Judgment Aggregation in Creative Production: Evidence from the Movie Industry*

Hong Luo† Jeffrey Macher‡ Michael Wahlen§

July 1, 2020

Abstract

We study a novel, low-cost approach to aggregating judgment from a large number of industry experts on ideas that they encounter in their normal course of business. Our context is the movie industry, in which customer appeal is difficult to predict and investment costs are high. The Black List, an annual publication, ranks unproduced scripts based on anonymous nominations from film executives. This approach entails an inherent trade-off: low participation costs enable high response rates; but nominations lack standard criteria, and which voters see which ideas is unobservable and influenced by various factors. Despite these challenges, we find that such aggregation is predictive: listed scripts are substantially more likely to be released than observably similar but unlisted scripts, and, conditional on release and investment levels, listed scripts generate higher box office revenues. We also find that this method mitigates entry barriers for less-experienced writers as: (i) their scripts are more likely to be listed than those by experienced writers and to rank higher if listed; (ii) within scripts by less-experienced writers, being listed is associated with a higher release rate. Yet, the gap in release probabilities relative to experienced writers remains large, even for top-ranked scripts. These results can be explained by the premise that scripts from less-experienced writers are more visible among eligible voters than scripts from experienced writers. This highlights idea visibility as an important determinant of votes and surfaces the trade-offs, as well as potential limitations, associated with such methods.

*We thank Dan Gross, Aseem Kaul, Ramana Nanda, Catherine Turco, Ezra Zuckerman, and seminar participants at Harvard Business School, Strategic Research Forum, Strategy Science Conference, University of Toronto, and MIT Economic Sociology Working Group for helpful comments. Tejas Ramdas and Esther Yan provided excellent research assistance. All errors and omissions are our own.

†Harvard Business School; email: hluo@hbs.edu.

‡Georgetown University (McDonough); email: jeffrey.macher@georgetown.edu.

§MIT Sloan; email: jmwahlen@mit.edu.

1 Introduction

Successfully evaluating and selecting early-stage ideas is immensely important and yet extremely difficult. This is especially true in domains—such as creative industries or those with low R&D-intensity—that lack concrete and objective quality signals (Waldfogel, 2017; Scott et al., 2020). In these settings, even industry experts may be limited in their ability to predict success. Moreover, the volume of ideas is often large, making screening costly. To address these challenges, decision makers often limit their searches to ideas from creators with strong, observable track records or rely on reputable gatekeepers. However, such remedies may lead to high entry barriers for those lacking prior experience or industry connections, as well as to an inefficient allocation of resources (Caves, 2000; Kaplan et al., 2009; Stuart and Sorensen, 2010).

One promising approach to addressing these challenges is the so-called 'wisdom of crowds' (Surowiecki, 2005). Though specific applications differ, the basic idea is rooted in the statistical principle that the aggregation of noisy but relatively independent estimates systematically increases precision (Condorcet, 1785; Bates and Granger, 1969). With more-precise estimates of idea quality, decision makers may find it less necessary to rely on the reputations of the creators, which can potentially reduce entry barriers for outsiders (Mollick and Nanda, 2015; Sorenson et al., 2016).

Despite this intuitive appeal, firms do not commonly use crowd-based judgment, especially in high-risk and high-cost settings. When stakes are high, decision makers—who tend to have substantial experience themselves—may naturally be skeptical of the judgment of an amateur crowd (Larrick and Soll (2006)). Thus, to generate meaningfully better and credible predictions in such settings, a sizable number of *high-quality* participants appears critical. However, as Csaszar (2018) suggests, this may be difficult because: (i) it may be prohibitively expensive to recruit such crowds; and (ii) in innovative settings, firms may not want to broadcast their early-stage projects due to appropriation concerns (Arrow, 1962).

In this paper, we study a novel approach to crowd creation that seeks to overcome the above obstacles in a high-stakes, hard-to-predict, and economically significant setting: the movie industry. Whether a movie will appeal to customers is notoriously difficult to predict (hence the award-winning screenwriter William Goldman's famous statement: "Nobody knows anything."), and investment costs are high (the median budget is $25m in our sample). The process is further complicated by the need to convince multiple parties: the producer, having purchased or commissioned a script, needs to attract key talent and, crucially, to obtain financing from studio executives. Our data indicate that only 15 percent of sold scripts are theatrically released.

The Black List—the judgment aggregator that we examine—is an annual, anonymous survey of "executives at major studios, major financiers, or production companies that have had a film in major release in the past three years." Survey participants are asked to nominate "up to ten of their favorite scripts that were written in, or are somehow uniquely associated with [a particular year], and will not be released in theaters during this calendar year." Scripts receiving more than a threshold number of nominations are published as a list and ranked by the number of nominations received. Nearly 80 percent of listed scripts are already sold by the annual release of the List, but, given the release rate of only 15 percent even for sold scripts, the aggregated judgment may prove valuable for studio executives in their financing decisions and in attracting talent. In his 2005 TED Talk, Franklin Leonard recounted his motivation for creating the Black List: a junior executive at a production company at the time, Leonard wanted to spend his time more efficiently by focusing on high-quality scripts and to be more inclusive by casting a wider net, thereby challenging conventional wisdom of where scripts are found and how they are evaluated.[1]

By relying on industry practitioners to nominate scripts that they became aware of in their normal course of business and by asking them only for nominations, the Black List requires minimal respondent time and effort. This enables it to achieve a nearly 50-percent response rate among busy industry insiders, resulting in a "crowd" of 250-300 credible voters. Yet such an approach also presents important tradeoffs: (i) the nominations are not based on uniform or extensive rating criteria, which may limit the List's predictive power; (ii) agents' promotional efforts and routine social interactions may undermine the independence criteria for effective judgment aggregation; and (iii) instead of centrally allocating ideas to judges, as in NIH grant reviews or venture idea competitions, which and how many eligible voters read a given script may be influenced by various factors in the idea-circulation process and is generally unobservable.

In light of Leonard's motivations and the tradeoffs of the Black List's approach, we examine two questions. First, regarding the "wisdom" of the List: Can such an aggregation system actually differentiate idea quality? Second, regarding barriers to entry: Does the List help surface ideas from less-experienced writers? And does being listed and highly ranked help close the gap in release outcomes by writer experience?

Our empirical analysis is based on two samples. The first includes all Black-Listed scripts over 2007-2014. This sample provides the exact number of nominations and a detailed picture of the most selective set of scripts. The second includes a sample of scripts that had been sold over the same 2007-2014 period (of these, 11 percent were Black-Listed). This second sample allows for comparison between listed and unlisted scripts, even though it is still conditional on scripts that were good enough to be sold.

Our first set of findings is that, despite its potential flaws, the Black List appears to be predictive, in that being listed helps further differentiate the quality of scripts that are otherwise observably similar. In a regression analysis that controls for factors observable to the industry before Black List publication—including the characteristics of the script, writer, and agent, as well as the producers or studios that own the script—we find that listed scripts are 70-percent more likely than unlisted scripts to be produced. Furthermore, conditional upon being released, listed movies generate greater returns to investment than unlisted ones with the same

[1] https://www.ted.com/talks/franklin_leonard_how_i_accidentally_changed_the_way_movies_get_made_nov_2018?language=en.

budget, providing further evidence that listed scripts are of higher quality.

Our second set of findings contrasts scripts by writers of different experience levels. The Black List appears quite effective at highlighting scripts from relatively novice writers: Conditional on sold scripts, those by less-experienced writers are more likely to be listed than are those by experienced writers; and conditional on being listed, they also rank higher. Among sold scripts by less-experienced writers, being listed is also associated with a higher release probability. Interestingly, however, the gap in the release probability relative to experienced writers does not shrink with more nominations, even for the highest-ranked scripts. This last finding is quite surprising, as the prior literature suggests that observable signals tend to be less important when idea quality is more certain; and, thus, we should expect the outcome gap by writer experience to decrease, at least for scripts on which industry experts reach relatively high levels of consensus (Simcoe and Waguespack, 2011; Azoulay et al., 2013).

While several factors may likely work together to explain the patterns described above, the most compelling explanation relates to the fact that the Black List is embedded in the normal course of script sales and development, which implies that the visibility of a given script among potential buyers, talent, and financiers (many of whom are also Black List voters) will be a systematic determinant of the nominations. If scripts from less-experienced writers are more visible among the voters, they will, indeed, garner more nominations. Adjusting for the wider visibility, decision makers would infer an inferior posterior quality for scripts from less-experienced writers, which may explain why the gap relative to experienced writers remains large, even for the top-ranked scripts. Such an explanation seems plausible because, unlike experienced writers, who can sell their ideas without talking to many buyers, less-experienced writers may need much greater exposure in order to find interested buyers or the best match.

Overall, our results suggest that despite its weaknesses, the Black List can help to further differentiate quality among observably similar ideas in the mainstream segment of a major creative industry. Apart from this specific context, such methods may be valuable in settings that share the following broad characteristics: (i) idea quality is difficult to predict (even for experts); (ii) assembling a large number of experienced but diverse evaluators via a structured process is costly and hard to scale; and (iii) ideas are circulated in a decentralized manner, with individuals producing judgments in their normal course of business—e.g., in industries with an active market for ideas (Arora et al., 2001; Gans and Stern, 2003). Our results also suggest that such methods may help to mitigate barriers to entry for lesser-known idea providers, though with important caveats. By providing more-precise information, such a list might allow the better ideas from these providers to stand out. Furthermore, in contexts in which ideas from lesser-known providers tend to be circulated more widely, casting a wide net (and treating the votes with equal weights) appears to also have some corrective effect, in that good ideas from those who find it necessary to sell their ideas harder and wider in the marketplace face greater chances to rank on top. But such methods, by design, do not change how ideas circulate in the marketplace. Thus, they are unlikely to meaningfully aid those who find it difficult to obtain an audience in the first place. Furthermore, because idea visibility is likely to be influenced by various factors and not easily observable, decision makers may make crude adjustments that further limit the intended impact.

Related Literature

This paper contributes to the body of research on the evaluation of early-stage ideas. An important topic is whether idea potential is predictable. Scott et al. (2020) show that upon reading new venture summaries, experts are able to differentiate quality, but only for R&D-intensive sectors and not for consumer product, consumer web, mobile, and enterprise software sectors. Astebro and Elhedhli (2006) suggest that while expert panels are effective in forecasting the commercial success of early-stage ventures, structured processes and extensive evaluator training seem necessary. Our paper shows that aggregating industry-expert opinions via a simple nomination method, but on a large scale, can predict the quality of mainstream projects in a major creative industry. A second research stream examines how reviewers' characteristics and incentives may systematically influence their evaluations: Li (2017) shows that reviewers are both better-informed and more positively biased about projects in their own areas, while Boudreau et al. (2016) find that evaluators systematically give lower scores to proposals that are closer to their own expertise or highly novel. While prior research typically takes the formation of reviewer panels as given, our paper highlights that such formation can be endogenous and unobservable and that idea exposure may systematically influence aggregated votes.

Recent research applies the wisdom of crowds to prediction problems in a variety of settings, including sales volume and project completion dates (Cowgill and Zitzewitz, 2015). Our paper provides novel evidence on the application of this principle to predicting the success of early-stage ideas. As such, our paper closely relates to a growing research stream on the evaluation of crowd-funded and crowd-sourced projects (Mollick, 2014; Agrawal et al., 2015; Sorenson et al., 2016). Mollick and Nanda (2015) find that crowds' evaluations, while generally coinciding with those of trained experts, may reduce the incidence of "false negatives" (i.e., viable projects that experts turn down). Riedl et al. (2013) compare the quality of aggregated judgments when users rate crowdsourced projects under different criteria. As crowd-funded and crowd-generated projects involve mainly small budgets, our contribution to this literature is to examine a novel, low-cost approach that aims to assemble a credible, high-quality crowd and to aid in decision-making in a high-stakes setting.

Finally, our paper relates to research examining the film industry. In contrast to studies examining film critics' ability to predict the box office success of completed movies (e.g., Eliashberg and Shugan (1997)), our paper focuses on early-stage projects, the quality of which is much harder to discern. Another line of research focuses on the use of writers' observable quality signals as a means of evaluating screenplays (Goetzmann et al., 2013; Luo, 2014). Our paper builds on this research by examining the ability of an information aggregator to pool judgments in order to arrive at a more precise judgment beyond observable signals.

2 Background and Conceptual Framework

2.1 Background

Making a movie is a long, costly, and uncertain process. It begins with a script. Broadly speaking, producers acquire scripts from two sources. First, producers option or purchase finished scripts from writers. Anecdotally, thousands of new scripts enter the industry every year, but only a few hundred are acquired. Second, producers may hire writers to adapt previous works (e.g., a novel) into scripts. We refer to scripts acquired through both means as 'sold.' With a script in hand, producers are tasked with finding talent and securing financing. The most important financiers are movie studios. Most producers are independent, whereas some—typically the more-established—are affiliated with specific studios. Only about 15 percent of sold scripts are released; and for released movies, the median time from script sale to release is 2.1 years.

Agents—intermediaries representing writers—have long been important industry gatekeepers. When selling a script, agent (and agency) reputation, experience, and connections are critical to identifying potential buyers and generating interest. Large agencies may also leverage their client portfolio by helping producers assemble a director and acting team. But obtaining agency representation is challenging; well-connected agents are busy, generally rely on referrals and do not read unsolicited materials by unknown writers.

As mentioned, the Black List survey is embedded in the existing sales and development process of movie projects. Potential voters (executives of production companies and studios) typically read scripts when they are being actively shopped around, either by the writer and his/her agent in the sales process, or by the producer in searching for talent and in securing financing after the script is sold. An article in *The Atlantic* noted that a script must get "into the hands of executives so that they may, in turn, like it and vote for it."[2]

Because the List is published annually, close to eighty percent of the listed scripts are already sold at the time of publication. Thus, in a typical case, the potential effect of being listed is no longer about getting the script sold. But with only a 15-percent release rate even for sold scripts, the List's stamp of approval could prove beneficial in convincing studio executives and attracting talent. For instance, Helen Estabrook, a producer of *Whiplash*, said, "[T]he spot on the Black List offered a level of validation that proved, 'Hey, I'm not a crazy person—many other people agree with me.'" A script's presence on the List "reassures financiers, executives, and producers that they are not going too far out on a limb" (*The Atlantic*, 2017).

Our data show that 28 percent of listed scripts are subsequently released, which is consistent with what the Black List reports. Listed movies have also achieved critical success, garnering 241 Oscar nominations and 48 wins (McGee and Mcara, HBS Case 317-027). While some observers are skeptical that such movies

[2] "The Hollywood List Everyone Wants to Be On" by Alex Wagner, *The Atlantic*, March 2017.

would not have been successful without the List (e.g., *Slate's Culture Blog*, 2011), it is not surprising that the *LA Times* declared in 2014 that "[t]he Black List has become a Hollywood institution."

2.2 Conceptual Framework

The above context suggests that the data-generating process consists of three broad stages. In the first stage, a script may encounter potential Black List voters through the normal sales and development process. In the second stage, at the end of the year, the Black List aggregates the individual judgments formed throughout the year in the first stage and publishes the number of nominations that the scripts received. Finally, in the third stage, studio executives may update their beliefs about the quality of projects after observing the number of nominations, which, in turn, may affect their decision making about which projects to finance.

Empirically, we do not observe this whole process; rather, we observe the writer's experience level, the outcome of the second stage of the above process (the Black List nominations), and the eventual outcome of the third stage (whether a project is produced and its box office revenues). In this section, we provide a parsimonious model that generates some basic intuitions to help structure the empirical analysis based on what we can observe in the data and surface potential mechanisms. The model focuses on (i) how the votes may be determined in the second stage; and (ii) Bayesian updating by the studio on script quality based on realized nominations in the third stage.

