving Equation of a Hit Film Script, With Data," May 6, https://www.nytimes.com/2013/05/06/business/media/solving-equation-of-a-hit-film-script-with-data.html, Accessed: 2024-11-10.

U.S. Bureau of Economic Analysis (2025). "Gross Domestic Product: Motion Picture and Sound Recording Industries (512) in the United States [USMOTPICSNDNGSP]," https://fred.stlouisfed.org/series/USMOTPICSNDNGSP, Retrieved from FRED, Federal Reserve Bank of St. Louis.

Weaver, Andrew J. (2011). "The role of actors' race in white audiences' selective exposure to movies," *Journal of Communication*, 61 (2), 369–385.

Weinstein, Mark (1998). "Profit-sharing contracts in Hollywood: Evolution and analysis," *The Journal of Legal Studies*, 27 (1), 67–112.

Yahr, Emily (2016). "It's hard to predict a movie's profitability, but you learn some lessons along the way," https://www.washingtonpost.com/news/arts-and-entertainment/wp/2016/05/16/its-hard-to-predict-a-movies-profitability-but-you-learn-some-lessons-along-the-way/, Accessed: 2024-11-10.

Zussman, Asaf (2013). "Ethnic discrimination: Lessons from the Israeli online market for used cars," *The Economic Journal*, 123 (572), F433–F468.

Tables and Figures

Table 1: Model Predictions

Discrimination Source	**Mathematical Definition**	**Comparative Statics: Expected Value**	**Comparative Statics: Variance**
Taste-based The studio bears a utility loss producing non-white movies.	$\pi_{0b} > \pi_{0w}$ The production threshold is relatively higher for non-white movies.	$E_b > E_w$	$V_b < V_w$
Customer The viewing public has a preference for white movies over non-white movies.	$\mu_b < \mu_w$ The distribution of box-office revenues for white movies is shifted to the right, relative to that of non-white movies.	$E_b < E_w$	$V_b < V_w$
Statistical The studio has "less" or "worse" information on non-white movies' potential.	$\sigma^2_{yb} > \sigma^2_{yw}$ The signal for non-white movies is less informative.	$E_b < E_w$	$V_b > V_w$

Note: Summary of the model predictions. In our notation, w (b) denotes white (non-white) movies; π_{0b}, π_{0w} are the type-specific production thresholds; μ_b, μ_w denote the type-specific means of the box-office revenue distributions; $\sigma^2_{yb}, \sigma^2_{yw}$ stand for the type-specific signal variances. See Section 2 for the derivations.

Table 2: Summary Statistics

Variable	(1) N	(2) Mean	(3) SD	(4) Min	(5) Max
PANEL A: Classification of movies by type					
Share of non-white performers	7,840	0.12	0.24	0	1
At least one non-white	7,840	0.26	0.44	0	1
At least two non-whites	7,840	0.08	0.27	0	1
Distribution of the number of non-white performers (percentages):					
0	74.3				
1	17.9				
2	4.5				
3	2.1				
4	1.3				
PANEL B: Other variables					
Gross revenue (in Millions of 2005 Dollars)	7,205	25.9	54.1	$1.94 * 10^{-5}$	804
Ln(Gross revenue)	7,205	14.14	3.39	2.97	20.50
Cost (in Millions of 2005 Dollars)	3,955	37.4	44.7	$1.1 * 10^{-3}$	907
Ln(Cost)	3,955	16.72	1.45	7.00	20.63
Run time (minutes)	6,804	103.51	18.49	38	600
IMDB score	4,915	6.25	0.97	1.50	9
Metacritic score	4,915	51.58	17.10	1	100
Average age of billed performers	7,715	41.92	10.42	10	99
Star power (in Millions)	7,840	262	303	0	2,350
Ln(Star power)	7,840	17.50	4.54	0	21.58
Number of weeks	6,491	11.62	14.66	1	476
Ln(Number of screens)	6,078	4.88	2.95	0.69	8.43
Distribution of movies by genre (percentages):					
Action	16.62				
Animation	0.17				
Comedy	26.27				
Drama	36.57				
Other	20.37				
"Big Six" (percentages):					
Share of movies	31.70				
Share of total revenues	77.69				

Note: Source: authors' calculations. Data sources are described in Section 3.

Table 3: The non-white revenue premium

Sample:	(1) Full Ln(Gross Rev)	(2) Full Ln(Gross Rev)	(3) Full Ln(Gross Rev)	(4) Full Ln(Gross Rev)	(5) Non-missing cost Ln(Gross Rev)	(6) Non-missing cost Ln(Gross Rev)
Race: at least two non-white	0.914*** (0.201)	0.819*** (0.201)	0.379*** (0.081)	0.377*** (0.083)	0.529*** (0.163)	0.433*** (0.082)
Share of female		-1.018*** (0.234)	0.105 (0.074)	0.100 (0.073)	-0.638*** (0.216)	0.063 (0.103)
Average age		-0.045*** (0.007)	0.000 (0.003)	0.001 (0.004)	-0.016*** (0.006)	-0.005 (0.003)
ln(Cost)			0.471*** (0.056)	0.469*** (0.053)		0.565*** (0.048)
= 1 if ln(Cost)			5.684*** (0.883)	5.762*** (0.817)		
ln(Star Power)				0.005 (0.008)		-0.034*** (0.009)
Metacritic score				0.020*** (0.003)		0.020*** (0.002)
Movie controls			Y	Y		Y
Distributor FEs			Y	Y		Y
N	6943	6943	6943	6943	3856	3856
R^2	0.006	0.026	0.303	0.332	0.012	0.349

Note: Data sources and specification are described in Sections 3 and 4. Cast control variables include the share of females and the average age of the four top-billed performers. "Star power" is defined as the log of performers' cumulative box office revenues up to the movie release date. Movie control variables include indicators for movie genre, indicator of whether the movie is from the "Big 6", run time, MPAA rating, year fixed effects, and indicators for missing run time, or MPAA rating. When the Metacritic score is controlled for, an indicator for missing values is also included. The movie budget cost (in the log) is included among the control variables when revenue is the dependent variable. Standard errors clustered by distributor in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Robustness with respect to different definitions of "non-white" movies

	(1)	(2)	(3)	(4)
Race:	At least two non-white Ln(Gross Revenue)	At least one non-white Ln(Gross Revenue)	Share of non-white Ln(Gross revenue)	Leading role is non-white Ln(Gross revenue)
Race	0.455***	0.239***	0.643***	0.474***
	(0.090)	(0.061)	(0.119)	(0.080)
Q10	0.531***	0.234**	0.596***	0.497***
	(0.114)	(0.110)	(0.213)	(0.119)
Q25	0.535***	0.259***	0.817***	0.524***
	(0.129)	(0.063)	(0.200)	(0.128)
Q50	0.398***	0.224***	0.609***	0.331***
	(0.079)	(0.055)	(0.120)	(0.090)
Q75	0.304***	0.203***	0.536***	0.336***
	(0.099)	(0.050)	(0.117)	(0.086)
Q90	0.254***	0.178***	0.414***	0.273***
	(0.064)	(0.057)	(0.112)	(0.067)
Cast controls	Y	Y	Y	Y
Movie controls	Y	Y	Y	Y
N	6943	6943	6943	6943

Note: Data sources and specification are described in Sections 3 and 4. Rows Q10, Q25, Q50, Q75, and Q90 report quantile-regression estimates at the 10th, 25th, 50th, 75th, and 90th percentiles of the conditional distribution of the dependent variable, respectively. Cast control variables include the share of females, the average age of the four top-billed performers, and "star power" (defined as the log of performers' cumulative box office revenues up to the movie release date). Movie control variables include indicators for movie genre, indicator of whether the movie is from the "Big 6", run time, Metacritic score, MPAA rating, year fixed effects, and indicators for missing run time, Metacritic score, or MPAA rating. The movie budget cost (in the log) is included among the control variables when revenue is the dependent variable. Standard errors clustered by distributor in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Robustness to different dependent variables

Sample:	(1) Non-missing cost variable Ln(Profit Margin+1)	(2) Non-missing cost variable Profit(in million)	(3) Non-missing cost variable Revenue(in million)
Race: At least two non-white	0.562***	8.891***	2.235
	(0.096)	(2.183)	(2.966)
Q10	0.448***	5.445***	3.558***
	(0.120)	(1.495)	(0.961)
Q25	0.360***	7.307***	6.355***
	(0.093)	(1.375)	(1.487)
Q50	0.493***	9.473***	6.900***
	(0.107)	(1.423)	(1.806)
Q75	0.352***	10.465***	7.596**
	(0.077)	(3.412)	(3.349)
Q90	0.282***	9.946***	6.366
	(0.078)	(3.113)	(5.520)
Cast controls	Y	Y	Y
Movie controls	Y	Y	Y
N	3856	3856	3856

Note: Data sources and specification are described in Sections 3 and 4. Rows Q10, Q25, Q50, Q75, and Q90 report quantile-regression estimates at the 10th, 25th, 50th, 75th, and 90th percentiles of the conditional distribution of the dependent variable, respectively. Cast control variables include the share of females, the average age of the four top-billed performers, and "star power" (defined as the log of performers' cumulative box office revenues up to the movie release date). Movie control variables include indicators for movie genre, indicator of whether the movie is from the "Big 6", run time, Metacritic score, MPAA rating, year fixed effects, and indicators for missing run time, Metacritic score, or MPAA rating. The movie budget cost (in logs) is included among the control variables when revenue is the dependent variable. Standard errors clustered by distributor in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Conditional residual variance regressions:
robustness with respect to different definitions of non-white movies