2.2.1 Model Setup

Assume that the inherent quality of a movie script is either high or low: $q \in \{q_H, q_L\}$.[3] A voter, having read a given script, obtains a binary quality signal $s \in \{s_H, s_L\}$, where $\mathrm{P}(s_H|q_H) = p_1$ and $\mathrm{P}(s_H|q_L) = p_2$. Assume that $p_1 > p_2$; that is, the likelihood of obtaining a positive signal (s_H) is greater when the true quality is high than when the true quality is low. For simplicity, assume that voters (and, hence, their signal structures) are homogeneous. In addition, any two voters' signals are conditionally independent given the true script quality. A voter will nominate a script as long as a positive signal is received.

We focus on two factors that determine the number of nominations a script receives, and both may vary with the writer's experience level, $w \in \mathbb{R}$. The first factor is the prior belief about the probability that a script is high-quality, $\pi(w)$. $\pi(w)$ is an exogenous factor, as it reflects the probability of success understood by the industry based on past projects by writers of similar characteristics. We assume that scripts from experienced writers are more likely to be high-quality than scripts from less-experienced writers; i.e.,

Assumption 1. $\pi'(w) > 0$.

This assumption simplifies the analysis greatly and is supported by empirical evidence.[4]

[3] Generally speaking, quality may mean different things to different people. Given that the outcome variables that we consider are release probabilities and box office revenues, we deem 'quality' to be about commercial success.

[4] In particular, using scripts from the DDP database in time periods prior to the Black List, we find a monotonic *increasing* rela-

The second factor is the number of voters who have actually read a particular script, $n(w)$. Script visibility is important because the probability of obtaining a nomination from a voter who has not yet read a script is (in principle) zero. As discussed above, script visibility among Black List voters is likely to depend on how widely the script is circulated; and the breadth of circulation is endogenous in the sense that it is affected by multiple players' decisions that may depend on the writer's experience level. For simplicity, we assume that visibility is a reduced-form function of w and that it is (weakly) negatively correlated with writer experience:

Assumption 2. $n'(w) \leq 0$.

In the Online Appendix, we present a simple model that describes one out of potentially multiple mechanisms through which Assumption 2 may hold: less-experienced writers will approach more buyers than experienced writers will because, under Assumption 1, the chance of any individual buyer receiving a positive signal about their scripts is lower.[5] As explained previously, it is important to emphasize that multiple factors, including strategic behaviors, may affect the ultimate scope of script exposure. Some of these forces may push script visibility to be greater for less-experienced writers, while others may push in the opposite direction. The sign of $n'(w)$ is, ultimately, an empirical question. As we discuss in detail later, however, what we need in order to interpret the empirical results is the *possibility* of $n'(w) < 0$. To this end, it suffices to know that Assumption 2 may hold under plausible conditions.

Let m be the number of nominations, the expectation of which is a function of writer experience, w:

$$\mathrm{E}[m|w] = \pi(w)n(w)p_1 + (1-\pi(w))n(w)p_2. \tag{1}$$

In each state of the world (i.e., the true quality being either q_H or q_L), the number of nominations follows a binomial distribution. Given each state, the expected number of nominations is, thus, the number of voters who have read a given script ($n(w)$) multiplied by the probability that each voter draws a positive signal given this state. The above equation weighs the two states by their respective prior probabilities, $\pi(w)$ and $1-\pi(w)$.

For a script that receives m nominations and is from writer of experience level w, the updated belief about the probability that it is of high quality (according to Bayes' rule) is:

$$\mathrm{P}(q_H|m,w) = \frac{\pi(w)\mathrm{P}(m|w,q_H)}{P(m|w)} = \frac{\pi(w)\mathrm{P}(m|w,q_H)}{\pi(w)\mathrm{P}(m|w,q_H) + (1-\pi(w))\mathrm{P}(m|w,q_L)}. \tag{2}$$

In words, this updated belief is the probability that the true script quality is high and that it receives m nominations ($\pi(w)\mathrm{P}(m|w,q_H)$), divided by the total probability that the script receives m nominations ($P(m|w)$),

tionship between release likelihood and writer experience, which is consistent with Assumption 1. Furthermore, quantile regressions using movies that are released also do not indicate a heavier right tail in the distribution of box office revenues for scripts from less-experienced writers (see results in Online Appendix Table B5).

[5] In this simple model, the writer endogenously chooses n as he attempts to sell a script. We assume that marketing to an additional buyer incurs an extra cost; apart from monetary costs, disutility of effort, and the opportunity cost of time, industry accounts also suggest that sellers have incentives to avoid an unnecessarily wide sales scope due to concerns of information leakage. We show that the optimal number of buyers to market a script to decreases with writer experience under plausible conditions.

which is the sum of the probabilities of receiving *m* nominations in the two possible states of the world.

In the following, we derive two sets of results that we take to the data. Proposition 1 suggests that the Black List is predictive, which we examine empirically in Section 4.1; and Propositions 2 to 4 relate to writer experience, which we examine in Section 4.2. See Online Appendix A for all the proofs.

2.2.2 Number of Nominations and Updated Beliefs About Script Quality

Given writer experience *w*, the following result shows that a greater number of nominations implies a higher likelihood that the script is of high quality (that is, $\frac{\partial \mathrm{P}(q_H|m,w)}{\partial m} > 0$):

Proposition 1. *Given writer experience w, the posterior belief about the probability of the script's quality being high increases with the number of nominations, m.*

We cannot directly test Proposition 1 because it is impossible to observe the true quality of a script. What we can do is to correlate nominations with release probability. But because the List may not only predict, but also influence release outcomes, there are three possible interpretations if this correlation is, indeed, positive.

The first possibility is that the Black List is purely predictive and does not influence decision making. Consider a studio with two scripts in its portfolio: if the Black List ranks these two scripts in the same order as the studio's internal (private) ranking, though still predictive, the List would have little impact on decision making. In other words, listed scripts will be more likely to be produced than unlisted scripts, even without the existence of the Black List.

The second possibility is that the Black List is predictive and has a causal impact on decision making. If the studio above could not further differentiate the two projects on its own, but one is Black-Listed while the other is not, the probability of release for the listed script would be greater. Or if the studio ranks the two projects internally in the opposite order, the Black List rankings may provide a sufficiently strong signal to reverse the studio's preference.

The third and final possibility is that the Black List does not predict script quality but somehow still influences decision making. Note that Proposition 1 is based on the assumptions that voters' signals are informative about true script quality and that they are conditionally independent. However, as discussed in the Introduction, these assumptions are not obvious. Due to the Black List's unstructured approach, voters may not share a consistent understanding of idea quality, and their signals may be correlated, thereby limiting the informativeness of their signals. For example, if $p_1 = p_2$, the number of nominations would have a zero correlation with an idea's quality. However, even if the List is not predictive, it is still possible that it influences decision making. This may happen if the studio (mistakenly) assumes that voter signals are informative of true quality, or if the Black List, by providing a focal point that attracts attention, helps to overcome any coordination failures.

We may exploit data on resource allocation and box office revenues to rule out the third possibility—that the Black List is not predictive. Specifically, suppose that a movie's box office revenue is a positive function of three inputs (as in a typical production function): script quality; physical capital (measured by production budget); and human capital (measured by the characteristics of the director and main cast). We can test the following null hypothesis: *If the Black List is not predictive, conditional upon release, listed and unlisted scripts should generate similar box office revenues, given the same physical and human capital allocations.*[6] If the data reject the null hypothesis, we may conclude that the Black List is, indeed, predictive. To distinguish between the first and the second possibilities, however, we need to know a studio's internal rankings of their portfolio projects and how confident it is about its own information relative to external input such as the Black List. Unfortunately, such data are impossible to obtain. As a result, we cannot pin down whether the Black List, along with being predictive, also influences decision making in a causal sense.

2.2.3 Nominations, Writer Experience, and Updated Beliefs About Script Quality

Assumptions 1 and 2 imply that writer experience w affects our outcome variables via two different channels—(i) they have different prior probabilities of being high-quality; and (ii) they may also have different levels of visibility among Black List voters. We write all of the following propositions in two parts. The first part focuses on the case of $n'(w) = 0$; that is, all scripts are read by the same number of voters. This scenario provides a null hypothesis that rules out visibility as a systematic determinant of votes. In the second part, we add the *additional* effect due to the possibility that visibility is also different. As we discuss later, in Section 4.2, this additional channel appears to be necessary to explain our empirical results on writer experience.

The following proposition is about how the number of nominations may vary by writer experience:

Proposition 2. *If $n'(w) = 0$, the expected number of nominations is greater for experienced writers than for less-experienced writers. If $n'(w)$ is sufficiently negative, the expected number of nominations can be greater for less-experienced writers.*

To see this, take the derivative of $\mathrm{E}[m|w]$ with respect to w:

$$\frac{\partial \mathrm{E}[m|w]}{\partial w} = \underbrace{\frac{\partial \mathrm{E}[m|w]}{\partial \pi(w)}}_{+} \pi'(w) + \underbrace{\frac{\partial \mathrm{E}[m|w]}{\partial n(w)}}_{+} n'(w). \tag{3}$$

$\frac{\partial \mathrm{E}[m|w]}{\partial \pi(w)} > 0$ is intuitive because when voters' signals are positively correlated with true script quality, more positive signals will be drawn if the prior probability of being high-quality is higher. $\frac{\partial \mathrm{E}[m|w]}{\partial n(w)} > 0$ is also intuitive, as scripts that have been read by more voters are likely to receive more nominations. Recall that $\pi'(w) > 0$ under Assumption 1. Thus, when $n'(w) = 0$, we expect that the probability of being Black-Listed

[6]The Online Appendix provides more-formal reasoning for this null hypothesis. There, we explicitly model the studio's decision to produce a movie as a function of writer experience and the realized number of nominations.

will be higher for experienced writers and that their scripts should be ranked higher. However, when scripts from less-experienced writers are read by sufficiently more voters (i.e., $n'(w)$ is sufficiently negative), we may observe the opposite result, despite their lower prior probability of success.

The next proposition shows how updated beliefs about the script's quality may vary by writer experience:

Proposition 3. *Consider scripts that receive the same number of nominations, m. If $n'(w) = 0$, the posterior belief about the probability that the script is high-quality is greater for experienced writers than for less-experienced writers. If $n'(w) < 0$, this positive gap by writer experience is even larger.*

This result is obtained by taking the derivative of $\mathrm{P}(q_H|m,w)$ with respect to w:

$$\frac{\partial \mathrm{P}(q_H|m,w)}{\partial w} = \underbrace{\frac{\partial \mathrm{P}(q_H|m,w)}{\partial \pi(w)}}_{+} \pi'(w) + \underbrace{\frac{\partial \mathrm{P}(q_H|m,w)}{\partial n(w)}}_{-} n'(w) \tag{4}$$

$\frac{\partial \mathrm{P}(q_H|m,w))}{\partial \pi(w)} > 0$ shows that for scripts receiving the same number of nominations, one should infer a higher quality for scripts with a higher ex-ante expectation. $\frac{\partial \mathrm{P}(q_H|m,w)}{\partial n(w)} < 0$ shows that the quality inferred should be discounted for scripts that are believed to have been read by more voters. For instance, consider two scripts that both receive five nominations, but ten voters have read the first and twenty voters the second. One should infer a lower quality for the second script, as it receives more negative signals than the first relative to the same number of positive signals. Thus, under Assumption 1 that $\pi'(w) > 0$, the first term in the above equation is positive; and with $n'(w) \leq 0$ (Assumption 2), the second term is non-negative. Thus, while w operates through two different channels, both produce the same directional predictions for $\mathrm{P}(q_H|m,w)$. Consequently, in contrast to Proposition 2 in which the two channels generate opposite predictions for $\mathrm{E}[m|w]$, Proposition 3 does not allow these two channels to be empirically distinguished.

The final result is about how the gap in updated beliefs by writer experience in the previous proposition may change as the number of nominations increases (that is, how $\frac{\partial \mathrm{P}(q_H|m,w)}{\partial w}$ changes as m increases).

Proposition 4. *If $n'(w) = 0$, there exists a unique m^* such that the (positive) gap by writer experience first increases with m for $m < m^*$ and then decreases with m for $m > m^*$. If $n'(w) < 0$, the threshold after which the gap starts to decrease is higher than m^*.*

Intuitively, when the number of nominations is relatively small ($m < m^*$), the signal provided by the Black List is quite weak (or even negative, as only a small number of voters having read the script received a positive signal). In this range, an incremental increase in m is more credible for scripts from writers with a higher prior probability of success. In other words, prior probability π and nominations m are complements in this range; and the gap by writer experience will actually increase with m. However, when the number of nominations is relatively large ($m > m^*$), Black List nominations provide relatively precise and positive signals. In this range, nominations and prior probability become substitute signals, and the gap by writer experience should decrease as m increases. In the case of very large n and m, we can show that the gap would converge to zero. The converging result in the $m > m^*$ range illustrates the intuition from the prior literature that, when uncertainty about idea quality becomes smaller, the importance of observable quality should decrease.

When $n'(w) < 0$, we can show that the threshold required for the gap to start to converge is higher; in other words, there will be a greater range of m's in which the gap will continue to increase. This is also quite intuitive, as people expect more voters to have read scripts from less-experienced writers, further discounting the updated quality belief for less-experienced writers. This discount, due to greater visibility, helps to maintain the size of the gap even for a relatively high number of nominations.

3 Data

3.1 Samples

Our data come primarily from two sources. Our first sample (*Sample BL*) uses annual Black List publications from 2007-2014 and includes 701 listed scripts. We exclude the first two years (2005 and 2006) because these publications do not provide any buyer-side information even if they are sold (as mentioned previously, 79 percent of them are). We also cut off the data after 2014 to provide a sufficient window for scripts to be produced before our final outcome data collection date (October 2017).

Our second sample (*Sample DDP*) uses Done Deal Pro (DDP), an internet database that tracks script transactions on a daily basis. DDP is recognized by various industry organizations (e.g., the Writers Guild of America) as one of the leading movie project information sources. We exclude DDP records on adaptation contracts to avoid confounding different degrees of uncertainty, as we cannot systematically observe whether these scripts were later completed and were good enough to advance to the next stage (i.e., a stage comparable to that of the majority of the Black-Listed scripts). Thus, Sample DDP is more representative of scripts based on original ideas and less representative of the most expensive movies. We also exclude rewrite contracts. In addition, for a cleaner interpretation, we exclude 21 observations of Black-Listed scripts that were sold after the List was published. We examine this small subset of scripts separately in Online Appendix B.4. After these exclusions, Sample DDP includes 1,566 observations over 2007-2014.

We can think of the two samples as covering different segments of the nomination distribution. Sample BL provides a detailed picture of the top end of the distribution (above the cutoff threshold), with the key variation being the number of nominations (and ranking) that a script receives. With Sample DDP, we can compare listed to unlisted scripts, though it is still conditional on scripts being good enough to be sold.[7]

The Black List and DDP both provide information on the writer, the representing agent and agency, and

[7] Sample BL is not a subset of Sample DDP, however, because the latter is a comprehensive, but incomplete and selected, sample of all scripts that are sold and, thus, does not capture all Black-Listed scripts.

the buyers (production companies and/or movie studios that purchased or commissioned scripts). We further collect data from complementary sources, including: (i) the Internet Movie Database (IMDb), which provides writers' industry experience; and (ii) The Numbers, which offers outcome information, including whether scripts have been theatrically released and, if so, the box-office revenues generated.

3.2 Variables

Table 1 provides summary statistics, summarizing Sample BL in the left panel and Sample DDP on the right. In the following, we explain in detail the key variables and then briefly describe the control variables.