	(1)	(2)	(3)	(4)	(5)	(6)
Race definition	At least two	At least two	At least two	At least one	Share	Leading role
	Dependent variable: Ln(residual square)					
Race	-0.480***	-0.483***	-0.345***	-0.106*	-0.362***	-0.222***
	(0.100)	(0.099)	(0.095)	(0.059)	(0.109)	(0.083)
Cast Controls		Y	Y	Y	Y	Y
Movie Controls			Y	Y	Y	Y
N	6943	6943	6943	6943	6943	6943

Note: Data sources and specification are described in Sections 3 and 4.5. Cast control variables include the share of females, the average age of the four top-billed performers, and "star power" (defined as the log of performers' cumulative box office revenues up to the movie release date). Movie control variables include indicators for movie genre, indicator of whether the movie is from the "Big 6", run time, Metacritic score, MPAA rating, year fixed effects, and indicators for missing run time, Metacritic score, or MPAA rating. The movie budget cost (in the log) is included among the control variables when revenue is the dependent variable. Standard errors are clustered by distributor in the main regressions only. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Heterogeneity Analysis

	(1) Distributor Not Big-6	(2) Distributor Big-6	(3) Genre: Action/Adventure	(4) Genre: Comedy	(5) Genre: Drama
Race: At least two non-white	0.565***	0.372**	0.053	0.843***	0.439***
	(0.131)	(0.099)	(0.202)	(0.176)	(0.126)
P-Value of the difference	0.224		0.011		
Q25	0.856***	0.249**	0.056	0.757***	0.481***
	(0.139)	(0.097)	(0.204)	(0.190)	(0.167)
Q75	0.457***	0.160**	0.033	0.402**	0.413***
	(0.125)	(0.081)	(0.098)	(0.183)	(0.133)
Cast and movie controls	Y	Y	Y	Y	Y
N	4488	2455	1135	1880	2590

	(6) Period: Pre-2007	(7) Period: Post-2007	(8) Gender: ≤50% female	(9) Gender: >50% female
Race: At least two non-white	0.563***	0.314**	0.430***	0.691**
	(0.097)	(0.131)	(0.086)	(0.302)
P-Value of the difference	0.121		0.370	
Q25	0.541***	0.462***	0.529***	0.772**
	(0.164)	(0.119)	(0.136)	(0.314)
Q75	0.362***	0.160	0.255***	0.507
	(0.122)	(0.133)	(0.084)	(0.358)
Cast and movie controls	Y	Y	Y	Y
N	2774	4169	5933	1010

Note: Data sources and specification are described in Sections 3 and 4.5. In all specifications, the sample is restricted to observations with non-missing data on production costs. Rows Q25 and Q75 report quantile-regression estimates at the 25th and 75th percentiles of the conditional distribution of the dependent variable, respectively. Cast control variables include the share of females, the average age of the four top-billed performers, and "star power" (defined as the log of performers' cumulative box office revenues up to the movie release date). Movie control variables include indicators for movie genre, indicator of whether the movie is from the "Big 6", run time, Metacritic score, MPAA rating, year fixed effects, and indicators for missing run time, Metacritic score, or MPAA rating. The movie budget cost (in the log) is included among the control variables when revenue is the dependent variable. Standard errors clustered by distributor in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Regressions of Log First-Weekend Revenues on Log Opening-Weekend Theaters

VARIABLES	(1) First-weekend revenues, log	(2) First-weekend revenues, log	(3) First-weekend revenues, log
First-weekend # theaters, log	0.990*** (0.016)	0.980*** (0.017)	0.841*** (0.016)
Cast controls	N	Y	Y
Movie controls	N	N	Y
Residuals: white vs non-white			
Average white	-0.014	-0.014	-0.016
Average non-white	0.151	0.152	0.181
Average difference	-0.165	-0.166	-0.197
p-value of t-test (two-sided)	0.001	0.001	0.000
N	6,276	6,276	6,276
R^2	0.889	0.890	0.928

Note: Data sources and specification are described in Sections 3 and 5. Cast control variables include the share of females, the average age of the four top-billed performers, and "star power" (defined as the log of performers' cumulative box office revenues up to the movie release date). Movie control variables include indicators for movie genre, indicator of whether the movie is from the "Big 6", run time, Metacritic score, MPAA rating, year fixed effects, movie budget cost (in the log), and indicators for missing run time, Metacritic score, MPAA rating, and movie budget cost. Relative to the baseline sample used in Table 3, 633 observations are excluded due to missing data on first-weekend revenues or theaters, while 55 are lost due to taking logs (zero values.) Standard errors clustered by distributor in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

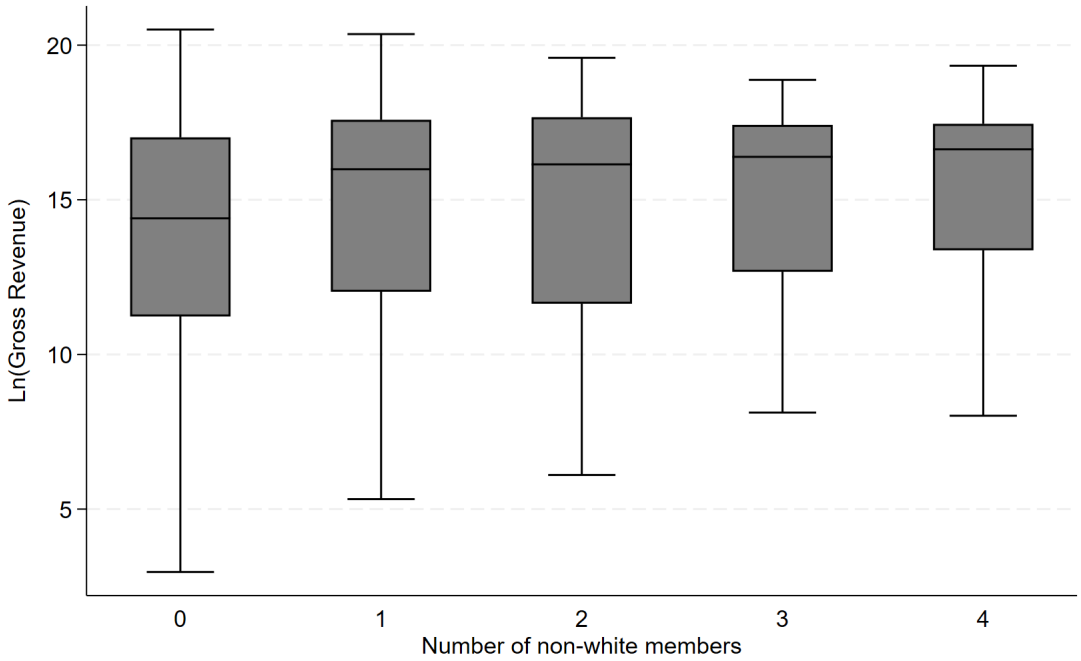

Figure 1: Revenue Distribution by Number of Non-White Top-Billed Performers

Note: The figure reports the distribution of log domestic gross box-office revenue, measured in 2005 dollars, by the number of non-white performers among the four top-billed cast members. The box denotes the interquartile range, the horizontal line denotes the median, and the whiskers denote adjacent values.

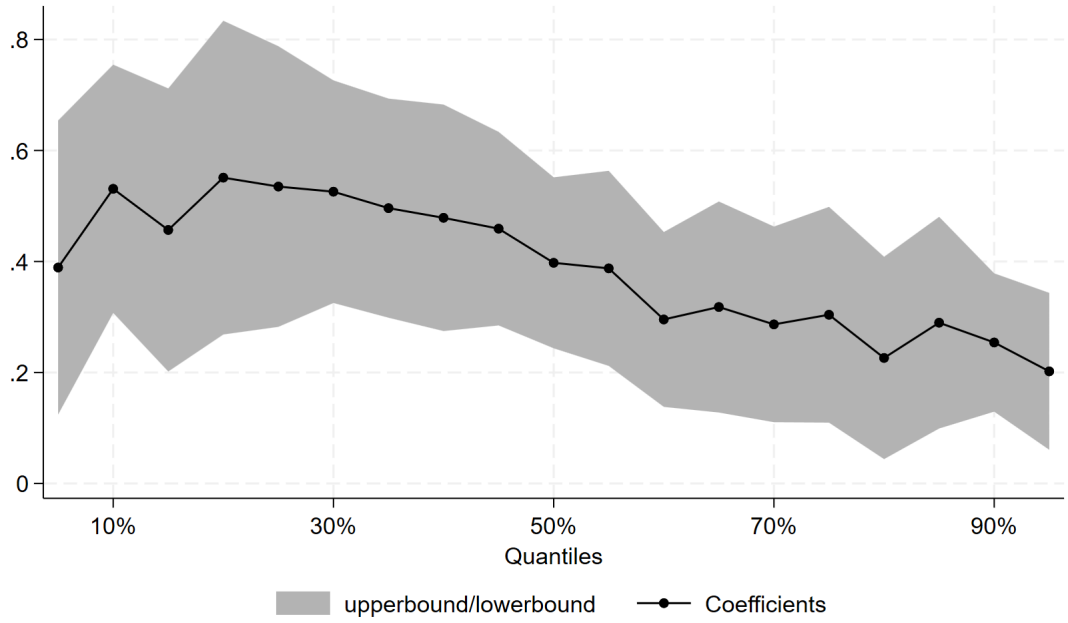

Figure 2: Non-White Revenue Premium by Conditional Quantile

Note: The figure plots quantile-regression estimates of the coefficient on the non-white movie indicator at quantiles 0.05 through 0.95 of log domestic gross box-office revenue. The shaded region reports 95 percent confidence intervals. Specifications include the same cast and movie controls as in the baseline controlled specification.