For Black-Listed scripts, because the difference in total respondents may result in variations in nominations across years, we create a categorical variable, *Black List Rank Group* to indicate where listed scripts place in a survey year: 4 = top 5 (5.7 percent of all scripts in Sample BL); 3 = top 6 to top 20 (19.4 percent); 2 = top 21 to just above the cutoff number used in a given year (53.4 percent); and 1 = those at the cutoff (21.5 percent). We also check the robustness of our results to using a continuous variable, defined as the number of nominations received by a script divided by the total number of respondents in that year. For scripts in Sample DDP, *Black-Listed* indicates listed scripts, which account for 11 percent of the sample.

Released indicates whether a script has been released in U.S. theaters and has generated positive box office revenues. For released movies, we have the movie's worldwide (U.S. plus international) box-office revenues and, for about two thirds of them, the production budget; both variables are adjusted for inflation.

We measure writer experience by the number of writing credits the writer obtained in the previous ten years for movies distributed by the top-30 movie studios (*Writer Major Credits*). If there is more than one writer, we use the maximum of the writers' experience. The ten-year restriction better captures writers' recent industry experience and current status versus simply measuring industry tenure. The restriction to movies distributed by the top-30 studios avoids inflating writer experience with small-budget independent movies. *Experienced Writer* indicates whether a writer has had one or more major writing credits in the previous ten years.

Control variables include the experience of the writer's agent: *Top-10 Agency* indicates whether the agency representing the writer ranks in the top ten in terms of market share, calculated using all transactions in the DDP database; *Experienced Agent* indicates whether the agent has more than 15 script transactions prior to a given sale. For buyers, *Producer Major Credits* is the number of producing credits obtained in the previous ten years for movies distributed by the top-30 studios, and *Movie Studio Buyer* indicates whether a movie studio is listed among the buyers. For the 21 percent of scripts in Sample BL that have not yet sold at the time of the List's publication, the two buyer variables are replaced with zero in the regressions. *Whether Reported* indicates whether any articles about a given script are published in the top two trade magazines—*Hollywood Reporter* and *Variety*—according to Google search results two years before the Black List publication date in the year the script was sold. For Sample DDP, we have additional characteristics of the scripts, including whether they were based on original content (versus an adaptation); the source of content if not original (e.g., books and short stories); and genre. Three percent of observations are indicated as having been purchased through an auction process. For half of the records, some talent is committed to the project at the time of sale.

For released movies, apart from production budget, we collect other measures of investments. These include the quality or popularity of the leading cast;[8] the total number of screens showing the movie during the opening weekend; and a seasonality index for weekly fluctuations in movie demand.[9]

4 Results

In addressing the two questions posed in the Introduction, we present and discuss two sets of results: Section 4.1 examines the relationship between nominations and release outcomes to determine whether the Black List is predictive, while Section 4.2, by examining the results by writer experience, explores whether the Black List reduces entry barriers. Finally, we discuss the generalizability of the Black List method in Section 4.3.

4.1 Black List Nominations and Release Outcomes

We first focus on whether or not a movie is released, as this is a systematic outcome measure available for all observations. Figure 1a, using Sample DDP, shows that Black-Listed scripts are 10.8 percentage points more likely to be released than unlisted scripts, representing a 76-percent difference (p-value = 0.001). Using Sample BL, Figure 1b shows that among listed scripts, there is a monotonic increasing relationship between Black List rank group and release probability. The biggest difference is between top-five scripts (47.5 percent) and those in the three lower-ranked groups (25.0 percent on average). The differences between lower-ranked groups, though not statistically significant at the conventional level, are economically non-trivial.

The raw data, thus, show a positive correlation between Black List nominations and release probabilities. Our goal is to understand whether the List, by aggregating voters' judgments formed after reading a script, helps to differentiate scripts beyond what readers can glean from the scripts' observable characteristics. For this reason, we control, to the extent possible, for factors observable to the industry *prior* to the List's publication. Consider the following specification, which uses Sample DDP as an example:

$$\text{Release}_i = \delta \text{Black-Listed}_i + \beta X_i + c_{jt} + \varepsilon_i, \tag{5}$$

where Release_i indicates whether script i is theatrically released; Black-Listed_i indicates whether script i is listed; and X_i represents the comprehensive list of observable factors that may correlate with the likelihood of release. These include the characteristics of the writers, their agents, and the scripts, as described in

[8] These include two variables: *Award-winning cast* indicates whether any of the leading cast has won any award (Academy Award, Golden Globe, and BAFTA) prior to the release year of the focal movie; and *Star score* is an index generated by the-numbers.com, based on the total box office revenues of movies an actor or actress has been in prior to the release year of the focal movie.

[9] Following Basuroy et al. (2003), the *Seasonality* index is based on the total number of tickets sold for all movies shown during a given weekend, averaged across years.

Section 3.2; whether the article is reported (as significant publicity may result in voter signals being highly correlated); and year fixed effects. It is important to highlight that we also control for the buyer of a given script because this is generally common knowledge by the time of the Black List survey. It is, therefore, possible that voters adjust their opinions about scripts based on this information, as they may know that standards differ among buyers. We do this by controlling for producer experience and by including studio fixed effects.[10] In some specifications, we also include a full set of studio× year fixed effects (c_{jt}). For all regressions, to be conservative, we use two-way clustered standard errors at the studio and year levels to allow for free correlations between scripts developed within a studio or in the same year (Cameron et al. (2011)).

Columns 1-3 in Table 2 report results using Sample DDP: Column 1 controls for year and studio fixed effects; Column 2 uses studio×year fixed effects; and Column 3 further excludes top-20 listed scripts, under the assumption that lower-ranked scripts are more likely than top-ranked scripts to be observably similar to unlisted scripts and less likely to be subject to publicity hype. All specifications produce similar results. The most conservative estimate shows that listed scripts are 10.9 percentage points, or 72 percent, more likely to be released than unlisted scripts.

The last three columns in Table 2 report regression results using Sample BL. Columns 4 and 5 show that top-five scripts are 28 percentage points more likely to be released than those in the default group (scripts receiving only the threshold number of nominations), representing a 136-percent increase. The coefficient of 'Top 6-20' represents a 50-percent increase relative to the default group (p-value is 0.154); and the 'Top 21-above threshold' coefficient represents a 32-percent difference (p-value is 0.208). Column 6 confirms this positive correlation using the ratio between the number of nominations received by a script and the total number of respondents in this year. The coefficient suggests that when this ratio increases by one standard deviation, the release likelihood increases by 6.7 percentage points.[11]

4.1.1 Box Office Performance and Production Budget

The above results show a significant positive relationship between Black List nominations and the likelihood of release for observably similar projects. But, as discussed in Section 2.2.2, because the Black List may both influence and predict outcomes, we cannot yet conclude that it is predictive. In this section, we provide further evidence by examining the box office performance of released movies.

Figure 2a shows that for released movies in Sample DDP, the density distribution of log(Worldwide Box Office) for listed movies is to the right of that for unlisted ones; listed movies are associated with both a higher mean (p-value = 0.003) and a smaller standard deviation (p-value $<$ 0.001). Online Appendix Fig-

[10]If no movie studios are included as a buyer in the DDP record, we pool these scripts together as a single default group.

[11]It is worth noting that the addition of the Black List variables substantially increases the amount of variation explained by these regressions. Online Appendix Table B1 reports the marginal increase in R^2 and adjusted R^2 by including the Black List variables in the regressions in Table 2. The percent increase ranges from 3.2% to 16.0%, depending on the specification, and the median is 8.9%.

ure B2a shows that the density distribution of log(Production Budget) is statistically similar between listed and unlisted movies, which suggests that better performance of listed scripts is not driven by greater investment. Panel (a) of Table 3 puts the above analysis into regression form. Column 1 confirms that released Black-Listed movies generate significantly higher box office revenues than unlisted movies do. Column 2 shows that listed scripts are not associated with a higher production budget. Similarly, regression results reported in Online Appendix Table B3 show that listed and unlisted movies are also statistically similar for other measures of investment level: the quality of the leading cast; the number of opening screens; and the release window. Controlling for all measures of investments, Column 3 shows that released listed movies generate significantly higher log(ROI)—where ROI (return over investment) is worldwide box office revenue divided by production budget—than unlisted ones generate. Columns 4-6 report quantile regression results for ROI, also controlling for investment. The Black List coefficients are all positive and statistically significant, confirming a statistically superior distribution for investment returns for listed movies.

Among Black-Listed scripts, the raw data show that the density distributions of log(Worldwide Box Office) and log(Production Budget) are statistically similar across all rank groups (see Figures 2b and Online Appendix Figure B2b). Panel (b) of Table 3 shows that the coefficients of all rank group dummies for regressions on ROI, though noisily estimated, are mostly positive relative to the default group of scripts receiving the threshold number of nominations.

4.1.2 Does The Black List Differentiate Idea Quality?

As discussed in Section 2.2.2, a positive correlation between release probability and Black List nominations is consistent with three interpretations: (i) the List is purely predictive; (ii) it is predictive and has a causal impact on decision making; and (iii) it influences outcomes without being predictive. Our results above on box office performance for released movies help to rule out the third possibility—that is, they reject the null hypothesis that "if the Black List is not predictive, conditional upon release, listed scripts and unlisted scripts should generate similar box office revenues, given the same physical and human capital allocations." It is important to note, however, that without knowledge of how studios rank projects internally prior to the List's publication, we are unable to conclude whether the Black List, while being predictive, also influences resource allocation in a causal sense.

4.2 The Black List and Writers of Different Experience Levels

This section presents the results on writer experience in two stages: first, how Black List nominations correlate with writer experience; and, second, how the relationship between nominations and release probability varies by writer experience. We discuss potential explanations and the implications after presenting the results.

4.2.1 Nominations and Writer Experience

The results suggest that the Black List is able to highlight ideas from writers that lack strong track records. In particular, Figure 3a shows that in Sample DDP, scripts from less-experienced writers are substantially more likely to be listed than those from experienced writers (13.5 vs. 7.0 percent, p-value = 0.002); Figure 3b shows that, if listed, the likelihood of achieving a top-20 ranking is also higher for less-experienced writers (27.6 vs. 14.4 percent, p-value = 0.002). Table 4 confirms this negative correlation between Black List nominations and writer experience with OLS regression results. Columns 1 and 2, using Sample DDP, show that scripts from experienced writers are about eight percentage points less likely to be Black-Listed, representing a 67-percent difference. Columns 3 and 4 show that, among Black-Listed scripts, those from experienced writers are 5.2 percentage points less likely to be in the top-five and 12.5 percentage points less likely to be in the top-20. Finally, Column 5 confirms this negative correlation using the ratio between the number of nominations and the total respondents in a year as the dependent variable.

4.2.2 Nominations, Release Probability, and Writer Experience

Figure 4 shows the following three patterns. First, similar to the average results, the positive correlation between Black List nominations and release probability also holds when considering scripts from less-experienced and experienced writers separately. Second, scripts from less-experienced writers are uniformly less likely to be released regardless of whether or not they are listed and, if listed, their rank group. Third, the gap in the release probability by writer experience remains large for listed scripts, even for the highest-ranked ones.

These patterns are confirmed by regression results presented in Table 5. Examining less-experienced and experienced writers in Sample DDP separately, Columns 1 and 2 show that Black-Listed scripts are significantly more likely to be released for both types of writers. The coefficient of Black-Listed for experienced writers is 13.5 percentage points greater than that for less-experienced writers. Given the noisier estimate, this difference, though economically large, is not statistically significant. Columns 3 and 4, using Sample BL and including 'nomination $\times$ survey year' fixed effects, compare scripts that receive exactly the same number of nominations in the same year. The results show that, for scripts listed below the top 20, the gap in release likelihood between experienced and less-experienced writers is 18 percentage points, representing about a 100-percent difference. The gap by writer experience becomes economically larger for top-20 scripts (by 9.3 or 14.5 percentage points, depending on the specification), though the difference between the gaps is not statistically significant.

4.2.3 Potential Explanations for Results by Writer Experience

This section discusses potential explanations for the results on writer experience. Note that any explanation must explain all of the following patterns: (i) the greater likelihood of being listed and ranking high for less-experienced writers; (ii) the lower likelihood of release for less-experienced writers relative to experienced ones, given the same number of nominations; and (iii) the persistence of this gap even among top-ranked scripts. In the following, we first propose an explanation, informed by Propositions 2-4 in Section 2.2.3, and then describe a potential alternative explanation.

Proposed Explanation: Script Visibility

Recall that the first part of each of the three propositions provides predictions when the breadth of script circulation is independent of the writer's experience level (i.e., $n'(w) = 0$ part of Assumption 2), whereas the second part is for when ideas from less-experienced writers are circulated more widely (i.e., $n'(w) < 0$ part of Assumption 2). According to these predictions, the best explanation of our results is that scripts from less-experienced writers are significantly more widely circulated than scripts from experienced writers (i.e., $n'(w)$ is sufficiently negative).

With significantly wider visibility, less-experienced writers would obtain more nominations despite a lower prior probability of success (the second part of Proposition 2). The wider visibility may also help to explain why the gap does not shrink for the highest-ranked scripts. The median number of nominations for top-20 ranked scripts is 20 votes. Though impossible to know, it seems reasonable that such numbers are likely to fall in the range of $m > m^*$ (as specified in the first part of Proposition 4); that is, the number of nominations is likely to produce relatively precise and positive signals of idea quality.[12] If so, we may expect observable quality signals to be less important and, hence, the gap by writer experience to start to shrink. Yet the data do not support this. Under the assumption that scripts from less-experienced writers are much more widely circulated, decision makers may discount the votes for less-experienced writers, which would sustain the gap even for the highest-ranked scripts.[13]

That the breadth of circulation is wider for scripts by less-experienced writers seems plausible, as it may be harder for them to find an interested buyer due to their lower (prior) probability of success. Moreover, though relatively selective, eligible voters are still likely to be heterogeneous in their experience and status in the industry. As mentioned, less-established producers/executives may be more willing to read materials from less-experienced writers, whereas more-established producers/executives may be more strict about the number of scripts they read and from which writers. Because all votes are treated equally, the likely segmentation in the market for scripts and the different levels of stringency in the voting criteria may further contribute to the greater number of votes for less-experienced writers.

The importance of visibility in determining the Black List nominations is consistent with the industry

[12]It is important to note that the typical number of Black List voters who have read any given script would be much smaller than the total number of respondents in a given year.

[13]Finally, Proposition 3 notes that both determinants of votes—prior probability π and script visibility n—predict that, given the same number of nominations, the likelihood of release is lower for less-experienced writers relative to experienced ones. This particular pattern is, thus, not discriminating between the two channels.

accounts quoted previously, such as: "[F]or a script to top that list, it needs to have been read by many of those people." It is also consistent with the results in Table 4, which show that the experience of the writer's agent and the size of the agency are, overall, positively correlated with Black List nominations. The positive association between agency size/agent experience and visibility is intuitive, as larger agencies and more-established agents are more likely to have the connections, resources, and reputation to get their client's script "into the hands of executives so that they may, in turn, like it and vote for it."

Of course, agent affiliation may correlate with other factors, as well; for example, writers associated with larger agencies are also likely to be higher-quality. To investigate this further, we conduct a split-sample analysis that fixes agency size and agent experience (Table 6). The idea is: (i) to restrict each subsample to writers who are likely to be similar in their overall quality; and (ii) to take advantage of the cross-subsample differences in the extent of resources, connections, and experience of writers' agents. The results reveal that the significant negative correlation between writer experience and Black List nominations occurs only for writers represented by large agencies, and the effect is the greatest for writers also represented by experienced agents.[14] Considering the baseline likelihoods (presented at the bottom of each column), *writers who receive the most nominations are the least experienced but associated with experienced agents at top agencies.* This cross-subsample pattern is consistent with the idea that the necessity for wider exposure needs to be complemented by the ability to reach the audience in order to achieve high visibility. Larger agencies and experienced agents can aid their less-experienced writers to obtain sufficient exposure, whereas smaller agencies and less-experienced agents are less able to do so.