Appendix A Proofs

Throughout this appendix, we suppress the type subscript t whenever doing so does not create ambiguity. Thus, write

$$a \equiv \sigma_{\pi t}^2, \qquad b \equiv \sigma_{yt}^2, \qquad \mu \equiv \mu_t, \qquad \pi_0 \equiv \pi_{0t}.$$

Let

$$\lambda(x) \equiv \frac{\phi(x)}{1 - \Phi(x)}$$

denote the inverse Mills ratio.

A.1 Proof of Proposition 1

From equation (1) in the main text, the posterior mean is

$$E(\pi \mid y, t) = \frac{a}{a+b} y + \frac{b}{a+b} \mu.$$

The production rule $E(\pi \mid y, t) > \pi_0$ is therefore equivalent to

$$y > \bar{y}_t \equiv \pi_0 + (\pi_0 - \mu)\frac{b}{a}.$$

The comparative statics follow directly:

$$\frac{\partial \bar{y}_t}{\partial \mu} = -\frac{b}{a} < 0, \qquad \frac{\partial \bar{y}_t}{\partial \pi_0} = 1 + \frac{b}{a} > 0,$$

and

$$\frac{\partial \bar{y}_t}{\partial b} = \frac{\pi_0 - \mu}{a}.$$

Thus, $\bar{y}_t$ increases in $b = \sigma_{yt}^2$ if $\pi_0 > \mu$ and decreases in $b = \sigma_{yt}^2$ if $\pi_0 < \mu$.

A.2 Posterior updating and selected moments

The posterior mean can be written as

$$m(y) \equiv E(\pi \mid y, t) = \frac{a}{a+b} y + \frac{b}{a+b} \mu.$$

Since $y \mid t \sim N(\mu, a + b)$, it follows that

$$m(y) \mid t \sim N(\mu, s^2), \qquad s^2 \equiv \frac{a^2}{a+b} = \frac{\sigma_{\pi t}^4}{\sigma_{\pi t}^2 + \sigma_{yt}^2}.$$

The posterior variance is constant in y and equal to

$$v \equiv \mathrm{Var}(\pi \mid y, t) = \frac{ab}{a+b}.$$

By the law of total variance, $a = s^2 + v$. The production rule is equivalent to $m(y) > \pi_0$. Define

$$x \equiv \frac{\pi_0 - \mu}{s}.$$

Using the standard moments of a lower-truncated normal random variable,

$$E[Z \mid Z > x] = \lambda(x), \qquad \mathrm{Var}(Z \mid Z > x) = 1 + x\lambda(x) - \lambda(x)^2, \qquad Z \sim N(0,1),$$

we obtain

$$\begin{aligned} E(\pi \mid m(y) > \pi_0, t) &= E[m(y) \mid m(y) > \pi_0, t] = \mu + s\lambda(x), \qquad (7) \\ \mathrm{Var}(\pi \mid m(y) > \pi_0, t) &= v + \mathrm{Var}(m(y) \mid m(y) > \pi_0, t) \\ &= v + s^2\left[1 + x\lambda(x) - \lambda(x)^2\right] \\ &= a + s^2\left[x\lambda(x) - \lambda(x)^2\right]. \qquad (8) \end{aligned}$$

Equivalently, since $a = s^2(1 + b/a)$,

$$\mathrm{Var}(\pi \mid m(y) > \pi_0, t) = s^2\left[1 + \frac{b}{a} + x\lambda(x) - \lambda(x)^2\right]. \qquad (9)$$

This establishes the mean and variance formulas used in the main text.

A.3 Auxiliary facts about the inverse Mills ratio

We use three elementary properties of $\lambda(x)$. First,

$$\lambda'(x) = \lambda(x)\big(\lambda(x) - x\big) > 0. \qquad (10)$$

Second, for $x > 0$,

$$x < \lambda(x) < x + \frac{1}{x}. \qquad (11)$$

This follows, for example, from the Mills-ratio bounds in Baricz (2008, Theorem 2.3). In inverse-Mills-ratio form, that theorem implies

$$\frac{3x + \sqrt{x^2+8}}{4} < \lambda(x) < \frac{x + \sqrt{x^2+4}}{2}, \qquad x > 0,$$

and the right-hand side is strictly smaller than $x + 1/x$ for $x > 0$.

Third, define

$$k(x) \equiv x\lambda(x) - \lambda(x)^2.$$

Then

$$k'(x) < 0, \qquad \forall x \in \mathbb{R}. \tag{12}$$

To see this, differentiating $k(x)$ and using (10) gives

$$k'(x) = \lambda(x)\left[1 - x^2 + 3x\lambda(x) - 2\lambda(x)^2\right]. \tag{13}$$

The term in brackets is a concave quadratic in $\lambda(x)$ with larger root

$$r(x) \equiv \frac{3x + \sqrt{x^2 + 8}}{4}.$$

Thus it is enough to show $\lambda(x) > r(x)$ for all x.

For $x \geq 0$, this is exactly the lower bound implied by Baricz (2008, Theorem 2.3). For $x \leq -1$, $r(x) \leq 0 < \lambda(x)$, so the result is immediate. It remains only to consider $-1 < x < 0$. Let $z = -x \in (0, 1)$ and let $c = \phi(0) = 1/\sqrt{2\pi}$. Since $\phi(z) = c\exp(-z^2/2) \geq c(1 - z^2/2)$ and

$$\Phi(z) = \frac{1}{2} + \int_0^z \phi(u)du \leq \frac{1}{2} + cz,$$

we have

$$\lambda(-z) = \frac{\phi(z)}{\Phi(z)} \geq \frac{c(1 - z^2/2)}{1/2 + cz}. \tag{14}$$

Moreover,

$$\frac{c(1 - z^2/2)}{1/2 + cz} - \frac{3}{4}(1 - z) = \frac{(c - 3/8) + (3/8 - 3c/4)z + (c/4)z^2}{1/2 + cz} > 0,$$

where the inequality follows because $3/8 < c < 1/2$ and $z \in (0, 1)$. Hence

$$\lambda(-z) > \frac{3}{4}(1 - z).$$

Finally,

$$r(-z) = \frac{\sqrt{z^2 + 8} - 3z}{4} \leq \frac{3}{4}(1 - z),$$

because $z \leq 1$ implies $\sqrt{z^2 + 8} \leq 3$. Therefore $\lambda(-z) > r(-z)$ for $z \in (0, 1)$, completing the proof that $k'(x) < 0$ for all $x \in \mathbb{R}$.

A.4 Proof of Proposition 2

Using (7) and (8), write

$$E = \mu + s\lambda(x), \qquad V = a + s^2 k(x), \qquad x = \frac{\pi_0 - \mu}{s}.$$

We consider each parameter in turn, holding the other primitive parameters fixed.

Effect of μ. Because $\partial x/\partial \mu = -1/s$,

$$\frac{\partial E}{\partial \mu} = 1 - \lambda'(x) = 1 + x\lambda(x) - \lambda(x)^2.$$

The last expression is the variance of a standard normal variable conditional on exceeding x, and is therefore strictly positive. Similarly,

$$\frac{\partial V}{\partial \mu} = s^2 k'(x)\left(-\frac{1}{s}\right) = -sk'(x) > 0,$$

where the inequality follows from (12). Hence both the conditional mean and the conditional variance increase in μ.

Effect of π_0. Because $\partial x/\partial \pi_0 = 1/s$,

$$\frac{\partial E}{\partial \pi_0} = \lambda'(x) > 0,$$

by (10). Similarly,

$$\frac{\partial V}{\partial \pi_0} = s^2 k'(x)\left(\frac{1}{s}\right) = sk'(x) < 0,$$

again by (12). Hence the conditional mean increases in π_0, while the conditional variance decreases in π_0.

Effect of $b = \sigma_{yt}^2$. The parameter b affects the selected moments through s, where

$$s = \frac{a}{\sqrt{a+b}}, \qquad \frac{\partial s}{\partial b} = -\frac{s}{2(a+b)} < 0.$$

Differentiating the conditional mean with respect to s gives

$$\begin{aligned}\frac{\partial E}{\partial s} &= \lambda(x) - x\lambda'(x) \\ &= \lambda(x)\left[1 + x^2 - x\lambda(x)\right] > 0.\end{aligned}$$

If $x \leq 0$, the inequality is immediate. If $x > 0$, it follows from (11), which implies $x\lambda(x) < x^2 + 1$. Since $\partial s/\partial b < 0$, we obtain

$$\frac{\partial E}{\partial b} < 0.$$

For the conditional variance, differentiating $V = a + s^2 k(x)$ with respect to s gives

$$\begin{aligned}\frac{\partial V}{\partial s} &= s\left[2k(x) - xk'(x)\right] \\ &= s\lambda(x)\left[x - 2\lambda(x)\right]\left[1 + x^2 - x\lambda(x)\right] < 0.\end{aligned}$$

The final factor is positive by the same argument used for $\partial E/\partial s$. The factor $x - 2\lambda(x)$ is negative: this is immediate if $x \leq 0$, and follows from $\lambda(x) > x$ if $x > 0$. Therefore $\partial V/\partial s < 0$. Since $\partial s/\partial b < 0$, it follows that

$$\frac{\partial V}{\partial b} > 0.$$

Thus, as the signal variance σ_{yt}^2 increases, the conditional mean decreases and the conditional variance increases.