Alternative Explanation: (Aesthetic) Novelty A second alternative explanation is that survey respondents nominate scripts that appeal to them for non-monetary reasons; for example, they may find scripts from less-experienced writers more novel or interesting, which will result in the Black List overrepresenting less-experienced writers. But these scripts may be less financially viable and, hence, less likely to be produced. We investigate this explanation in detail in the Online Appendix; available data and novelty measures used by prior work do not support the notion that scripts from less-experienced writers are systematically more novel.

4.2.4 Does The Black List Mitigate Barriers to Entry?

Our results suggest that the Black List helps to mitigate barriers to entry for less-experienced writers, but with some important caveats. Two key pieces of evidence support an affirmative answer. First, the finding that being listed is associated with a significantly higher release probability among less-experienced writers suggests that the List allows better scripts by less-experienced writers to stand out. In the absence of the List, all similar scripts from less-experienced writers are likely to be pooled together, and, due to the difficulty of differentiating among them, are equally less likely than those by experienced writers to be green-lighted by

[14] Almost all experienced agents work for a top-10 agency, so we do not split the non-top-10 agency sample by agent experience.

decision makers.

Second, the Black List does feature less-experienced writers more prominently. Given our explanation, less-established writers (albeit only those associated with large agencies) tend to have wider visibility among Black List voters. By casting a wider net and by giving equal weights to all votes, the Black List seems to have a *corrective effect*, in that those who need to sell their ideas more broadly in the normal circulation process may have a greater chance of making it to the List. Even if these scripts are not produced, such attention may prove valuable in the long run (e.g., obtaining future writing assignments).

There are two important caveats, however. First, the Black List, by design, builds upon the existing channels through which ideas circulate. Scripts that have a hard time obtaining an audience in the first place (in particular, those without an agent, let alone a top agent at a large agency) are unlikely to be substantially helped. Second, there may be unintended consequences when it comes to making inferences. Because the visibility of a specific script is not observable, decision makers may have to make crude adjustments based on coarse characteristics. Consequently, they may discount the nominations indiscriminately for less-experienced writers, even for those who have not shopped their scripts widely. Similarly, they may *not* discount the nominations for experienced writers, even if their scripts have had wide exposure.

To fully break down industry barriers, multiple approaches would likely need to work in combination, each addressing different segments and each with its different but inherent tradeoffs. The Black List annual survey that we study in the paper focuses on the mainstream segment of the industry—i.e., the ideas that have already been circulating among relatively successful producers. Given the demands on film executives' time, it would be difficult to impose a more-structured evaluation process without increasing the participation costs to a fatal extent. Recognizing these limitations, in 2012, the Black List began a separate platform on which any writer can host his or her script and have it rated by a reader, both for a nominal fee. In contrast to the survey, this platform provides much greater accessibility and transparency, and the evaluation is based on extensive, uniform criteria. But, because the platform focuses on a segment far outside the mainstream—91 percent of the hosted scripts are from writers without any representation either by an agency or a manager—it does not provide useful information about the set of scripts that is immediately relevant to the studios. Consequently, anecdotal evidence suggests that the key decision makers pay significantly less attention to the platform than they do to the annual survey. Examining the effect of the platform and how these different approaches could interact with each other is beyond the scope of this paper but would be a fascinating avenue for future research.

4.3 Generalizability

In discussing the generalizability of our findings, it is helpful to distinguish between two different questions: 1) what are the characteristics of settings in which a Black-List type method is likely to succeed? 2) are the mechanisms and effects in these other settings likely to be similar to what we observe in the movie industry?

4.3.1 Settings in Which a Black List-Type Method Could Succeed.

In its essence, the Black List survey aggregates, on a large scale, individual judgments about idea quality that are formed during the normal sales and development process of these ideas. We believe that similar methods could also apply to settings that share the following three characteristics: (i) idea quality is difficult to predict (even for experts) and hard to verify; (ii) assembling the judgment of a sizable number of experienced, but diverse evaluators via a structured process is costly and hard to scale; and (iii) many ideas are circulated in a decentralized manner, with individuals forming judgments (albeit noisy) in their normal course of business.

The first characteristic is applicable to many settings. In situations in which alternative means may produce relatively precise information (e.g., using physical experiments to test the viability of the technologies in science-based industries), the added value of such an information-aggregation approach may be limited. But the commercial viability of startups in domains such as consumer technologies or enterprise software, or even of science-based startups, often remains highly uncertain before investors make costly investments.[15]

The second characteristic comes from the simple fact that people—especially successful industry insiders whose opinions are likely to be seen as credible—are busy. Moreover, the further disclosure of ideas (to a large number of evaluators) may be undesirable. A key insight of the Black List, in our view, is that *it utilizes judgment that has already been formed in the marketplace.* Soliciting such judgments by requesting simple nominations can lower participation costs while requiring no significant additional exposure of ideas.

Finally, characteristic three—requiring the setting to have ideas in circulation upon which judgments are formed—is also relatively general. For example, to obtain financing, advice, and endorsement, founders often circulate the details of their startups among various types of experts, including mentors at incubators and accelerators, judges in startup competitions (often hosted by non-profit organizations such as universities), scientific and domain experts, angel investors, and executives at early-stage VC firms. Similarly, within large corporations, product and project ideas may also circulate in decentralized ways and encounter a nontrivial number of people through formal and informal processes.

In light of the above discussion, one can imagine conducting a Black List-like survey in the venture investment community, whereby a large number of participants nominate up to ten of their favorite startups that they have encountered in a given year, but which have not yet closed their current funding round (e.g., their series A). One key concern is that these voters may have strategic incentives to influence votes; for example, early investors may want to promote their investments to follow-on investors. But similar concerns are also raised with respect to the Black List: producers and agents may vote for their own projects and encourage friends to support their favorites. The opportunity to promote projects is, in itself, a potentially

[15]For example, using internal data from a large, successful venture capital firm, Kerr et al. (2014) find that the correlation between the partners' ratings of every deal at the time of the first investment and the startup's ultimate performance was 0.1, suggesting that even successful professionals struggle to distinguish the most promising startups at the earliest stages of investment.

powerful incentive to participate in such surveys. How strong these strategic incentives are and how much they may threaten the usefulness of the aggregated signals are likely to be context-dependent and must be taken seriously. In our setting, at least, the Black List appears to be able to further differentiate similar ideas despite these strategic incentives.

4.3.2 Should We Expect Similar Findings in Other Settings?

Because a Black List-type method depends on judgments formed in the decentralized marketplace, the particular effects that such methods may have, as well as their mechanisms, will depend on how ideas circulate in specific settings. Our results suggest that the visibility of an idea among survey respondents is a systematic determinant of votes received, and, in the movie industry, it seems to be wider for less-experienced providers (i.e., $n'(w) < 0$). What might we expect if, instead, the relationship between idea visibility and seller experience is positive (i.e., Assumption 2 is violated and $n'(w) > 0$)?

Our conceptual framework suggests which results we may expect to be similar and which may change. Proposition 1 holds regardless of the relationship between visibility and experience, because, by comparing *within* sellers of similar characteristics (i.e., fixing w), we effectively hold their $\pi(w)$ and $n(w)$ constant. Thus, the first key objective of such methods—to further differentiate ideas that are observably similar—still holds under the basic premise that the voters' judgments are informative. Also similar to The Black List, this effect helps to mitigate barriers to entry, because it allows better ideas among those by less-experienced sellers to stand out instead of being pooled together with worse ones and being under-funded.

Propositions 2 and 3 will be different, and these differences have two implications.[16] The first is with respect to the types of sellers that are more likely to be nominated. Equation (3) shows that in situations in which $n'(w) > 0$, the 'corrective' effect—which, in The Black List, highlights less-experienced sellers who had to sell harder in the idea-circulation stage—would disappear. Instead, we would expect to see a 'rich gets richer' situation, in that the more-established sellers will receive additional nominations due to their wider visibility.[17] Thus, to the extent that being nominated (and highly ranked) proves valuable in the long run (e.g., obtaining future writing assignments), the Black List method would be less helpful for less-experienced sellers via this channel.

The second implication regards whose ideas should be discounted. Equation (4) shows that if $n'(w) > 0$, for ideas receiving the same nominations, one should discount those from more-experienced sellers, to some extent, because of the expectation that these nominations are partially due to higher visibility.[18] This result is

[16]Proposition 4 considers secondary derivatives and has less direct implications than Propositions 2 and 3. For simplicity, we do not discuss the potentially different dynamics implied by this result.

[17]The second term in Equations (3), $\frac{\partial \mathrm{E}[m|w]}{\partial n(w)} n'(w)$, is positive if $n'(w) > 0$. Thus, the expected number of nominations would be greater for experienced than for less-experienced idea sellers, which is opposite to our result with the Black List.

[18]If $n'(w) > 0$, the second term in equation (4) is negative, making $\frac{\partial \mathrm{P}(q_H|m,w)}{\partial w}$ ambiguous. Thus, for ideas receiving the same nominations, it is even possible that those from less-experienced sellers could be *more* likely to be inferred as having a high quality.

in contrast to the Black List, where less-experienced writers appear to be discounted. As we discuss in Section 4.2.4, this channel would, to some extent, help less-experienced sellers when decision makers do not observe the visibility of a specific idea and, hence, do *not* discount the nominations for less-experienced sellers, even if their ideas have had wide exposure.

More generally, in order to interpret data generated by a Black List-type method, one needs to understand how individual judgments are formed in the decentralized marketplace in the first place—e.g., which types of voters are included in the survey; what types of startups these voters typically encounter and how; and which voters tend to see which ideas. How this process works can only be understood in context. Factors such as how the survey questions are framed and the scope of voters may imply different idea visibility-seller experience relationships, even within the same industry setting. In the startup context, for example, a survey that asks voters to nominate startups that have pitched to them or that they have mentored will likely capture startups, idea-circulation patterns, and quality of judgment differently from a survey that seeks nominations of startups that the voters are just generally aware of. Similarly, a more-inclusive survey scope—for example, including mentors and judges from a relatively large and diverse number of accelerators, startup competitions, and venture funds—will capture a set of startups that selects to participate in these institutions, which will be different from startups to which a more-exclusive scope of voters likely have access.

5 Conclusion

In this paper, we study a novel, low-cost approach to solicit and aggregate opinions from a large number of industry experts on ideas that they encounter in their normal course of business. Our empirical context is the movie industry, in which investment costs are high and consumer tastes are difficult to predict. The Black List, an annual publication, ranks unproduced scripts based on the number of anonymous nominations they receive from film executives. There are trade-offs to such an approach: on the one hand, it requires a minimal time commitment that leads to a high participation rate by knowledgable but busy voters; on the other hand, the unstructured process and lack of standard criteria, as well as the fact that one cannot observe which voters see which ideas, are risks that may reduce the power of information aggregation and make the aggregated votes hard to interpret.

We find that, despite the above challenges, the Black List helps to further differentiate quality among observably similar ideas. This finding is important, as it suggests that such a method may aid evaluation in settings that share characteristics with the film industry, as discussed in the previous section. One potential limitation, revealed by our results on writer experience, is that the visibility of ideas among voters is an important determinant of votes, which may be influenced by various factors or frictions and not be easily observable. Determinants of visibility, and their consequences, are likely to be context-dependent. Although potential limitations of this feature may be mitigated by improving the process, structure, or transparency of the aggregation methods, one needs to be mindful of the tradeoffs implied by the necessity to keep participation time- and cost-effective.

References

Agrawal, A., C. Catalini, and A. Goldfarb (2015). Crowdfunding: Geography, social networks, and the timing of investment decisions. *Journal of Economics and Management Strategy 24*, 253–274.

Arora, A., A. Fosfuri, and A. Gambardella (2001). *Markets for Technology: The Economics of Innovation and Corporate Strategy*. Cambridge, MA: MIT Press.

Arrow, K. J. (1962). *The Rate and Direction of Inventive Activity: Economic and Social Factors*, Chapter Economic Welfare and the Allocation of Resources for Invention, pp. 609–626. Princeton University Press.

Astebro, T. and S. Elhedhli (2006). The effectiveness of simple decision heuristics: Forecasting commercial success for early-stage ventures. *Management Science 52*, 395–409.

Azoulay, P., T. Stuart, and Y. Wang (2013). Matthew: Effect or fable? *Management Science 60*, 92–109.

Basuroy, S., S. Chatterjee, and S. A. Ravid (2003). How Critical are Critical Reviews? The Box Office Effects of Film Critics, Star Power, and Budgets. *Journal of Marketing 67*(4), 103–117.

Bates, J. M. and C. W. Granger (1969). The combination of forecasts. *Journal of the Operational Research Society 20*(4), 451–468.

Boudreau, K. J., E. C. Guinan, K. R. Lakhani, and C. Riedl (2016). Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. *Management Science 62*, 2765–2783.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011). Robust inference with multiway clustering. *Journal of Business and Economic Statistics 29*, 238–249.

Caves, R. (2000). *Creative Industries: Contracts Between Art and Commerce*. Cambridge, MA: Harvard University Press.

Condorcet, J. A. N. (1785). *Foundations of social choice and political theory*. (Translated and edited by I. McLean and F. Hewitt in 1994) Brookfield, VT: E. Elgar.

Cowgill, B. and E. Zitzewitz (2015). Corporate prediction markets: Evidence from google, ford, and firm x. *The Review of Economic Studies 82*(4), 1309–1341.

Csaszar, F. A. (2018). Limits to the wisdom of the crowd in idea selection. *Organization Design.*, 275–297.

Eliashberg, J. and S. M. Shugan (1997). Film Critics: Influencers or Predictors? *Journal of Marketing 61*(2), 68–78.

Gans, J. S. and S. Stern (2003). The product market and the market for "ideas": Commercialization strategies for technology entrepreneurs. *Research Policy 32*, 333–350.

Goetzmann, W. N., S. A. Ravid, and R. Sverdlove (2013). *The pricing of soft and hard information: Economic lessons from screenplay sales*, Volume 37.

Kaplan, S. N., B. A. Sensoy, and P. Stromberg (2009). Should investors bet on the jockey or the horse? evidence from the evolution of firms from early business plans to public companies. *Journal of Finance 64*, 75–115.

Kerr, W. R., R. Nanda, and M. Rhodes-Kropf (2014). Entrepreneurship as experimentation. *Journal of Economic Perspectives 28*, 25–48.

Larrick, R. P. and J. B. Soll (2006). Intuitions about combining opinions: Misappreciation of the averaging principle. *Management Science 52*(1), 111–112.

Li, D. (2017). Expertise versus bias in evaluation: Evidence from the nih. *American Economic Journal: Applied Economics 9*, 60–92.

Luo, H. (2014). When to Sell Your Idea : Theory and Evidence from the Movie Industry. *Management Science 60*(12), 3067–3086.

Mollick, E. (2014). The dynamics of crowdfunding: Determinants of success and failure. *Journal of Business Venturing 29*(1), 1–16.

Mollick, E. and R. Nanda (2015). Wisdom or Madness? Comparing Crowds with Expert Evaluation in Funding the Arts. *Management Science 62*(6), 1533–1553.

Riedl, C., I. Blohm, J. Leimeister, and H. Krcmar (2013). The effect of rating scales on decision quality and user attitudes in online innovation communities. *International Conference on Information Systems, St. Louis, Mi, USA 17*, 7–36.

Scott, E., P. Shu, and R. Lubynsky (2020). Entrepreneurial uncertainty and expert evaluation: An empirical analysis. *Management Science 66*, 1005–1507.

Simcoe, T. S. and D. M. Waguespack (2011). Status, quality, and attention: What's in a (missing) name? *Management Science 57*, 274–290.

Sorenson, B. O., V. Assenova, G.-c. Li, and J. Boada (2016). Expand Innovation Finance via Crowdfunding. *Science 354*, 1526–1528.

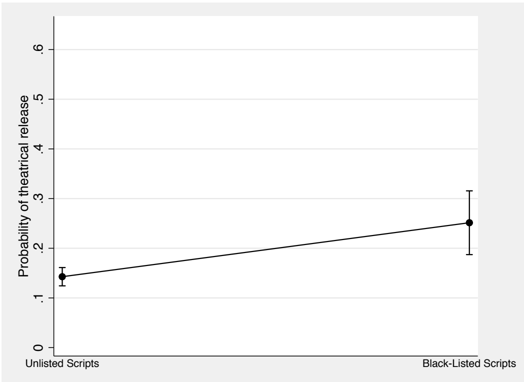
Stuart, T. E. and O. Sorensen (2010). Syndication networks and the spatial distribution of venture capital investments. *American Journal of Sociology 106*, 1546–1588.

Surowiecki, J. (2005). *The wisdom of crowds*. Anchor.

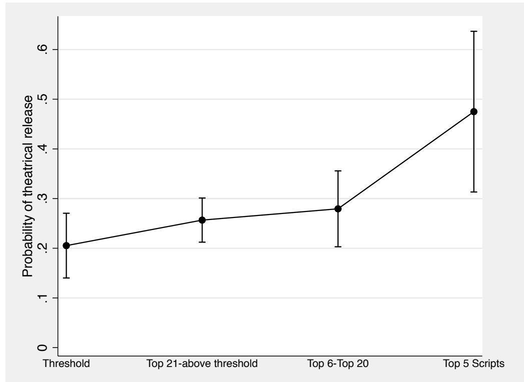
Waldfogel, J. (2017). How digitization has created a golden age of music, movies, books, and television. *Journal of Economic Perspectives 31*(3), 195–214.

Figure 1: Probability of Theatrical Release by Black List Nominations

(a) Sample DDP



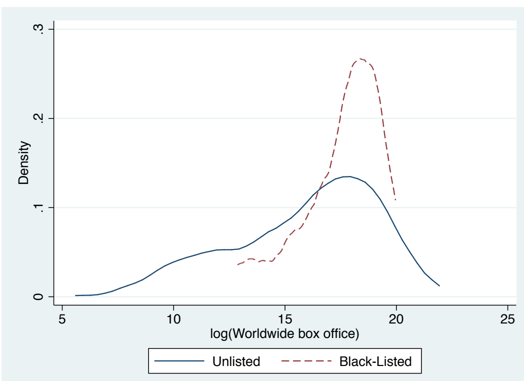
(b) Sample BL



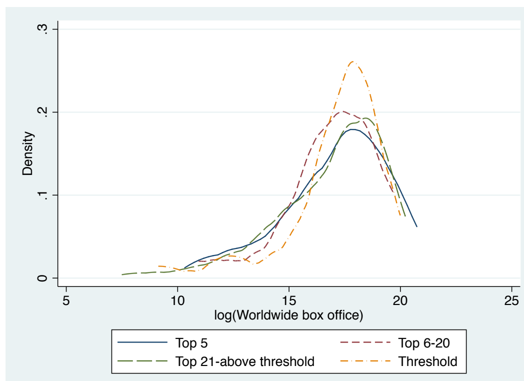
Note: Panel (a) uses Sample DDP and compares theatrical release probabilities for listed and unlisted scripts. Panel (b) uses Sample BL and compares theatrical release probabilities by Black List Rank Group. The figures plot the 95-percent confidence intervals.

Figure 2: Worldwide Box Office Revenues by Black List Nominations

(a) Sample DDP



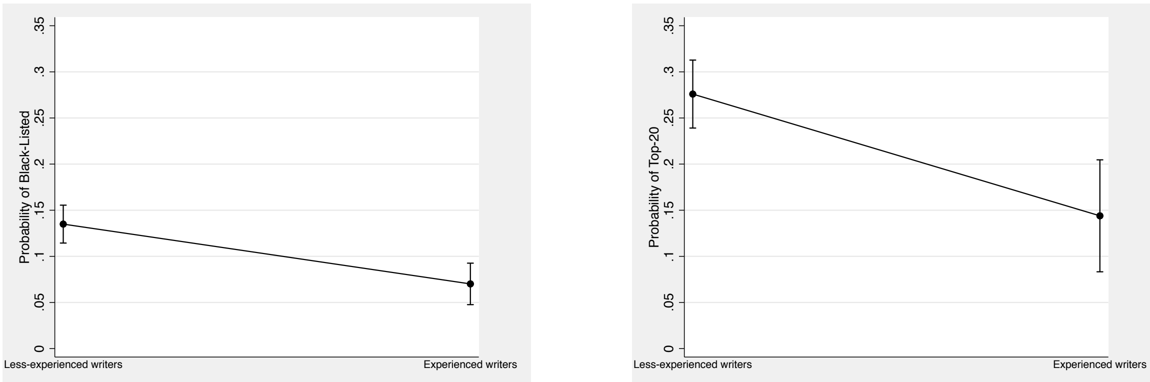
(b) Sample BL



Note: Panel (a) uses Sample DDP and plots the density distribution of log(Worldwide Box Office) for released movies by whether or not the script is Black-Listed, and Panel (b) plots the same distribution for Sample BL by Black List Rank Group.

Figure 3: Probability of Being Black-Listed (or Top-20 if Listed) by Writer Experience

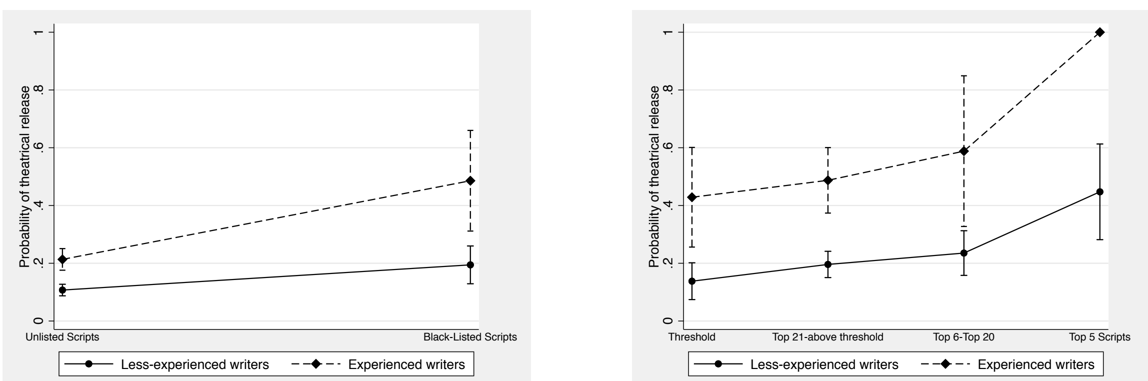
(a) Sample DDP | (b) Sample BL



Note: Panel (a) uses Sample DDP and compares the likelihood of being Black-Listed by writer experience. Panel (b) uses Sample BL and compares the likelihood of a top-20 Black List ranking by writer experience. Experienced writers are those with one or more major writing credits in the previous ten years; less-experienced writers are those with zero major writing credits in the previous ten years. The figures plot the 95-percent confidence intervals.

Figure 4: Probability of Theatrical Release by Black List Nominations and Writer Experience

(a) Sample DDP | (b) Sample BL



Note: Panel (a) uses Sample DDP and compares theatrical release probabilities for listed and unlisted scripts and by writer experience. Panel (b) uses Sample BL and compares theatrical release probabilities by Black List Rank Group and by writer experience. Experienced writers are those with one or more major writing credits in the previous ten years; less-experienced writers are those with zero major writing credits in the previous ten years. The figures plot the 95-percent confidence intervals. Note that in Panel (a), the difference in the slopes by writer experience is statistically significant at the one-percent level.

Table 1: Summary Statistics

	Sample BL			.	Sample DDP			
	N	Mean	S.D.	Median	N	Mean	S.D.	Median
Black List Rank Group	701	2.91	0.79	3				
Nomination Ratio	701	0.05	0.04	0.03				
Black-Listed					1566	0.11	0.32	0
Theatrically Released	701	0.26	0.44	0	1566	0.16	0.36	0
Writer Major Credits	701	0.36	0.93	0	1566	0.62	1.12	0
Experienced Writer	701	0.19	0.39	0	1566	0.32	0.47	0
Top-10 Agency	701	0.91	0.28	1	1566	0.54	0.50	1
Experienced Agent	701	0.34	0.48	0	1566	0.16	0.37	0
Sold Prior to Publication	701	0.79	0.41	1				
Producers Major Credits	701	5.11	6.21	3	1566	5.72	6.13	4
Movie Studio Buyer	701	0.37	0.48	0	1566	0.37	0.48	0
Whether Reported	701	0.09	0.28	0	1566	0.11	0.32	0
Original					1566	0.58	0.49	1
Bidding					1566	0.03	0.16	0
Talent Attached					1566	0.57	0.50	1
<u>Released movies</u>								
Worldwide Box Office ($m)	184	89.20	121.08	40.78	243	102.81	192.79	24.66
Budget ($m)	152	33.22	34.67	23.04	163	45.59	46.94	27.21
Star Score	184	103.11	111.35	85	243	67.73	100.01	0
Open Screen	184	1654.39	1421.06	2174	243	1648.72	1534.04	2027
Award-Winning Cast	184	0.63	0.49	1	243	0.56	0.50	1
Seasonality	184	0.70	0.15	0.70	243	0.70	0.14	0.70
Franchise	184	0.07	0.26	0	243	0.12	0.33	0

Note: The left panel provides summary statistics for Sample BL; the right panel provides the summary statistics for Sample DDP. Section 3.2 provides a description of the variables. Dummy variables for genre and content source materials are not summarized in the table.

Table 2: Black List Nominations and Release Probability

	Sample DDP			Sample BL		
			Exclude Top-20			
	Release	Release	Release	Release	Release	Release
	(1)	(2)	(3)	(4)	(5)	(6)
Black-Listed	0.135**	0.120**	0.109**			
	(0.047)	(0.039)	(0.042)			
Top-5				0.295***	0.280***	
				(0.055)	(0.048)	
Top 6-20				0.090	0.093	
				(0.072)	(0.058)	
Top 21-Above Threshold				0.065	0.060	
				(0.043)	(0.044)	
Nomination Ratio						1.536***
						(0.006)
Writer Major Credits	0.039	0.034	0.036	0.083***	0.086***	0.088***
	(0.022)	(0.022)	(0.023)	(0.008)	(0.010)	(0.013)
Top 10 Agency	0.055**	0.060***	0.063***	-0.006	0.014	0.014
	(0.018)	(0.016)	(0.016)	(0.018)	(0.020)	(0.018)
Experienced Agent	-0.031	-0.020	-0.017	-0.033	-0.032	-0.026
	(0.024)	(0.023)	(0.020)	(0.025)	(0.035)	(0.034)
Producer Major Credits	0.006*	0.005	0.005	0.005	0.004	0.004
	(0.003)	(0.004)	(0.003)	(0.004)	(0.005)	(0.006)
Original	0.040	-0.112	0.031			
	(0.029)	(0.086)	(0.028)			
Talent Attached	0.099***	0.098***	0.098***			
	(0.018)	(0.021)	(0.025)			
Bidding	0.039	0.016	-0.015			
	(0.076)	(0.083)	(0.104)			
Whether Reported	0.045*	0.061**	0.061*	-0.026	-0.090	-0.078*
	(0.020)	(0.021)	(0.029)	(0.021)	(0.057)	(0.036)
Sold Prior to Publication				0.134***	0.140***	0.136**
				(0.023)	(0.033)	(0.045)
Genre FE	Y	Y	Y			
Content Source FE	Y	Y	Y			
Studio FE	Y	N	N	Y	N	N
Year FE	Y	N	N	Y	N	N
Studio×year FE	N	Y	Y	N	Y	Y
N	1565	1537	1473	701	659	659
R-squared	0.113	0.163	0.162	0.213	0.268	0.272
Mean(DV|Baseline)	0.143	0.141	0.142	0.205	0.190	0.190

Note: OLS regressions. The dependent variable is whether the script is theatrically released. Columns 1-3 use Sample DDP, with Column 3 excluding top-20 listed scripts. Columns 4-6 use Sample BL; the excluded group in Columns 4-5 is scripts receiving the threshold number of nominations, and Column 6 uses the ratio of nominations received by the script divided by the total number of respondents in the year. The baseline group for Columns 1-3 consists of unlisted scripts; and the baseline group for Columns 4-6 consists of scripts receiving threshold nominations. Standard errors (in parentheses) in all columns are clustered in two dimensions: year and studio. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Box Office Performance and Production Budget

(a) Sample DDP

	OLS regressions			Quantile regressions		
				25th	50th	75th
	log(Box Office)	log(Budget)	log(ROI)	ROI	ROI	ROI
	(1)	(2)	(3)	(4)	(5)	(6)
Black-Listed	1.196**	-0.160	1.275***	0.529**	0.903*	1.961***
	(0.389)	(0.296)	(0.192)	(0.210)	(0.522)	(0.415)
log(Budget)			0.106	0.016	0.224	0.023
			(0.127)	(0.056)	(0.264)	(0.212)
Other controls	Y	Y	Y	Y	Y	Y
Other investment controls	N	N	Y	Y	Y	Y
Release Year FE	Y	Y	Y	Y	Y	Y
Studio FE	Y	Y	Y	Y	Y	Y
N	239	158	153	163	163	163
R-squared	0.580	0.590	0.469	0.446	0.490	0.493

(b) Sample BL

	OLS regressions			Quantile regressions		
				25th	50th	75th
	log(Box Office)	log(Budget)	log(ROI)	ROI	ROI	ROI
	(1)	(2)	(3)	(4)	(5)	(6)
Top 5	0.577	0.594	0.178	0.895*	0.310	0.338
	(0.525)	(0.399)	(0.266)	(0.468)	(0.378)	(1.082)
Top 6-20	0.515	0.010	0.440	0.200	0.093	0.660
	(0.440)	(0.258)	(0.598)	(0.500)	(0.824)	(1.501)
Top 21-Above Threshold	0.098	0.008	0.273	0.439	0.719*	0.616
	(0.383)	(0.327)	(0.264)	(0.374)	(0.379)	(0.408)
log(Budget)			-0.220***	-0.297	-0.333*	-0.658***
			(0.046)	(0.187)	(0.170)	(0.171)
Other controls	Y	Y	Y	Y	Y	Y
Other investment controls	N	N	Y	Y	Y	Y
Release Year FE	Y	Y	Y	Y	Y	Y
Studio FE	Y	Y	Y	Y	Y	Y
N	177	143	143	152	152	152
R-squared	0.440	0.451	0.293	0.156	0.321	0.310

Note: Panel (a) uses released movies for Sample DDP, and Panel (b) uses released movies for Sample BL. Columns 2-6 in both panels use released movies for which production budget information is available. ROI is defined as the ratio between worldwide box office revenues and production budget. In Panel (b), the excluded group includes scripts receiving the threshold number of nominations. Other controls for panel (b) include Writer Major Credits; Top-10 Agency; Experienced Agent; Producer Major Credits; Whether Reported; Sold prior to Publication; Franchise; and Genre dummies. Other controls for panel (a) are the same as panel (b) plus Original, Talent Attached, and Bidding. Other investment controls include Star Score; Award-Winning Cast; Open Screen; and Seasonality. For both panels, standard errors in columns 1-3 are clustered in two dimensions: year and studio; and those in columns 4-6 (quantile regressions) are clustered by studio. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Black List Nominations and Writer Experience