Appendix B Alternative distributional assumptions

In this appendix, we explore the robustness of our results to alternative distributional assumptions.

B.1 Beta-Binomial distribution

B.1.1 Setup

We first consider the case where the producer cares only about a binary criterion, whether the movie will be a "hit" or not.

Assume the object of interest is p, the probability that the movie is a hit. The prior distribution of p is Beta with parameters α and β:

$$p \sim Beta(\alpha, \beta)$$

Therefore:

$$\begin{aligned} f(p) &\propto p^{\alpha-1}(1-p)^{\beta-1} \\ E(p) &= \frac{\alpha}{\alpha+\beta} \\ V(p) &= \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2} \end{aligned}$$

Producers observe a signal y, which, conditional on the true p, is distributed binomial with parameters n and p. We can think of this as the producer consulting with n critics, and each one independently assessing whether the movie will be a hit or not, with probability p.

$$f(y|p) = \binom{n}{y} p^y (1-p)^{(n-y)}$$

It follows that the posterior density of p given y is

$$\begin{aligned} f(p|y) &\propto p^{\alpha-1}(1-p)^{\beta-1} p^y (1-p)^{n-y} \\ &\Rightarrow p|y \sim Beta(\alpha+y, \beta+n-y) \end{aligned}$$

Therefore,

$$E(p|y) = \frac{\alpha+y}{\alpha+y+\beta+n-y} = \frac{\alpha+y}{\alpha+\beta+n}$$

The producer will produce the movie if $E(p|y) > p_0$, for some predetermined p_0. Therefore, the signal threshold for production $\bar{y}$ is:

$$\bar{y} \equiv p_0(\alpha + \beta + n) - \alpha.$$

It is convenient to use a reparametrization, letting $\kappa = \alpha + \beta$. If $\alpha, \beta > 1$, then κ captures the spread of the distribution: for a given α, a higher value of κ means that the distribution is more concentrated, i.e., the prior is more informative.[1]

Let $s = y/n$ be the success rate of the signal. Expressing the signal threshold in terms of s, the movie is produced if and only if

$$s > \bar{s} \equiv p_0 + \frac{\kappa}{n}(p_0 - \frac{\alpha}{\kappa})$$

B.1.2 Comparative statics for production

(1) **Customer discrimination**. We interpret customer discrimination against non-white movies as $\alpha_b < \alpha_w$. That is, non white movies have a lower prior probability of being a hit. The signal threshold is decreasing in α:

$$\frac{\partial \bar{s}}{\partial \alpha} < 0.$$

Therefore, under customer discrimination, the signal threshold for non-white movies is higher than the signal threshold for white movies.

(2) **Taste-based discrimination**. We interpret taste-based discrimination against non-white movies as $p_{0b} > p_{0w}$. That is, non-white movies are held to a higher standard, and are produced only if the posterior probability of the movie being a hit exceeds a higher threshold. The signal threshold is increasing in p_0:

$$\frac{\partial \bar{s}}{\partial p_0} > 0.$$

Therefore, under taste-based discrimination, the signal threshold for non-white movies is higher than the signal threshold for white movies.

[1] Under this parameterization, the mean of the prior distribution is $E(p) = \frac{\alpha}{\kappa}$ and the variance is $V(p) = \frac{\alpha(\kappa-\alpha)}{\kappa^2(\kappa-1)}$. For $\alpha, \beta > 1$, the variance is strictly decreasing in κ.

(3) **Statistical discrimination**. Finally, we interpret statistical discrimination as $n_b < n_w$. That is, the signal for non-white movies being is less informative than that for white movies. The derivative of the signal threshold with respect to n is:

$$\frac{\partial \bar{s}}{\partial n} = -\frac{\kappa}{n^2}(p_0 - \frac{\alpha}{\kappa})$$

The sign of this derivative depends on $p_0 - \alpha/\kappa$. Under this parametrization, α/κ is the prior mean of p. In other words, we have the same qualitative result as in the log-normal model presented in the main text.

(i) If $p_0 > \alpha/\kappa$, (i.e., the producer wants to produce only movies with a very high probability of being a hit),

$$\frac{\partial \bar{s}}{\partial n} < 0;$$

that is, a less precise signal (lower n) raises the signal threshold. The signal threshold for non-white movies is higher.

(ii) If $p_0 < \alpha/\kappa$, (i.e. the producer wants to weed out the very low quality movies),

$$\frac{\partial \bar{s}}{\partial n} > 0;$$

now, a less precise signal (lower n) *lowers* the signal threshold. One can be a bit more tolerant of a bad signal for non-white movies, because it is difficult to say, based on the signal alone, whether the movie is really bad.

It is easy to see that the comparative statics with respect to the precision of the signal mirrors exactly what we had in the normal-normal case.

B.1.3 Comparative statics for the observed success rate, conditional on production: simulations

We only observe whether a movie is a hit, conditional on production. Therefore, as in the analysis in the main text, we need to characterize the the posterior distribution of p conditional on $s > \bar{s}$, and derive its comparative statics with respect to p_0, α and n. While it is not possible to derive an analytical solution for the comparative statics, we can proceed

by simulation. Specifically, for each set of parameter values, we draw a sample of L movies, apply the production decision rule, and report the mean and standard deviation of the posterior distribution p conditional on production. The results are presented in Table B.1.

Table B.1: Simulation results: Beta-binomial distribution

A: Taste based discrimination: $p_0 \uparrow$ for Non-white movies
Fixed $\alpha = 4$, $\kappa = 8$, $n = 5$

p_0	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6	Trend
mean	0.500	0.500	0.500	0.500	0.514	0.545	0.545	0.587	0.638	0.638	$\uparrow$
std	0.167	0.167	0.167	0.167	0.160	0.151	0.151	0.142	0.133	0.133	$\downarrow$

B: Customer discrimination: $\alpha \downarrow$ for Non-white movies
Fixed $p_0 = 0.5$, $\kappa = 8$, $n = 5$

α	6	5.5	5	4.5	4	3.5	3	2.5	2	1.5	
mean	0.752	0.692	0.650	0.597	0.587	0.542	0.555	0.515	0.538	0.497	$\downarrow$
std	0.142	0.151	0.148	0.151	0.142	0.143	0.136	0.137	0.135	0.135	$\downarrow$

C1: Statistical discrimination, $p_0 > \alpha/\kappa$: $n \downarrow$ for Non-white movies
Fixed $p_0 = 0.6$, $\kappa = 8$, $\alpha = 4$

n	11	10	9	8	7	6	5	4	3	2	
mean	0.673	0.656	0.679	0.659	0.638	0.662	0.638	0.666	0.637	0.600	$\downarrow$
std	0.115	0.119	0.117	0.122	0.128	0.126	0.133	0.131	0.139	0.148	$\uparrow$

C2: Statistical discrimination, $p_0 < \alpha/\kappa$: $n \downarrow$ for Non-white movies
Fixed $p_0 = 0.4$, $\kappa = 8$, $\alpha = 4$

n	11	10	9	8	7	6	5	4	3	2	
mean	0.549	0.557	0.540	0.548	0.528	0.535	0.545	0.520	0.527	0.500	$\downarrow$
std	0.145	0.143	0.149	0.148	0.154	0.153	0.151	0.159	0.158	0.167	$\uparrow$

Legend: simulated data with sample size $L = 10^6$, using R with seed 123. Mean, Std: sample average and standard deviation of the posterior distribution of $p|s, s > \bar{s}$ from the simulation.

Each panel in the table presents a different comparative statics exercise. For example, in Panel A, to examine the role of taste-based discrimination, we fix the values of α, κ and n, and study what happens to the mean and standard deviation of the posterior distribution of p as we increase the value of p_0. The results in Panel A show that as taste-based discrimination increases, the posterior expected value of p increases, and the posterior standard deviation decreases. These results match the predictions in the normal-normal model,

derived analytically.

In Panel B we look at the effect of increasing customer discrimination increases. As in the normal model, both the expected value and the standard deviation increase as the value of α decreases.[2]

In Panels C1 and C2 we study the effect of statistical discrimination, distinguishing between the case in which the producer only wants to produce very high quality movies ($p_0 > \alpha/\kappa$) so that the signal threshold decreases in n (case 3.i in Section B.1.2); and the one in which the producer wants to weed out very low quality movies ($p_0 < \alpha/\kappa$) so that the signal threshold increases in n. We see that in both cases the mean of p decreases and the standard deviation increases as the extent of statistical discrimination increases (the signal becomes less prescise, or n decreases). Again, the pattern of comparative statics results mirrors exactly what we obtained in the normal-normal case (Section 2 in the main text).

We conclude that all of the main predictions of the theoretical model based on the normal-normal case in the main text remain identical under the beta-binomial model.

B.2 Pareto-Normal distribution

B.2.1 Setup

We now return to the case considered in the main text, where the producer cares about (log) revenue, but we now depart from the normal-normal model. Specifically, we assume that ex-ante revenue $\tilde{\pi}$ (in dollars) follows a Pareto distribution:

$$\tilde{\pi} \sim Pareto(x_m, a)$$

where x_m is the minimum, and a is the shape parameter. The CDF is:

$$F(\tilde{\pi}) = \begin{cases} 1 - (\frac{x_m}{\tilde{\pi}})^a, & if\ \tilde{\pi} \geq x_m \\ 0, & if\ \tilde{\pi} < x_m \end{cases}$$

Then, log revenue $\pi \equiv \log(\tilde{\pi})$ has a shifted exponential distribution: $\pi \sim Exp(a) + log(x_m)$, or, equivalently, $\log(\frac{\tilde{\pi}}{x_m}) \sim Exp(a)$.