	Sample DDP		Sample BL		
	Black-Listed	Black-Listed	Top-5	Top-20	Nomination Ratio
	(1)	(2)	(3)	(4)	(5)
Experienced Writer	-0.083**	-0.087**	-0.052*	-0.125**	-0.012***
	(0.025)	(0.027)	(0.027)	(0.037)	(0.002)
Top-10 Agency	0.076**	0.074*	0.035***	0.101*	0.009*
	(0.030)	(0.032)	(0.007)	(0.048)	(0.004)
Experienced Agent	0.097**	0.092**	-0.007	-0.006	-0.002
	(0.038)	(0.036)	(0.015)	(0.030)	(0.003)
Producer Major Credits	0.003	0.003	0.001	0.002	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Whether Reported	0.096***	0.094***	0.090**	0.084	0.009
	(0.026)	(0.026)	(0.037)	(0.079)	(0.007)
Original	0.108*	0.114*			
	(0.047)	(0.051)			
Talent Attached	-0.023	-0.02			
	(0.018)	(0.018)			
Bidding	0.236***	0.226***			
	(0.052)	(0.059)			
Movie Studio Buyer	0.043				
	(0.024)				
Sold Prior to Publication			0.008	0.105***	0.009**
			(0.012)	(0.023)	(0.003)
Genre FE	Y	Y			
Content Source FE	Y	Y			
Year FE	Y	Y	Y	Y	Y
Studio FE	N	Y	Y	Y	Y
N	1566	1565	700	700	700
R-squared	0.168	0.186	0.063	0.072	0.102
Mean(DV|Exp writer =0)	0.135	0.135	0.067	0.276	0.048

Note: OLS (linear probability) regressions. Columns 1-2 use Sample DDP; Columns 3-5 use Sample BL. The dependent variable in Columns 1 and 2 is an indicator of whether the script is Black-Listed. The dependent variables of Column 3 and 4 are, respectively, indicators of whether the script achieves top-5 or top-20 Black List ranking in the survey year; the dependent variable in Column 5 is the ratio of nominations received by the script divided by the total number of respondents in the year. Experienced writers are defined as those with one or more major writing credits in the previous ten years. Standard errors (in parentheses) are clustered in two dimensions: year and studio. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Black List Nominations and Release Probability, by Writer Experience

	Sample DDP		Sample BL	
	Less-Experienced Writers	Experienced Writers		
	Release	Release	Release	Release
	(1)	(2)	(3)	(4)
Black-Listed	0.088**	0.223*		
	(0.037)	(0.116)		
Experienced Writer			0.178**	0.190***
			(0.056)	(0.049)
Top 20 × Experienced Writer			0.145	0.093
			(0.122)	(0.112)
Studio FE			N	Y
Year×Nominations FE			Y	Y
N	1023	459	611	608
R-squared	0.145	0.233	0.318	0.358
Mean(DV|Baseline)	0.188	0.438	0.178	0.178

Note: Columns 1 and 2 use, respectively, the subset of scripts from less-experienced writers and the subset by experienced writers in Sample DDP, and Columns 3-4 use Sample BL. The dependent variable in all columns is an indicator of whether the movie is theatrically released. Columns 1-2 include the same set of controls as in Column 2 in Table 2, including studio×year fixed effects; Columns 3-4 include the same set of controls as in Column 4 in Table 2, except for year fixed effects, as they are absorbed by Year×Nominations fixed effects. The baseline group for Columns 1 and 2 is unlisted scripts; and that for Columns 3 and 4 are less-experienced writers. Standard errors are clustered in two dimensions: year and studio. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Black List Nominations and Writer Experience: Subsample by Agency Size and Agent Experience

	Top-10 Agency = 0	Top-10 Agency = 1	
		Experienced Agent = 0	Experienced Agent = 1
	(1)	(2)	(3)
Panel 1: Sample DDP	Black-Listed	Black-Listed	Black-Listed
Experienced Writer	-0.019	-0.086*	-0.219**
	(0.017)	(0.041)	(0.078)
N	707	601	244
R-squared	0.284	0.128	0.401
Mean (DV| Exp writer = 0)	0.09	0.22	0.45
Panel 2: Sample BL	Top 20	Top 20	Top 20
Experienced Writer	-0.073	-0.092	-0.165**
	(0.173)	(0.069)	(0.059)
N	62	401	224
R-squared	0.176	0.094	0.069
Mean (DV| Exp writer = 0)	1.81	2.12	2.26

Note: Panel (a) uses Sample DDP. The dependent variable indicates whether the script is Black-Listed. All regressions include the same controls as in Column 2 in Table 4. Panel (b) conducts the same analysis as in Panel (a), using Sample BL. The dependent variable indicates top-20 ranked scripts on the Black List. All regressions include the same controls as in Column 3 in Table 4. Standard errors (in parentheses) are clustered in two dimensions: year and studio. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Online Appendices (not for publication)

Online Appendix A. Proofs of the Propositions

Before presenting the proofs of the propositions, first recall the basic setup of the model. The inherent quality of a movie script is either high or low: $q \in \{q_H, q_L\}$. A voter, having read a given script, obtains a binary quality signal $s \in \{s_H, s_L\}$. Let $\mathrm{P}(s_H|q_H) = p_1$ and $\mathrm{P}(s_H|q_L) = p_2$. Assume that $p_1 > p_2$; that is, the likelihood of obtaining a positive signal is greater when the true quality is high than when the true quality is low. Moreover, any two voters' signals are conditionally independent given the true script quality. A voter will nominate a script as long as a positive signal (s_H) is received. For a script from a writer of experience level w, the prior probability of high quality is $\pi(w)$, and the number of Black List voters that have read the script is $n(w)$. According to Assumptions 1 and 2, $\pi'(w) > 0$ and $n'(w) \leq 0$.

Given this set up, the expected number of nominations for a script from a writer of experience level w is (that is, equation (1)):

$$\mathrm{E}[m|w] = \pi(w)n(w)p_1 + (1-\pi(w))n(w)p_2.$$

Based on Bayes' rule, for a script that has received m nominations and is from writer of experience level w, the updated belief about the probability that it is of high quality is (that is, equation (2)):

$$\mathrm{P}(q_H|m,w) = \frac{\pi(w)\mathrm{P}(m|w,q_H)}{\pi(w)\mathrm{P}(m|w,q_H) + (1-\pi(w))\mathrm{P}(m|w,q_L)}.$$

Note that the denominator in the above equation—the probability that a script from writer of experience level w receives m nominations—can be written as:

$$\begin{aligned} & \pi(w)\mathrm{P}(m|w,q_H) + (1-\pi(w))\mathrm{P}(m|w,q_L) \\ = \; & \pi(w)\tbinom{n(w)}{m}p_1^m(1-p_1)^{(n(w)-m)} + (1-\pi(w))\tbinom{n(w)}{m}p_2^m(1-p_2)^{(n(w)-m)}. \end{aligned}$$

In words, when the true script quality is q_H, there are $\binom{n(w)}{m}$ ways to have m voters receive positive signals s_H, each with a probability of $p_1^m(1-p_1)^{(n(w)-m)}$. When the true script quality is q_L, there are also $\binom{n(w)}{m}$ ways to have m voters receive positive signals, each with a probability of $p_2^m(1-p_2)^{(n(w)-m)}$. We weight these two states by their respective prior probabilities, $\pi(w)$ and $(1-\pi(w))$. Thus, $\mathrm{P}(q_H|m,w)$ can be rewritten as:

$$\mathrm{P}(q_H|m,w) \quad = \quad \frac{\pi(w)p_1^m(1-p_1)^{(n(w)-m)}}{\pi(w)p_1^m(1-p_1)^{(n(w)-m)} + (1-\pi(w))p_2^m(1-p_2)^{(n(w)-m)}}, \qquad \text{(A1)}$$

where $\binom{n(w)}{m}$ is canceled out from all terms.

Proof of Proposition 1

Proof. We need to show that $\mathrm{P}(q_H|m,w)$ is an increasing function of m, holding w fixed. Taking the derivative of $\mathrm{P}(q_H|m,w)$ (in equation (A1)) with respect to m yields:

$$\begin{aligned}
&\frac{\partial \mathrm{P}(q_H|m,w)}{\partial m} \\
= \;& \frac{\frac{\partial(\pi(w)p_1^m(1-p_1)^{(n(w)-m)})}{\partial m}(1-\pi(w))p_2^m(1-p_2)^{(n(w)-m)} - \pi(w)p_1^m(1-p_1)^{(n(w)-m)}\frac{\partial((1-\pi(w))p_2^m(1-p_2)^{(n(w)-m)})}{\partial m}}{[\pi(w)p_1^m(1-p_1)^{(n(w)-m)} + (1-\pi(w))p_2^m(1-p_2)^{(n(w)-m)}]^2} \\
= \;& \frac{\pi(w)(1-\pi(w))p_1^m(1-p_1)^{(n(w)-m)}p_2^m(1-p_2)^{(n(w)-m)}[\ln p_1 - \ln(1-p_1) - (\ln p_2 - \ln(1-p_2))]}{[\pi(w)p_1^m(1-p_1)^{(n(w)-m)} + (1-\pi(w))p_2^m(1-p_2)^{(n(w)-m)}]^2} > 0
\end{aligned}$$

because $\ln p_1 - \ln(1-p_1) > \ln p_2 - \ln(1-p_2)$ when $p_1 > p_2$. □

Proof of Proposition 2

Proof. As is illustrated in equation (3), we need to show that $\mathrm{E}[m|w]$ is an increasing function of $\pi(w)$ and also an increasing function of $n(w)$. First,

$$\frac{\partial \mathrm{E}[m|w]}{\partial \pi(w)} = n(w)p_1 - n(w)p_2 > 0,$$

because $p_1 > p_2$. Second,

$$\frac{\partial \mathrm{E}[m|w]}{\partial n(w)} = \pi(w)p_1 + (1-\pi(w))p_2 > 0.$$

□

Proof of Proposition 3

Proof. As is illustrated in equation (4), we need to show that $\mathrm{P}(q_H|m,w)$ is an increasing function of $\pi(w)$ and also a decreasing function of $n(w)$.

Take the derivative of $\mathrm{P}(q_H|m,w)$ with respect to $\pi(w)$, we get:

$$\begin{aligned}
&\frac{\partial \mathrm{P}(q_H|m,w)}{\partial \pi(w)} \\
= \;& \frac{\frac{\partial(\pi(w)p_1^m(1-p_1)^{(n(w)-m)})}{\partial \pi(w)}(1-\pi(w))p_2^m(1-p_2)^{(n(w)-m)} - \pi(w)p_1^m(1-p_1)^{(n(w)-m)}\frac{\partial((1-\pi(w))p_2^m(1-p_2)^{(n(w)-m)})}{\partial \pi(w)}}{[\pi(w)p_1^m(1-p_1)^{(n(w)-m)} + (1-\pi(w))p_2^m(1-p_2)^{(n(w)-m)}]^2} \\
= \;& \frac{p_1^m(1-p_1)^{(n(w)-m)}p_2^m(1-p_2)^{(n(w)-m)}}{[\pi(w)p_1^m(1-p_1)^{(n(w)-m)} + (1-\pi(w))p_2^m(1-p_2)^{(n(w)-m)}]^2} > 0.
\end{aligned} \tag{A2}$$

The derivative of $\mathrm{P}(q_H|m,w)$ with respect to $n(w)$ is:

$$\begin{aligned}
&\frac{\partial \mathrm{P}(q_H|m,w)}{\partial n(w)} \\
= \;& \frac{\frac{\partial(\pi(w)p_1^m(1-p_1)^{(n(w)-m)})}{\partial n(w)}(1-\pi(w))p_2^m(1-p_2)^{(n(w)-m)} - \pi(w)p_1^m(1-p_1)^{(n(w)-m)}\frac{\partial((1-\pi(w))p_2^m(1-p_2)^{(n(w)-m)})}{\partial n(w)}}{[\pi(w)p_1^m(1-p_1)^{(n(w)-m)} + (1-\pi(w))p_2^m(1-p_2)^{(n(w)-m)}]^2} \\
= \;& \frac{\pi(w)(1-\pi(w))p_1^m(1-p_1)^{(n(w)-m)}p_2^m(1-p_2)^{(n(w)-m)}(\ln(1-p_1) - \ln(1-p_2))}{[\pi(w)p_1^m(1-p_1)^{(n(w)-m)} + (1-\pi(w))p_2^m(1-p_2)^{(n(w)-m)}]^2} < 0,
\end{aligned} \tag{A3}$$

because $\ln(1-p_1) < \ln(1-p_2)$ when $p_1 > p_2$.

□

Proof of Proposition 4

Proof. Taking the derivative of $\frac{\partial \mathrm{P}(q_H|m,w)}{\partial w}$ with respect to m, we get

$$\begin{aligned}
\frac{\partial^2 \mathrm{P}(q_H|m,w)}{\partial w \partial m} &= \frac{\partial^2 \mathrm{P}(q_H|m,w)}{\partial m \partial w} \\
&= \frac{\partial(\frac{\partial \mathrm{P}(q_H|m,w)}{\partial m})}{\partial \pi(w)}\pi'(w) + \frac{\partial(\frac{\partial \mathrm{P}(q_H|m,w)}{\partial m})}{\partial n(w)}n'(w) \\
&= \frac{\partial^2 \mathrm{P}(q_H|m,w)}{\partial \pi(w) \partial m}\pi'(w) + \frac{\partial^2 \mathrm{P}(q_H|m,w)}{\partial n(w) \partial m}n'(w).
\end{aligned} \tag{A4}$$

For the first part of Proposition 4—when scripts from writers of different experience levels are read by the same number of Black List voters (i.e., $n'(w) = 0$)—we consider only the first term in the above equation. Because $\pi'(w) > 0$ under Assumption 1, the sign of the first term in the above equation would depend on $\frac{\partial^2 \mathrm{P}(q_H|m,w)}{\partial \pi(w) \partial m}$. From equation (A2), we can show that

$$\begin{aligned}
\frac{\partial^2 \mathrm{P}(q_H|m,w)}{\partial w \partial m} \leq 0 \quad &\Leftrightarrow \quad \pi(w)p_1^m(1-p_1)^{(n(w)-m)} \geq (1-\pi(w))p_2^m(1-p_2)^{(n(w)-m)} \\
&\Leftrightarrow \quad (\tfrac{p_1}{p_2})^m(\tfrac{1-p_1}{1-p_2})^{n(w)-m} \geq \tfrac{1-\pi(w)}{\pi(w)} \\
&\Leftrightarrow \quad (\tfrac{p_1(1-p_2)}{p_2(1-p_1)})^m(\tfrac{1-p_1}{1-p_2})^{n(w)} \geq \tfrac{1-\pi(w)}{\pi(w)} \\
&\Leftrightarrow \quad m\ln(\tfrac{p_1(1-p_2)}{p_2(1-p_1)}) + n(w)\ln(\tfrac{1-p_1}{1-p_2}) \geq \ln\tfrac{1-\pi(w)}{\pi(w)} \\
&\Leftrightarrow \quad m\ln(\tfrac{p_1(1-p_2)}{p_2(1-p_1)}) \geq \ln\tfrac{1-\pi(w)}{\pi(w)} - n(w)\ln(\tfrac{1-p_1}{1-p_2}) \\
&\Leftrightarrow \quad m \geq \frac{\ln\frac{1-\pi(w)}{\pi(w)} + n(w)\ln(\frac{1-p_2}{1-p_1})}{\ln(\frac{p_1(1-p_2)}{p_2(1-p_1)})}.
\end{aligned}$$

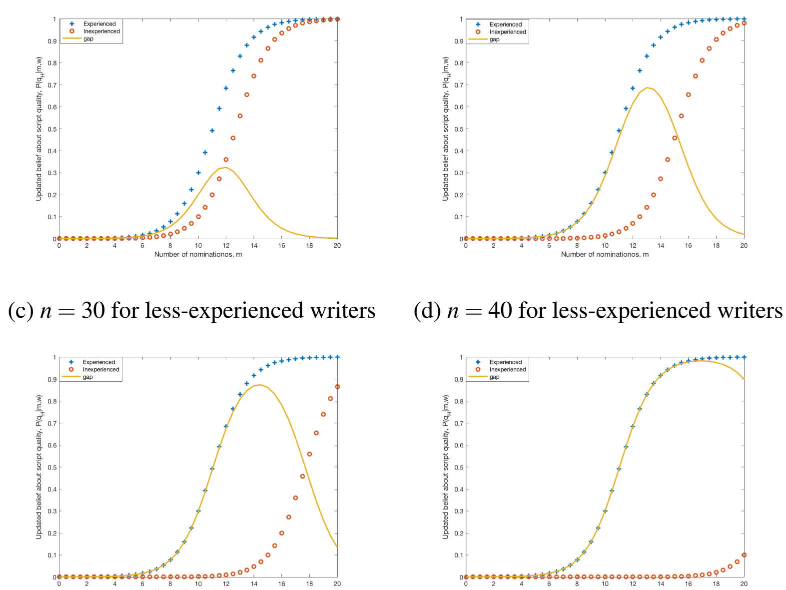
We can preserve the sign of inequality in the last line because $\frac{p_1(1-p_2)}{p_2(1-p_1)} > 1$ when $p_1 > p_2$.