[2] In each panel, as we move in the table from left to right, we *increase* the extent of discrimination. In the case of customer discrimination (Panel B) and statistical discrimination (Panel C), an increase in discrimination implies a decrease in the parameter of interest.

Therefore, the pdf of π is:

$$f(\pi) = \begin{cases} a * exp(-a(\pi - log(x_m))), \ if \ \pi \geq log(x_m) \\ 0, \ if \ \pi < log(x_m) \end{cases}$$

Producers observe a signal y, which, conditional on the true π is distributed $N(\pi, \sigma_y^2)$.

$$f(y|\pi) = \frac{1}{\sigma_y \sqrt{2\pi}} exp(-\frac{1}{2}(\frac{y-\pi}{\sigma_y})^2)$$

It follows that the posterior distribution of π given y is

$$f(\pi|y) \propto exp(-a(\pi - \log(x_m)) - \frac{1}{2}(\frac{y-\pi}{\sigma_y})^2), \text{ if } \pi \geq \log(x_m)$$

When $\pi \geq log(x_m)$:

$$\begin{aligned} f(\pi|y) &\propto exp(-a(\pi - log(x_m)) - \frac{1}{2}(\frac{y-\pi}{\sigma_y})^2) \\ &= exp(-\frac{1}{2\sigma_y^2}(\pi^2 - 2y\pi + y^2 + 2\sigma_y^2 a\pi - 2\sigma_y^2 a \log(x_m))) \\ &= exp(-\frac{1}{2\sigma_y^2}((\pi - (y - a\sigma_y^2))^2 - a^2\sigma_y^4 + 2ya\sigma_y^2 - 2\sigma_y^2 a \log(x_m))) \\ &= exp(-\frac{1}{2\sigma_y^2}((\pi - (y - a\sigma_y^2))^2)) \times exp(-a(y - \frac{a\sigma_y^2}{2} - \log(x_m))) \end{aligned}$$

Given y, the second term is constant. Therefore, putting everything together, we have that

$$f(\pi|y) \propto exp(-\frac{1}{2\sigma_y^2}((\pi - (y - a\sigma_y^2))^2)), \text{ for } \pi > \log(x_m).$$

This implies that the posterior distribution of π given the signal y is a truncated normal derived from a normal distribution with mean $y - a\sigma_y^2$, variance σ_y^2 and lower truncation point $\log(x_m)$. The posterior mean is therefore

$$E(\pi|y) = (y - a\sigma_y^2) + \sigma_y \frac{\phi(\log(x_m))}{1 - \Phi(\log(x_m))}.$$

The producer will produce the movie if $E(\pi|y) > \pi_0$, for some predetermined π_0. Therefore, the signal threshold for production $\bar{y}$ is:

$$\bar{y} \equiv \pi_0 + a\sigma_y^2 - \sigma_y \frac{\phi(\log(x_m))}{1 - \Phi(\log(x_m))}.$$

B.2.2 Comparative statics for production

(1) **Customer discrimination**: The expectation of an exponential distribution with parameter a is $1/a$. Therefore, we interpret customer discrimination against non-white movies as $a_b > a_w$. The signal threshold is increasing in a:

$$\frac{\partial \bar{y}}{\partial a} > 0$$

As in the normal-normal case, the signal threshold for non-white movies is higher than the signal threshold for white movies.

(2) **Taste-based discrimination**: We interpret taste-based discrimination against non-white movies as $\pi_{0b} > \pi_{0w}$ – non-white movies are held to a higher standard and are produced only if the posterior mean exceeds a threshold that is higher than that set for white movies. The signal threshold increases in π_0:

$$\frac{\partial \bar{y}}{\partial \pi_0} > 0.$$

Therefore, under taste-based discrimination, the signal threshold for non-white movies is higher than that for white movies. This result mirrors that of the normal-normal case.

(3) **Statistical discrimination**: We interpret statistical discrimination as $\sigma_{yb} > \sigma_{yw}$, i.e., the signal for non-white movies is less precise. The derivative of the signal threshold with respect to σ_y is:

$$\frac{\partial \bar{y}}{\partial \sigma_y} = 2a\sigma_y - \frac{\phi(\log(x_m))}{1 - \Phi(\log(x_m))}$$

The sign of this derivative depends on the magnitude of σ_y.

(i) If $\sigma_y > \frac{\phi(\log(x_m))}{2a(1-\Phi(\log(x_m)))}$, (i.e., the movie has a high variance on the potential outcome),

$$\frac{\partial \bar{y}}{\partial \sigma_y} > 0;$$

that is, a less precise signal (lager σ_y) raises the signal threshold. The signal threshold for non-white movies is higher.

(ii) If $\sigma_y < \frac{\phi(\log(x_m))}{2a(1-\Phi(\log(x_m)))}$, (i.e. the signal has good information on the movie revenue),

$$\frac{\partial \bar{y}}{\partial \sigma_y} < 0;$$

now, a less precise signal (larger σ_y) *lowers* the signal threshold. One can be a bit more tolerant of a bad signal for non-white movies, because it is difficult to say, based on the signal alone, whether the movie is really bad.

Although the signal threshold for non-white movies could still have been either higher or lower than that for white movies, the result does not depend on whether producers only want to produce very high quality movies, or they just want to weed out very low quality movies (i.e., it does not depend on whether the revenue threshold π_0 is above or below the prior mean of π, which is in contrast to the normal-normal model).

B.2.3 Comparative statics for observed revenue, conditional on production: simulations

As in Section B.1.3 we use simulations to characterize the posterior distribution of observed revenue, conditional on production. We choose the parameters of the Pareto distribution to roughly mimic the observed distribution of revenue in our sample. Therefore, in all simulations, we set $x_m = 20$ (roughly equal to the minimum observed revenue in our sample) and set the baseline value of a at 0.2.[3] In this case, $\frac{\phi(\log(x_m))}{2a(1-\Phi(\log(x_m)))} \approx 8.2$, so we choose $\sigma_y = 8$ as the baseline. The prior mean is $1/a + log(x_m) \approx 8$, so we choose $\pi_0 = 8$ as the baseline. The results of the simulations are presented in Table B.2.

[3]The maximum likelihood estimate of a in our full sample is 0.09; 0.18 if one excludes the bottom 10% of the distribution; and 0.22 if one excludes the bottom 25%. We chose a slightly higher value of a as the baseline in our simulations because lower values of a will result in an implausibly large fraction of movies with explosive revenues (the mean of a Pareto distribution with $a < 1$ is infinite).

Table B.2: Simulation results: Pareto distribution

A: Taste based discrimination: $\pi_0 \uparrow$ for Non-white movies Fixed $a = 0.2$, $\sigma_y = 8$, $x_m = 20$											Trend
π_0	5	7	9	11	13	15	17	19	21	23	
mean	8.091	8.164	8.252	8.374	8.548	8.767	9.041	9.381	9.800	10.309	$\uparrow$
std	5.041	5.081	5.108	5.173	5.270	5.362	5.521	5.686	5.888	6.137	$\uparrow$

Fixed $a = 0.5$, $\sigma_y = 8$, $x_m = 20$											Trend
π_0	5	7	9	11	13	15	17	19	21	23	
mean	5.697	5.858	6.020	6.215	6.474	6.747	7.105	7.373	7.844	8.518	$\uparrow$
std	2.545	2.670	2.781	2.951	3.141	3.327	3.621	3.731	4.094	4.364	$\uparrow$

B: Customer discrimination: $a \uparrow$ for Non-white movies Fixed $\pi_0 = 8$, $\sigma_y = 8$, $x_m = 20$											
a	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55	
mean	13.056	9.786	8.198	7.315	6.769	6.428	6.201	6.040	5.944	5.837	$\downarrow$
std	10.012	6.711	5.083	4.168	3.587	3.242	3.008	2.833	2.742	2.628	$\downarrow$

C: Statistical discrimination: $\sigma_y \uparrow$ for Non-white movies Fixed $a = 0.2$, $\pi_0 = 8$, $x_m = 20$											
σ_y	1	2	3	4	5	6	7	8	9	10	
mean	9.736	8.344	8.156	8.112	8.114	8.139	8.156	8.203	8.254	8.324	?
std	5.082	5.072	5.045	5.026	5.027	5.061	5.078	5.087	5.128	5.168	?

Fixed $a = 0.5$, $\pi_0 = 8$, $x_m = 20$											
σ_y	1	2	3	4	5	6	7	8	9	10	
mean	6.749	5.469	5.327	5.361	5.447	5.585	5.750	5.930	6.151	6.360	?
std	2.216	2.167	2.152	2.203	2.278	2.412	2.565	2.725	2.933	3.118	?

Legend: simulated data with sample size $L = 10^6$, using R with seed 123. Mean, Std: sample average and standard deviation of the posterior distribution of $\pi|y, y > \bar{y}$ from the simulation.

In Panel A we examine the role of taste-based discrimination. We fix the values of a and σ_y and study what happens to the posterior mean and standard deviation of (log) revenue conditional on production as we increase π_0. The posterior mean increases (as in the normal-normal case), while the standard deviation also increases which is different from the normal-normal case.

In Panel B we look at the effect of increasing customer discrimination by letting a in-

crease. Both the posterior mean and standard deviation decrease as the extent of customer discrimination increases, matching the predictions of the normal-normal model.

Finally, in panel C we vary the extent of statistical discrimination by letting σ_y increase, i.e., making the signal less precise. Here the results stand in contrast with those of the normal-normal model: as the signal becomes less precise, both the posterior mean of log revenue and the posterior standard deviation have a U-shaped pattern, first decreasing and then increasing in the extent of noise in the signal.

The comparative statics in the Pareto model are not identical to those in the normal-normal model presented in the main text. However, the simulations show that both the mean and the variance of log box-office revenue always move in the same direction as we change the discrimination parameter, under all three forms of discrimination. This is in contrast with the observed patterns in the data, where the mean of log revenue is higher for non-white movies, but the variance of log revenue is smaller (see Tables 6 in the text).

B.3 Studio's Choice of Race

B.3.1 Setup

The studio can choose to cast a single, race-neutral script either as a white or non-white movie (the "race-of-version" choice).

We begin with one race-neutral script. The studio may produce it as a white version ($t = w$), a non-white version ($t = b$), or not at all.

Step 1: Script and potential revenues. If cast as version $t \in \{w, b\}$, the log-revenue follows:

$$\pi_t \sim N\big(\mu_t, \sigma^2_{\pi t}\big), \qquad t \in \{w, b\}.$$

Step 2: Signals and posteriors. The studio observes two independent signals:

$$y_t \mid \pi_t \sim N\big(\pi_t, \sigma^2_{yt}\big), \quad t \in \{w, b\},$$

and computes the posterior expectation:

$$E[\pi_t \mid y_t] = \alpha_t\, y_t + (1 - \alpha_t)\, \mu_t, \qquad \alpha_t = \frac{\sigma^2_{\pi t}}{\sigma^2_{\pi t} + \sigma^2_{yt}}.$$

Step 3: Production decision. Define the net return:

$$R_t \equiv E[\pi_t \mid y_t] - \pi_{0t}.$$

The studio chooses the option that maximizes net return among producing a white version, a non-white version, or not producing at all:

$$\begin{cases} \text{no production}, & 0 = \max\{R_w, R_b, 0\}, \\ \text{produce white}, & R_w = \max\{R_w, R_b, 0\}, \\ \text{produce non-white}, & R_b = \max\{R_w, R_b, 0\}. \end{cases}$$

B.3.2 Comparative statics for observed revenue, conditional on production: simulations

As in Section B.1.3, we use simulations to characterize the posterior distribution of observed revenue, conditional on production.

Denote DE as the difference in expected log-revenue between produced white and non-white movies:

$$DE = E\big(\pi_w \mid R_w > R_b,\ R_w > 0\big) - E\big(\pi_b \mid R_b > R_w,\ R_b > 0\big).$$

Similarly, denote $DVar$ as the difference in the variance of log-revenue:

$$DVar = \mathrm{Var}\big(\pi_w \mid R_w > R_b,\ R_w > 0\big) - \mathrm{Var}\big(\pi_b \mid R_b > R_w,\ R_b > 0\big).$$

Rather than investigating the trends of mean and variance with respect to parameters, we focus on the sign of DE and $DVar$, in line with the theoretical predictions in the main text.

Table B.3: Simulation results: Endogenous Choice of Race

A: Taste-based discrimination: $\pi_{0b} > \pi_{0w}$ Fixed $\mu = 6$, $\sigma_y = 5$, $\sigma_\pi = 3$, $\pi_{0b} = 9$										
π_{0w}	4	4.5	5	5.5	6	6.5	7	7.5	8	8.5
DE	-3.58	-3.35	-3.08	-2.77	-2.44	-2.08	-1.69	-1.29	-0.87	-0.44
$DVar$	1.30	1.11	0.91	0.72	0.56	0.42	0.31	0.21	0.14	0.07

B: Customer discrimination: $\mu_b < \mu_w$ Fixed $\pi_0 = 9$, $\sigma_y = 5$, $\sigma_\pi = 3$, $\mu_b = 6$										
μ_w	6.5	7	7.5	8	8.5	9	9.5	10	10.5	11
DE	0.06	0.13	0.21	0.31	0.42	0.56	0.73	0.92	1.16	1.42
$DVar$	0.07	0.14	0.21	0.31	0.42	0.56	0.72	0.91	1.11	1.30

C: Statistical discrimination: $\sigma_{yw} < \sigma_{yb}$ Fixed $\pi_0 = 9$, $\mu = 6$, $\sigma_\pi = 3$, $\sigma_{yb} = 5$										
σ_{yw}	2	2.3	2.6	2.9	3.2	3.5	3.8	4.1	4.4	4.7
DE	0.62	0.54	0.47	0.39	0.32	0.25	0.19	0.14	0.09	0.05
$DVar$	-3.02	-2.59	-2.18	-1.80	-1.45	-1.13	-0.85	-0.60	-0.37	-0.16

Legend: Simulated data with sample size $L = 10^7$, using R with seed 123. DE and $DVar$ report differences in the average and variance of log revenue across produced white and non-white movies.

In Panel A, we examine the role of taste-based discrimination. Across all scenarios, white movies tend to have lower mean and higher variance in revenue compared to non-white movies, consistent with the predictions in the main text. Furthermore, as π_{0w} increases, DE increases (becomes less negative), and $DVar$ decreases.

In Panel B, we explore customer discrimination by increasing μ_w while holding μ_b fixed. Both the expected revenue and variance for white movies increase relative to non-white movies, and both DE and $DVar$ increase with μ_w. These patterns align with the comparative statics of the normal-normal model.

In Panel C, we consider statistical discrimination by increasing signal precision for white movies, i.e., decreasing σ_{yw}. As the white signal becomes more informative, DE remains positive but decreases, while $DVar$ remains negative but increases. These trends are consistent with the model without endogenous choice of race.

Overall, the comparative statics under endogenous race-of-version choice are qualitatively identical to those in the baseline normal-normal model presented in the main text. Thus, the predictions are robust to this extension.

B.4 A Model with Market Segmentation

We consider two producers, $s \in \{W, B\}$, each with potentially different parameters for white ($t = w$) and non-white ($t = b$) movies. Ex ante, the log-box-office revenue for a movie of type t produced by s is

$$\pi_{ts} \sim N\big(\mu_{ts},\, \sigma^2_{\pi\, ts}\big), \qquad t \in \{w, b\},\ s \in \{W, B\}.$$

Upon observing two independent signals,

$$y_{ts} \mid \pi_{ts} \sim N\big(\pi_{ts},\, \sigma^2_{y\, ts}\big),$$

the producer forms the posterior mean

$$E[\pi_{ts} \mid y_{ts}] = \frac{\sigma^2_{\pi\, ts}}{\sigma^2_{\pi\, ts} + \sigma^2_{y\, ts}}\, y_{ts} + \frac{\sigma^2_{y\, ts}}{\sigma^2_{\pi\, ts} + \sigma^2_{y\, ts}}\, \mu_{ts}. \tag{15}$$

A movie is produced if and only if

$$E[\pi_{ts} \mid y_{ts}] > \pi_{0ts}.$$

Hence, the probability of production is

$$P\big(E[\pi_{ts} \mid y_{ts}] > \pi_{0ts}\big) = 1 - \Phi\Big(\big(\pi_{0ts} - \mu_{ts}\big)\, \frac{\sqrt{\sigma^2_{\pi\, ts} + \sigma^2_{y\, ts}}}{\sigma^2_{\pi\, ts}}\Big).$$

We assume script-writers choose the producer s that maximizes this probability.

Case 1: Different means (μ_{ts}). Without loss of generality, let $\mu_{wW} > \mu_{wB}$ and $\mu_{bW} < \mu_{bB}$. Then each producer specializes:

$$A \text{ produces only white movies}, \quad L \text{ produces only non-white movies}.$$

Moreover, if $\mu_{wW} > \mu_{bB}$ the sorting pattern mirrors customer-discrimination results.

Case 2: Comparative advantage in signals ($\sigma^2_{y\, ts}$). Assume $\sigma^2_{y\, wW} < \sigma^2_{y\, wB}$ and $\sigma^2_{y\, bW} > \sigma^2_{y\, bB}$. Again, full sorting obtains, with W handling whites and B handling non-whites; and if $\sigma^2_{y\, wW} < \sigma^2_{y\, bB}$, the pattern replicates statistical-discrimination results.

Case 3: Different production cutoffs (π_{0ts}). Let $\pi_{0\,wW} < \pi_{0\,wB}$ and $\pi_{0\,bW} > \pi_{0\,bB}$. Full sorting occurs as before, and if $\pi_{0\,wW} < \pi_{0\,bB}$, the outcome matches taste-based discrimination.

In all three cases, the model predicts complete market segmentation. Empirically, however, the between-studio variation in non-white shares is small (less than 1% among Big 6 studios; approximately 15% for non-Big 6 studios). Thus, while sorting may contribute to the racial gap, the persistent difference even within the Big 6 indicates that factors beyond pure sorting are at work.

Table B.4: Variance Decomposition of White/Non-White Movie Sorting

	Full sample	Big 6	Non-Big 6
Movies	6,943	2,455	4,488
Distributors	498	6	492
Mean `morethan2` rate (%)	8.5	9.7	7.8
Total variance	0.078	0.088	0.072
Between-distributor variance	0.008	0.000	0.011
Within-distributor variance	0.070	0.088	0.061
Share between distributors (%)	9.7	0.3	15.8
Share within distributors (%)	90.3	99.7	84.2

Notes: The table decomposes the total variance of the non-white dummy into between- and within-distributor components. The between-distributor share equals the R-squared from a regression of the non-white dummy on distributor fixed effects within the relevant sample.