Letting $m^* = \frac{\ln\frac{1-\pi(w)}{\pi(w)} - n(w)\ln(\frac{1-p_1}{1-p_2})}{\ln(\frac{p_1(1-p_2)}{p_2(1-p_1)})}$, we have the result for the first part of Proposition 4: when $n'(w) = 0$, $\frac{\partial^2 \mathrm{P}(q_H|m,w)}{\partial w \partial m} > 0$ when $m < m^*$; and $\frac{\partial^2 \mathrm{P}(q_H|m,w)}{\partial w \partial m} < 0$ when $m > m^*$. Furthermore, note that m^* is an increasing function of n and a decreasing function of π. As an example, when $p_2 = 1 - p_1$, $m^* = \frac{n}{2}$ when $\pi = 0.5$; $m^* < \frac{n}{2}$ when $\pi > 0.5$; and $m^* > \frac{n}{2}$ when $\pi < 0.5$.

Now consider the second part of Proposition 4—when scripts from less-experienced writers are read by more Black List voters (i.e., $n'(w) < 0$). Based on equation (A3), we can similarly show that $\frac{\partial^2 \mathrm{P}(q_H|m,w)}{\partial n(w) \partial m} \leq 0 \Leftrightarrow m \geq m^* = \frac{\ln\frac{1-\pi(w)}{\pi(w)} - n(w)\ln(\frac{1-p_1}{1-p_2})}{\ln(\frac{p_1(1-p_2)}{p_2(1-p_1)})}$. Thus, qualitatively, we also have $\frac{\partial^2 \mathrm{P}(q_H|m,w)}{\partial w \partial m} > 0$ when $m < m^*$; and $\frac{\partial^2 \mathrm{P}(q_H|m,w)}{\partial w \partial m} < 0$ when $m > m^*$. But as stated above, m^* increases with n. Thus, a greater number of voters for less-experienced writers pushes the inflection point m^* to be greater.

The second part of Proposition 4 is much easier to understand with the help of a numerical example in which we use *discrete* values of writer experience, as well as their respective extents of script visibility, rather than considering only marginal changes around $n(w)$. Figure A1 plots the updated beliefs about quality of scripts receiving a given number of nominations for experienced writers (indicated by the blue, plus signs) and less-experienced writers (indicated by the red, circle signs) separately. The figure also plots the gap in the updated beliefs between experienced and less-experienced writers (indicated by the yellow line). Figure A1a below presents a situation in which all writers have the same number of voters (20) who have read their scripts. Under the parameter values of this example, the gap by writer experience starts to decrease and converge to zero after 12. In Figure A1b, all parameters are of the same value, except that scripts from less-experienced writers are read by more Black List voters (25 instead of 20). In Figures A1c and A1d, we increase the number of voters who have read scripts from less-experienced writers to 30 and 40. As the graph shows, the inflection point m^*, after which the gap by writer experience starts to decline, gets increasingly larger as we increase the number of voters who have read scripts from less-experienced writers. In Figure A1d, for example, the gap continues to rise until when the number of nominations is about 18, after which it starts to decline.

Figure A1: Illustration of Proposition 4



(a) $n = 20$ for all writers

(b) $n = 25$ for less-experienced writers

(c) $n = 30$ for less-experienced writers

(d) $n = 40$ for less-experienced writers

Note: The horizontal axis is the number of nominations, and the vertical axis is the updated belief about script quality $\mathrm{P}(q_H|m,w)$. We plot two curves for experienced and less-experienced writers, as well as the gap between these two curves. The four figures use exactly the same parameter values, except for the number of voters who have read scripts from less-experienced writers. In particular, in (a), scripts from writers of different experience levels are read by the same number of voters ($n_{\text{experienced}} = n_{\text{less-experienced}} = 20$). In (b), scripts from less-experienced writers are read by more people ($n_{\text{experienced}} = 20$, while $n_{\text{less-experienced}} = 25$). In (c) and (d), we increased $n_{\text{less-experienced}}$ to 30 and 40, while keeping $n_{\text{less-experienced}}$ the same at 20. The other parameter values are: $\pi_{\text{experienced}} = 0.3$, $\pi_{\text{less-experienced}} = 0.1$, $p_1 = 0.6$, and $p_2 = 0.4$.

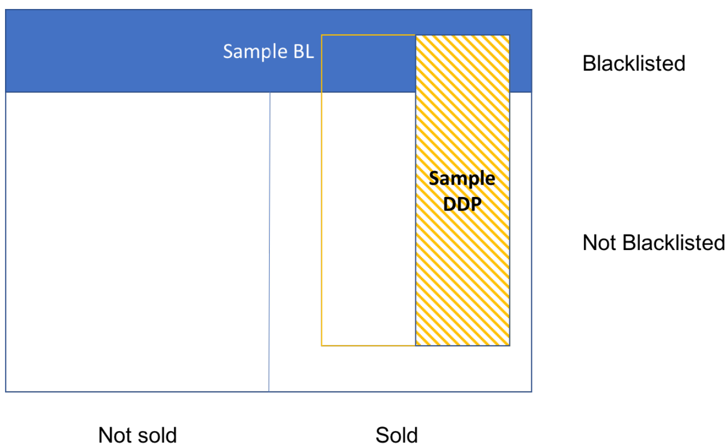
$\square$

Online Appendix B. Additional Results

B.1 Samples

Figure B1 provides a graphical illustration of the relationship between the two samples we use in the analysis. Consider the population of scripts that are up for sale (from the writers) or commissioned (by the producers). The solid, blue rectangle indicates Sample BL—it is a complete sample of scripts that are Black-Listed. These scripts may or may not be sold eventually (either before or after the List's publication). The yellow, shaded rectangle indicates Sample DDP, which is a subset of the DDP database—indicated by the larger yellow-framed rectangle and itself a comprehensive but incomplete sample of all sold scripts

Figure B1: Graphical Illustration of the Relationship between Sample BL and Sample DDP



B.2 Black List Nominations and Release Outcomes

Table B1 reports the marginal increase in R^2 measures by including Black List variables for Table 2, the set of regressions that examines the predictiveness of Black List nominations for the probability of release. This table reproduces the results of Table 2; only Black List variables are presented, and the coefficients of other controls are omitted for ease of presentation. For each column, we reran the regression without the specific Black List variables used in this column. The bottom of the table displays both the level differences and the percent changes in two different R^2 measures (unadjusted as well as adjusted by the number of predictors). The percent change is defined as (R^2 measure with Black List variables - R^2 measure without Black List variables)/R^2 measure without BL variables.

These comparisons reveal two results. First, the marginal increase in the regressions' explanatory power is economically significant. Out of the twelve different comparisons, the percent change ranges from 3.2% to 16.0%, with the median being 8.9%. Second, these percent increases become larger when using adjusted R^2 than when using unadjusted R^2. Because the adjusted R^2 penalizes the addition of new variables, the fact that the relative increase is larger for the adjusted R^2 suggests that the Black List variables provide disproportionately more explanatory power relative to the other predictors in the regressions.

Table B2 reports regression results analogous to those in Table 2 based on the coarsened exact matching

Table B1: Comparisons of R-squared with and without Black List Variables for Table 2 "Black List Nominations and Release Probability"

	Sample DDP			Sample BL		
			Exclude Top-20			
	Release	Release	Release	Release	Release	Release
	(1)	(2)	(3)	(4)	(5)	(6)
Black-Listed	0.135**	0.120**	0.109**			
	(0.047)	(0.039)	(0.042)			
Top-5				0.295***	0.280***	
				(0.055)	(0.048)	
Top 6-20				0.090	0.093	
				(0.072)	(0.058)	
Top 21-Above Threshold				0.065	0.060	
				(0.043)	(0.044)	
Nomination Ratio						1.536***
						(0.006)
		Other Controls Omitted				
R-squared						
R-squared with BL variables	0.113	0.163	0.162	0.213	0.268	0.272
R-squared without BL variables	0.102	0.155	0.157	0.194	0.252	0.252
Difference	0.012	0.008	0.005	0.019	0.016	0.021
Percent change	11.3%	5.2%	3.2%	9.7%	6.5%	8.1%
Adjusted R-squared						
Adjusted R-squared with BL variables	0.082	0.084	0.081	0.168	0.189	0.196
Adjusted R-squared without BL variables	0.071	0.076	0.076	0.152	0.175	0.175
Difference	0.011	0.008	0.005	0.016	0.014	0.021
Percent change	16.0%	10.8%	6.3%	10.6%	7.9%	12.0%

(CEM) method. The regressions generate a matched sample of listed and unlisted scripts in Sample DDP that are similar to each other in pre-specified dimensions (Iacus et al., 2012). Variables used to create the matched sample are Writer Major Credits, Top-10 Agency, Experienced Agent, Producer Major Credits, Whether Studio Buyer, Original, Talent Attached, and Whether Reported. Column 2 excludes top-20 listed scripts. The CEM estimates, at about 13 to 15 percentage points, are statistically significant and economically consistent with our baseline results reported in Table 2. The (unreported) propensity score matching method produces an estimate of the average treatment effect at 9.8 percentage points (significant at the five-percent level).

Table B2: Average Effects of Being Black-Listed on Release Probability (CEM)

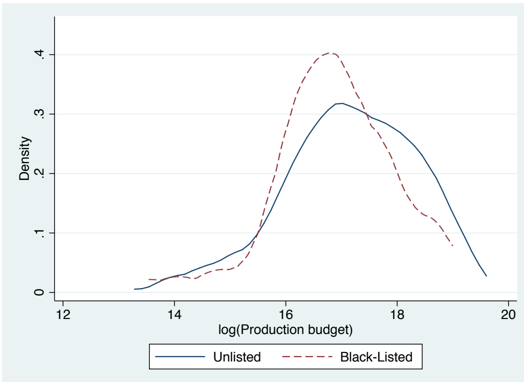
	Release	Exclude Top-20 Release
	(1)	(2)
Black-Listed	0.150**	0.152**
	(0.051)	(0.056)
Writer Major Credits	0.089*	0.087
	(0.047)	(0.062)
Top-10 Agency	0.015	0.036
	(0.046)	(0.060)
Experienced Agent	-0.045	-0.014
	(0.058)	(0.047)
Producer Major Credits	0.001	0.000
	(0.006)	(0.007)
Original	0.101**	0.083
	(0.034)	(0.052)
Talent Attached	0.054	0.061
	(0.045)	(0.048)
Bidding	-0.005	-0.057
	(0.050)	(0.074)
Whether Reported	0.021	0.016
	(0.053)	(0.063)
Genre FE	Y	Y
Content Source FE	Y	Y
Studio FE	Y	Y
Year FE	Y	Y
N	425	287
R-squared	0.168	0.171

Note: Standard errors (in parentheses) are clustered in two dimensions: year and studio. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure B2 plots the density distributions of log(Production Budget). Panel (a) uses released movies in Sample DDP and plot log(Production Budget) by whether or not the movie is Black-Listed; and Panel (b) plots the same density for Sample BL by Black List Rank Group. As discussed in the paper, the raw data show that the budget level is not statistically different between listed and unlisted scripts or by rank group.

Figure B2: Production Budget by Black List Nominations

(a) Sample DDP



(b) Sample BL

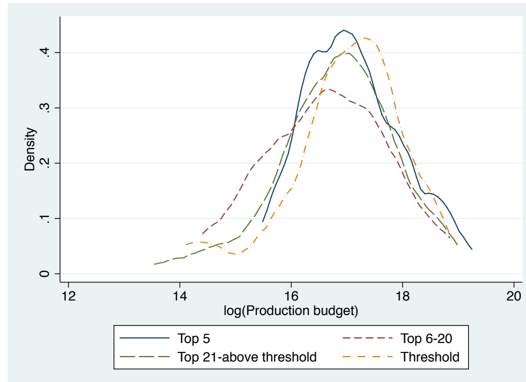


Table B3 reports regression results in which the dependent variables are measures of investments other than production budget: Open Screen is the number of theaters screening the movie during the opening weekend; Star Score is an index generated by the-numbers.com, based on the total box office revenues of the movies that an actor or actress has been in prior to the release year of the focal movie, averaged over all the leading cast of the movie; Award-Winning Cast is an indicator of whether any of the leading cast of the movie has won any award (Academy Award, Golden Globe, and BAFTA) prior to the release year of the focal movie; and following Basuroy et al. (2003), Seasonality is based on the total number of tickets sold for all movies shown during a given weekend, averaged across years. We use the box office revenue data of all the movies from *the-numbers.com* between 2005 and 2017. Dividing the average number of tickets for the 52 weekends by the maximum sold (Christmas weekend), we generate a normalized index ranging between zero and one.

As discussed in the paper, other measures of investments are also statistically similar between listed and unlisted scripts; and, for listed scripts, these measures are largely statistically similar across all rank groups.

Table B3: Other Measures of Investments

(a) Sample DDP

	Open Screen (1)	Star Score (2)	Award-Winning Cast (3)	Seasonality (4)
Black-Listed	47.661	5.775	0.043	0.044
	(229.226)	(29.280)	(0.079)	(0.035)
Other controls	Y	Y	Y	Y
Release Year FE	Y	Y	Y	Y
Studio FE	Y	Y	Y	Y
N	239	239	239	239
R-squared	0.619	0.312	0.153	0.234

(b) Sample BL

	Open Screen (1)	Star Score (2)	Award-Winning Cast (3)	Seasonality (4)
Top 5	-537.116	-18.019	0.110	0.011
	(407.522)	(34.322)	(0.200)	(0.048)
Top 6-20	-494.617	-64.982**	0.088	-0.021
	(300.380)	(24.775)	(0.122)	(0.029)
Top 21-Above Threshold	-596.002*	-17.841	0.162	0.005
	(264.951)	(27.679)	(0.152)	(0.037)
Other controls	Y	Y	Y	Y
Release Year FE	Y	Y	Y	Y
Studio FE	Y	Y	Y	Y
N	177	177	177	177
R-squared	0.493	0.240	0.245	0.326

Note: Panel (a) uses released movies in Sample DDP, and Panel (b) uses released movies in Sample BL. All regressions in Panel (a) include exactly the same controls as in Column 2 in Table 3a, and all regressions in Panel (b) include exactly the same controls as in Column 2 in Table 3b. For all regressions, standard errors are clustered in two dimensions: year and studio. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.3 The Black List and Writers of Different Experience Levels

Table B4 reports regression results analogous to those in Table 4 using different measures of writer's experience levels. These regressions also show a significant negative correlation between Black List nominations and writer experience. In both panels: Column 1 counts writing credits for movies distributed by top-30 distributors in the previous ten years ('writer major credits' defined in Section 3.2, thus replicating Table 4's results); Column 2 counts writing credits for movies distributed by top-15 distributors in the previous ten years; Column 3 counts all writing credits prior to survey (sale) year without time period or distributor size restrictions; Column 4 uses an indicator that the writer has one or more major writing credits (same as used in the paper) and includes indicators for writer other experience—e.g., any directing, producing, or acting credits for movies distributed by top-30 distributors in the previous ten years.