B.5 Taste-Based Discrimination and Inaccurate Beliefs: An Alternative Explanation

We consider an alternative explanation in which studios hold inaccurate beliefs about the revenue-generating potential of different types of movies. Specifically, studios believe the data-generating process follows:

Step 1: Script and potential revenues. If a script is cast as version $t \in \{w, b\}$, the (log) box-office revenue is believed to follow:

$$\pi_t \sim N(\hat{\mu}_t, \hat{\sigma}^2_{\pi t}).$$

Step 2: Signals and posterior beliefs. The producer observes a signal:

$$y_t \mid \pi_t \sim N(\pi_t, \hat{\sigma}^2_{yt}),$$

and forms a posterior expectation:

$$\hat{E}[\pi_t \mid y_t] = \hat{\alpha}_t \, y_t + (1 - \hat{\alpha}_t) \, \hat{\mu}_t, \quad \text{where} \quad \hat{\alpha}_t = \frac{\hat{\sigma}^2_{\pi t}}{\hat{\sigma}^2_{\pi t} + \hat{\sigma}^2_{yt}}.$$

Step 3: Production decision. A movie is produced if:

$$\hat{E}[\pi_t \mid y_t] > \pi_{0t}.$$

As in Bohren et al. (2023), when beliefs are accurate—that is, when $\hat{\mu}_t = \mu_t$ and $\hat{\alpha}_t = \alpha_t$—this framework coincides with the baseline model. However, in general, the actual posterior mean is:

$$E[\pi_t \mid y_t] = \hat{E}[\pi_t \mid y_t] + (\alpha_t - \hat{\alpha}_t)(y_t - \hat{\mu}_t) + (1 - \alpha_t)(\mu_t - \hat{\mu}_t).$$

Thus, the threshold π_{0t} applied to the perceived posterior $\hat{E}[\pi_t \mid y_t]$ corresponds to an effective threshold for $E[\pi_t \mid y_t]$ equal to:

$$\pi_{0t} + (\alpha_t - \hat{\alpha}_t)(y_t - \hat{\mu}_t) + (1 - \alpha_t)(\mu_t - \hat{\mu}_t).$$

If studios systematically underestimate non-white movies, this misbelief can arise in two ways:

Case 1: Underestimating the mean ($\hat{\mu}_b < \mu_b$). In this case, the industry undervalues the baseline potential of non-white movies. Since $(1-\alpha_b)(\mu_b-\hat{\mu}_b) > 0$, the effective threshold for $E[\pi_b \mid y_b]$ exceeds that for $E[\pi_w \mid y_w]$, even if $\pi_{0b} = \pi_{0w}$. This leads to underproduction of non-white movies, mirroring the outcomes associated with taste-based discrimination.

Case 2: Underestimating signal precision ($\hat{\sigma}_{yb} > \sigma_{yb}$). Studios may also underestimate the informativeness of signals associated with non-white movies. As they gain experience, the true signal variance σ^2_{yb} may be lower than believed, implying $\alpha_b > \hat{\alpha}_b$.

Suppose studios only greenlight projects when signals are sufficiently positive, i.e., when $\hat{E}[\pi_t \mid y_t] > \pi_{0t} > \hat{\mu}_t$. Then production occurs only if $y_t > \mu_t$. Under this condition,

$$(\alpha_b - \hat{\alpha}_b)(y_b - \hat{\mu}_b) > 0,$$

so again, the effective threshold for $E[\pi_b \mid y_b]$ is higher than that for $E[\pi_w \mid y_w]$, despite equal nominal cutoffs.

In both cases—and especially when both misbeliefs are present—the studio applies a systematically higher threshold to non-white movies. Thus, even absent explicit animus, inaccurate beliefs can generate outcomes observationally equivalent to taste-based discrimination.

Appendix C Machine learning algorithm for facial classification

For performers that were not unambiguously classified by the human raters, we applied the facial classification algorithm proposed by Anwar and Islam (2017)[1]. The algorithm is based on a machine learning architecture that combines a convolutional neural network (CNN) and support vector machine (SVM), described below.

Step 1. We started with a sample of more than 7000 motion pictures released in the United States between 1997 and 2017, taken from Opus Data,[2] a private company that collects data on the industry. For each movie, we took the names of the four top-billed performers. We then scraped and cropped the image appearing on each performer's page on the popular website IMDB.[3]

Step 2. We used the Visual Geometry Group[4] (V.G.G.) technique to locate the actor's face on each picture. The output of this step is a vector of information extracted from each image, or a "feature vector."

Step 3. We repeated step 2 on our training data set, the Chicago Face Database (CFD).[5] This database is intended for use in scientific research. It is useful as it contains images of 597 unique individuals (both male and female) who self-identify as White, Black, Asian, or Latino/a.

Step 4. We used CFD to train our algorithm using the Support Vector Machine (SVM) approach.[6] Intuitively, the purpose of SVM is to find the "best separation line," meaning the hyper-plane that correctly separates white from non-white performers when such performers are located in a multi-dimensional space through their feature vectors.

Step 5. We applied our trained algorithm to the pictures obtained from Steps 1 and 2. We validated our algorithm on a subsample of actors for which we manually coded the racial groups and obtained a success rate of 95%. A few examples of the outcomes of our

[1] Link: https://arxiv.org/ftp/arxiv/papers/1709/1709.07429.pdf.
[2] www.opusdata.com
[3] www.imdb.com
[4] See for reference https://www.robots.ox.ac.uk/~vgg/.
[5] The CFD is available at https://www.chicagofaces.org/.
[6] See for reference https://scikit-learn.org/stable/modules/svm.html.

classification algorithm are presented in Figure C.1.

Figure C.1: Output of facial classification

Appendix D Other Figures and Tables

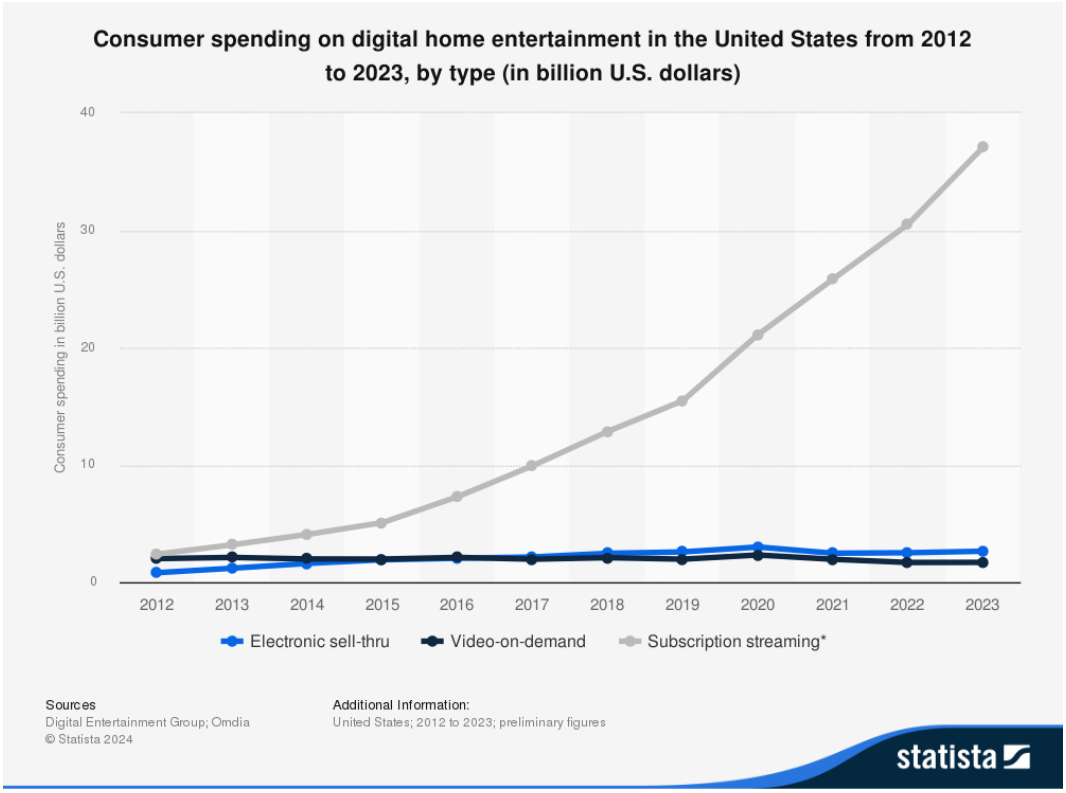

Figure D.1: Trend in consumer spending on digital home entertainment, by category
Source: Statista (link.)

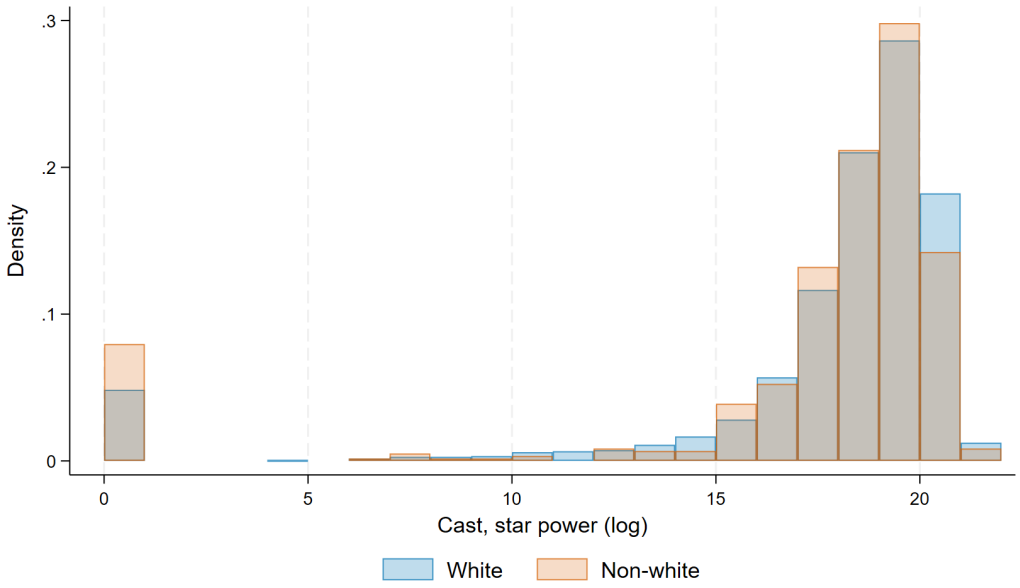

Figure D.2: Distribution of log cast star power across racial movie types
Note: A one-sided t-test on the means calculated off the full distributions fails to reject the null hypothesis that the white average is larger than the non-white average. Excluding the left tail of the distributions (i.e., truncating the distributions from below at 5) makes the means non significantly different.

Table D.1: Descriptive statistics by distributor

Panel A: Non Big-6

Variable	N	Mean	SD	Min	Max
Run time (minutes)	3,940	101.51	17.69	38	600
IMDb score	2,753	6.21	0.95	1.5	8.9
Metacritic score	2,753	51.78	17.29	1	100
Ln(Cost)	1,736	16.04	1.47	7.00	20.63
Ln(Gross revenue)	4,488	12.80	3.09	2.96	19.94
Average age of billed performers	4,488	42.38	10.78	10	99
Gross revenue (in Millions of 2005 Dollars)	4,488	9.27	28.18	1.94×10^{-5}	457.65
Star power (in Millions)	4,488	211.76	263.47	0.00	2,027.67
Ln(Star power)	4,488	16.89	5.06	0	21.43
At least two non-white	4,488	0.08	0.27	0	1
At least one non-white	4,488	0.35	0.43	0	1
Share non-white performers	4,488	0.12	0.25	0	1

Panel B: Big-6

Variable	N	Mean	SD	Min	Max
Run time (minutes)	2,320	108.23	18.92	39	219
IMDb score	2,162	6.30	0.99	2	9
Metacritic score	2,162	51.31	16.85	6	96
Ln(Cost)	2,120	17.35	1.08	10.32	20.18
Ln(Gross revenue)	2,455	16.98	1.79	6.53	20.50
Average age of billed performers	2,455	40.34	8.67	15	77
Gross revenue (in Millions of 2005 Dollars)	2,455	59.04	73.99	6.85×10^{-4}	803.60
Star power (in Millions)	2,455	381.48	351.28	0.00	2,351.62
Ln(Star power)	2,455	18.78	3.15	0	21.58
At least two non-white	2,455	0.10	0.30	0	1
At least one non-white	2,455	0.31	0.46	0	1
Share non-white performers	2,455	0.13	0.23	0	1

Table D.2: Descriptive statistics by genre

Panel A: Action/Adventure					
Variable	N	Mean	SD	Min	Max
Run time (minutes)	1,081	106.45	19.90	38	201
IMDb score	945	6.19	1.10	2	9
Metacritic score	945	50.04	17.08	6	96
Ln(Cost)	948	17.82	0.95	12.20	19.78
Ln(Gross revenue)	1,135	16.69	2.82	6.17	20.46
Average age of billed performers	1,128	41.63	8.81	12.5	85
Gross revenue (in Millions of 2005 Dollars)	1,135	76.40	94.20	4.81×10^{-4}	771.00
Star power (in Millions)	1,135	394.00	369.00	0.00	2,350.00
Ln(Star power)	1,135	18.68	3.35	0	21.58
At least two non-white	1,135	0.09	0.29	0	1
At least one non-white	1,135	0.37	0.48	0	1
Share non-white performers	1,135	0.15	0.25	0	1
Panel B: Comedy					
Variable	N	Mean	SD	Min	Max
Run time (minutes)	1,712	99.06	11.81	72	165
IMDb score	1,432	6.02	0.90	1.9	8.3
Metacritic score	1,432	48.02	16.44	1	90
Ln(Cost)	1,021	16.48	1.32	8.94	18.92
Ln(Gross revenue)	1,880	14.19	3.33	4.28	19.48
Average age of billed performers	1,854	40.47	9.61	15.67	99
Gross revenue (in Millions of 2005 Dollars)	1,880	20.90	36.30	7.24×10^{-5}	290.00
Star power (in Millions)	1,880	249.00	291.00	0.00	2,000.00
Ln(Star power)	1,880	17.31	4.82	0	21.42
At least two non-white	1,880	0.09	0.29	0	1
At least one non-white	1,880	0.26	0.44	0	1
Share non-white performers	1,880	0.12	0.24	0	1
Panel C: Drama					
Variable	N	Mean	SD	Min	Max
Run time (minutes)	2,286	108.65	20.85	39	600
IMDb score	1,758	6.56	0.85	2.1	8.8
Metacritic score	1,758	57.08	16.33	5	100
Ln(Cost)	1,189	16.33	1.42	8.89	20.63
Ln(Gross revenue)	2,590	13.49	3.04	5.57	19.81
Average age of billed performers	2,538	41.76	10.04	11	82
Gross revenue (in Millions of 2005 Dollars)	2,590	13.00	31.90	2.63×10^{-4}	401.00
Star power (in Millions)	2,590	246.00	285.00	0.00	1,920.00
Ln(Star power)	2,590	17.31	4.81	0	21.38
At least two non-white	2,590	0.10	0.30	0	1
At least one non-white	2,590	0.25	0.43	0	1
Share non-white performers	2,590	0.12	0.25	0	1

Table D.3: Descriptive statistics by period

Panel A: Pre-2007

Variable	N	Mean	SD	Min	Max
Run time (minutes)	2,302	104.79	20.83	38	600
IMDb score	2,133	6.23	1.01	2	8.9
Metacritic score	2,133	50.37	17.09	5	94
Ln(Cost)	1,754	16.87	1.38	7.00	20.63
Ln (Gross revenue)	2,774	15.14	2.95	2.97	20.51
Average age of billed performers	2,750	39.73	9.32	10	99
Gross revenue (in Millions of 2005 Dollars)	2,774	31.60	56.10	1.94×10^{-5}	804.00
Star power (in Millions)	2,774	205.00	225.00	0.00	1,850.00
Ln(Star power)	2,774	17.29	4.78	0	21.34
At least two non-white	2,774	0.10	0.29	0	1
At least one non-white	2,774	0.30	0.46	0	1
Share of non-white performers	2,774	0.13	0.24	0	1

Panel B: Post-2007

Variable	N	Mean	SD	Min	Max
Run time (minutes)	3,958	103.55	16.88	39	319
IMDb score	2,782	6.27	0.93	1.5	9
Metacritic score	2,782	52.50	17.06	1	100
Ln(Cost)	2,102	16.66	1.47	8.94	19.77
Ln(Gross revenue)	4,169	13.71	3.49	4.10	20.46
Average age of billed performers	4,085	42.95	10.58	11	95
Gross revenue (in Millions of 2005 Dollars)	4,169	23.70	53.80	6.05×10^{-5}	771.00
Star power (in Millions)	4,169	316.00	346.00	0.00	2,350.00
Ln(Star power)	4,169	17.74	4.41	0	21.58
At least two non-white	4,169	0.08	0.27	0	1
At least one non-white	4,169	0.25	0.43	0	1
Share non-white performers	4,169	0.11	0.24	0	1

Table D.4: Descriptive statistics by female share in cast

Panel A: $\leq$ 50% female					
Variable	N	Mean	SD	Min	Max
Run time (minutes)	5,382	104.21	18.61	38	600
IMDb score	4,330	6.27	0.97	1.5	9
Metacritic score	4,330	51.57	17.07	1	100
Ln(Cost)	3,441	16.83	1.41	7.00	20.63
Ln(Gross revenue)	5,933	14.44	3.36	2.97	20.51
Average age of billed performers	5,848	42.03	10.01	12.5	99
Gross revenue (in Millions of 2005 Dollars)	5,933	29.10	57.80	1.94×10^{-5}	804.00
Star power (in Millions)	5,933	288.00	317.00	0.00	2,350.00
Ln(Star power)	5,933	17.71	4.45	0	21.578
At least two non-white	5,933	0.09	0.29	0	1
At least one non-white	5,933	0.28	0.45	0	1
Share non-white performers	5,933	0.12	0.24	0	1
Panel B: $>$ 50% female					
Variable	N	Mean	SD	Min	Max
Run time (minutes)	878	102.72	17.33	40	319
IMDb score	585	6.10	0.95	2.1	8.1
Metacritic score	585	51.62	17.37	6	95
Ln(Cost)	415	16.14	1.44	9.50	19.27
Ln(Gross revenue)	1,010	13.36	3.22	4.10	19.52
Average age of billed performers	987	39.41	11.09	10	75
Gross revenue (in Millions of 2005 Dollars)	1,010	13.60	30.40	6.05×10^{-5}	299.00
Star power (in Millions)	1,010	175.00	227.00	0.00	1,780.00
Ln(Star power)	1,010	16.66	5.10	0	21.30
At least two non-white	1,010	0.06	0.24	0	1
At least one non-white	1,010	0.21	0.41	0	1
Share non-white performers	1,010	0.11	0.25	0	1