Table B4: Black List Nominations and Writer Experience: Alternative Experience Measures

(a) Sample DDP

	Black-Listed	Black-Listed	Black-Listed	Black-Listed
	(1)	(2)	(3)	(4)
Writer's Writing Experience	-0.035**	-0.031***	-0.005**	-0.076**
	(0.011)	(0.009)	(0.002)	(0.028)
Writer's Directing Experience				-0.030
				(0.019)
Writer's Producing Experience				-0.003
				(0.021)
Writer's Acting Experience				-0.027
				(0.020)
N	1565	1565	1565	1565
R-squared	0.185	0.183	0.181	0.188

(b) Sample BL

	Top 20	Top 20	Top 20	Top 20
	(1)	(2)	(3)	(4)
Writer's Writing Experience	-0.041**	-0.040**	-0.011**	-0.093*
	(0.017)	(0.015)	(0.004)	(0.041)
Writer's Directing Experience				-0.021
				(0.040)
Writer's Producing Experience				-0.090**
				(0.036)
Writer's Acting Experience				-0.033
				(0.065)
N	700	700	700	700
R-squared	0.068	0.067	0.066	0.075

Note: Panel (a) uses Sample DDP, and the dependent variable of all columns is an indicator of whether a script is Black-Listed. All regressions include the same control variables as in Column 2 in Table 4. Panel (b) uses Sample BL, with the dependent variable of all columns being an indicator for top-20 scripts. All regressions include the same control variables as in Column 5 in Table 4. Standard errors (in parentheses) are clustered in two dimensions: year and studio. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B5 provides a test for Assumption 1 (that is, the prior probability of being high-quality is greater for ideas from experienced writers. We also use this analysis to investigate the first alternative explanation we discuss in the paper for our results by writer experience (see Section 4.2.3). In particular, were Assumption 1 not true—that is, if scripts from less-experienced writers were actually more likely to be very good—scripts from less-experienced writers would be more likely to be Black-Listed and to rank higher if listed without having wider visibility. This explanation, together with risk aversion, would explain why, if listed, scripts from less- experienced writers are less likely to be released.

To investigate this further, we use all sales records in the DDP database before 2005, the year in which the Black List was started. In Panel (a) of Table A4, the dependent variable of all columns is whether the script is theatrically released, and the default group includes writers with zero prior major experience. The results show a monotonic increasing relationship between release likelihood and writer experience, which is consistent with Assumption 1. Panel (b) uses movies that are released and for which budget information is available. Column 1 reports OLS regression results, which show that released movies from experienced writers generate either statistically similar or significantly higher box office revenues than writers with no prior major experience (the default group). Furthermore, quantile regressions (in Columns 2-5) also do not indicate a heavier right tail in the distribution of box office revenues for scripts from less-experienced writers. These results, collectively, provide supporting evidence for Assumption 1.

In Table B6b, we investigate the second alternative explanation discussed in Section 4.2.3. Suppose that survey respondents nominate scripts that appeal to them for non-monetary reasons; for example, they may find scripts from less-experienced writers more novel or interesting, which will result in the Black List over-representing less- experienced writers. But these scripts may be less financially viable and, hence, less likely to be produced. This explanation seems to resonate with the following quote from an industry veteran: "[I]nnovative scripts by young writers tend to score high in the Black List, but these scripts are usually not properly 'put together' so when it is time to actually make a multi-million-dollar movie, there is preference for less innovative, more formulaic and better structured screenplays."[1]

To investigate this explanation, we construct two types of novelty measures: one based on the genre combinations of the script and the other based on the script's logline. The genre-based novelty measures follow the patent literature (Verhoeven et al., 2016). In particular, we create two dummy variables that indicate the least-used genre combinations accounting for ten or 25 percent of the scripts sold in a given year, with the frequency of a genre combination based on all scripts sold prior to that year. The logline-based measures follow the machine learning literature (Mikolov et al., 2013); see Appendix C.2.5 for a detailed explanation.

Panel (a) in Table B6 shows that, consistent with what we would expect, Black-Listed scripts have a higher value of these measures than unlisted scripts and, for released movies, the correlations between these measures and the mean score of critic reviews are also all positive.[2] Panel (b) in Appendix Table B6 shows that these measures are not significantly correlated with writer experience; the only correlation with a p-value is less than 0.10 between more-novel scripts and *experienced writers*. It is certainly possible that our measures are not precise enough, as novelty is a complex concept, but available data and measures used by prior work do

[1]We thank an anonymous referee who provided us with this quote.

[2]Three out of the eight correlations are significant at the five-percent or ten-percent level; and for another three, the p-value is less than 0.2. The critics' ratings data are from *metacritic.com*, an aggregator that coverts movie reviews by major newspapers and magazines to a standardized score. The metacritic.com data are used in the prior literature, including Brown et al., (2013).

Table B5: Writer Experience and Release Outcomes (DDP database before 2005)

(a) Release Probability

	Release (1)	Release (2)
Writer Major Credits = 1	0.040	0.025
	(0.024)	(0.051)
Writer Major Credits = 2	0.139**	0.086*
	(0.063)	(0.044)
Writer Major Credits = 3	0.206**	0.172**
	(0.093)	(0.055)
Writer Major Credits $\geq$ 4	0.147*	0.149
	(0.080)	(0.085)
Other controls	N	Y
Studio$\times$year FE	N	Y
Year FE	Y	N
N	1112	1095
R-squared	0.031	0.136

(b) Box Office Performance for Released Movies

	OLS Regressions	Quantile Regressions			
		25th	50th	75th	90th
	log(Box Office) (1)	Box Office (2)	Box Office (3)	Box Office (4)	Box Office (5)
log(Budget)	0.824***	37.277**	43.712	67.558**	57.713***
	(0.155)	(18.727)	(29.871)	(26.397)	(10.114)
Writer Major Credits = 1	0.146	27.192	62.232**	38.018	-3.707
	(0.230)	(131.344)	(25.348)	(87.878)	(15.216)
Writer Major Credits = 2	-0.518	-17.476	84.220	81.846	186.484***
	(0.689)	(29.217)	(59.655)	(430.800)	(46.667)
Writer Major Credits = 3	0.672**	74.349	80.002***	48.688	229.417***
	(0.279)	(127.634)	(30.772)	(77.180)	(31.332)
Writer Major Credits $\geq$ 4	0.288	-14.733	57.526	32.481	52.953
	(0.210)	(49.914)	(84.900)	(153.847)	(36.933)
Other controls	N	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
Studio FE	Y	Y	Y	Y	Y
N	111	115	115	115	115
R-squared	0.695	0.516	0.558	0.560	0.469

Note: Other controls include Top-10 Agency, Experienced Agent, Producer's Major Credits, Original Idea, Talent Attached, Bidding, and Genre and Content Source fixed effects. Standard errors (in parentheses) are clustered in two dimensions: year and studio. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

not support the typical impression that scripts from less-experienced writers are systematically more novel.

Table B6: Novelty Measures

(a) Correlations between Novelty Measures and Black-Listed and Mean Critic Reviews

	Logline-Based Novelty Measures		Genre-Based Novelty Measures	
	Not Weighted	Weighted	10 Percent	25 Percent
Black Listed	0.041	0.050	0.034	0.021
(p-value)	(0.141)	(0.074)	(0.176)	(0.402)
N	1289	1289	1564	1564
Released Scripts Only				
Mean (critic reviews)	0.111	0.045	0.094	0.141
(p-value)	(0.081)	(0.481)	(0.114)	(0.017)
N	248	248	287	287

(b) Correlations between Novelty Measures and Writer Experience

	Logline-Based Novelty Measures		Genre-Based Novelty Measures	
	Not weighted	Weighted	10 Percent	25 Percent
Experienced Writer	0.050	-0.010	0.020	0.031
	(0.071)	(0.713)	(0.442)	(0.228)
	1289	1289	1564	1564
Black-Listed Scripts Only				
Experienced Writer	-0.009	-0.002	0.012	0.037
	(0.907)	(0.977)	(0.878)	(0.626)
	159	159	178	178

Note: Sample DDP. Panel (a) presents correlations between different novelty measures and whether a script is Black-Listed and the mean critical reviews for released movies. These measures are not available for all scripts in Sample DDP because we exclude observations without a logline or with a logline of fewer than six words. Panel (b) correlates these measures with writer experience, first for all observations in Sample DDP and then for only Black-Listed scripts in this sample.

B.4 Black-Listed Scripts Unsold Prior to Publication

Of the 145 Black-Listed scripts unsold at the time of publication, only eight (six percent) were eventually theatrically released. To further understand why so few eventually made it to the theater, we unpacked the low release rate into two components: (i) the probability of making a sale after being listed; and (ii) the probability of theatrical release conditional on being sold.

To obtain a sale-rate estimate, we matched these unsold scripts to the entire DDP database. We matched 26 percent: a lower-bound estimate of the sale rate because DDP does not capture all sold scripts. By further assuming that DDP's completeness in capturing all sold scripts is not systematically different for Black-Listed scripts that are sold and unsold before publication, we obtained a point estimate of a 40-percent sale rate.[3]

[3] 356 out of 556 (64 percent) of Black-Listed scripts that are sold before publication are captured by the DDP database.

While not definitive, these estimates suggest that the Black List could have a sizable impact on garnering attention for unsold scripts. Using the point estimate above, we obtained a conditional release rate of 13.5 percent (8/59). This release rate (with notable small- sample caveats) is far below the (conditional) release rate for Black-Listed scripts that were already sold (32 percent) and is lower even than that for unlisted sold scripts from Sample DDP (16 percent).

The above statistics indicate that unsold Black-Listed scripts have a low release rate not because they do not get sold. Instead, these scripts (once sold) have a release rate lower than scripts never listed. There are at least three possible reasons that this outcome could occur: (i) there may be a strong selection effect that results in these scripts having less-viable material than listed but sold scripts;[4] (ii) relatedly, due to the fact that these scripts are not yet sold, the market may infer lower quality and, hence, hesitate to commit resources; and (iii) the Black List may carry momentum only for a short time, which suggests that, due to the time it takes for these scripts to sell, previous years' Black List scripts may already be forgotten, thus making it difficult to overcome the hurdles in the remainder of the production process. Unfortunately, the small sample size and data incompleteness restrict our ability to investigate these conjectures further.

Table B7: Sold versus Unsold Black-List Scripts

	N	Black List rank group	Experienced writers	Top 10 agency	Agent experience	Released
Unsold	145	1.93	0.07	0.84	10.53	0.06
Sold	556	2.13	0.22	0.93	13.38	0.32
(p-value)		(0.01)	(0.00)	(0.00)	(0.02)	(0.00)

Note: Raw data based on Sample BL. Sold is an indicator of whether a script is associated with a production company or studio at the time of the List's publication.

B.5 Voter Incentives

Voters may have strategic incentives when they vote, and social stigma/reputation concerns may not be a sufficient deterrent due to the survey's anonymity. For example, voters may want to keep "hidden gems" to themselves for scripts that they want to buy, or they may want to nominate their own scripts given the potentially large rewards from being listed. Without detailed data on who votes for which scripts, however, it is hard to tease apart these various incentives. That said, we examined the two incentives discussed above using the available data.

First, among Black-Listed scripts, regression results controlling for nomination-year fixed effects (Column 1 in Table B8) show that sold scripts receive 2.60 more nominations than unsold scripts (p-value of 0.011). The result is similar when including other controls (Column 2 in Table B8). This result is consistent with the idea that voters may have incentives to keep unsold 'hidden gems' to themselves. However, the gap may also

[4] The data suggest some selection effects at least on observables. In particular, Table B7 shows that unsold scripts are ranked lower than sold scripts, are less likely to come from experienced writers, and are more likely to have representation by small agencies and less-experienced agents.

be due to other reasons that a given script is not sold yet—for example, it may have less-viable materials to start with (which is consistent with evidence presented in the next section on listed scripts that are not sold prior to publication).

Second, we wanted to see whether scripts purchased by producers who are more likely to be eligible voters are systematically more likely to be listed and ranked higher than those purchased by producers who are less likely to be eligible voters. We defined producers as eligible voters if they had a major release (that is, a movie distributed by the top-30 studios) in the previous three years. This definition is likely to be a reasonable, though noisy, measure of the Black List's criteria because we do not know exactly how it defines 'major release.' Because of potential confounding factors—e.g., larger and more-prominent producers are both more likely to be eligible voters and more likely to be matched to better scripts—we control for other factors to the extent that we can, including the producer's experience in the previous ten years. Columns 3 and 4 in Table B8 show that, among Black-Listed scripts that are sold prior to publication, scripts associated with producers who are more likely to be eligible do not receive significantly more nominations than scripts associated with producers that are unlikely to be eligible to vote.

Table B8: Potential voter incentives

Dependent Variable	Nominations (1)	Nominations (2)	Nominations (3)	Nominations (4)
Sold Prior to Publication	2.605**	1.984***		
	(0.837)	(0.611)		
Eligible Voter			0.562	-0.983
			(2.319)	(1.345)
Other controls	N	Y	N	Y
Year FE	Y	Y	Y	Y
Studio FE	N	Y	N	Y
N	701	701	556	556
R-squared	0.097	0.146	0.075	0.131

Note: OLS (linear probability) regressions using Sample BL. The dependent variable is the number of nominations received by a listed script. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

References

Basuroy, S., S. Chatterjee, and S. A. Ravid (2003). "How Critical are Critical Reviews? The Box Office Effects of Film Critics, Star Power, and Budgets." *Journal of Marketing* 67, 103-117.

Brown, A. L., C. F. Camerer, and D. Lovallo (2013). "Estimating structural models of equilibrium and cognitive hierarchy thinking in the field: The case of withheld movie critic reviews." *Management Science* 59, 733-747.

Iacus, S. M., G. King, and G. Porro (2012). "Causal inference without balance checking: Coarsened exact matching." *Political Analysis* 20, 1-24.

Mikolov, T., K. Chen, G. Corrado, and J. Dean (2013). "Efficient Estimation of Word Representations in Vector Space, " arXiv e-prints, arXiv:1301.3781.

Verhoeven, D., J. Bakker, and R. Veugelers (2016). "Measuring technological novelty with patent-based indicators. " *Research Policy* 45, 707-723.

Online Appendix C: Data Appendix

As Section 3.1 discusses, this study uses two different samples in its analysis. This Appendix further explains how each sample was created.

C.1 Sample Creation

C.1.1 Initial Data from Black List Website

The Black List is published annually on a website (https://blcklst.com/lists/) and in PDF format. The information provided includes script name, writer(s), publication year, writers' agent, agent's agency, and the number of nominations received. Figure C.1 provides a screenshot image from 2005. Since 2007, the Black List has also included information on production companies or movie studios associated with scripts. Over 2005-2016, 1,224 scripts were listed. We use the 701 scripts listed over 2007-2014 to take advantage of buyer-side information and to provide sufficient time for script release prior to our final- outcome data collection (October 2017).

Figure C1: Image from 2005 Black List PDF

THE BLACK LIST			***2005***
twenty-five mentions			
THINGS WE LOST IN THE FIRE	Allan Loeb	CAA	Carin Sage
twenty-four mentions			
JUNO	Diablo Cody	Gersh	Sarah Self
fifteen mentions			
LARS AND THE REAL GIRL	Nancy Oliver	UTA	Tobin Babst
fourteen mentions			
ONLY LIVING BOY IN NEW YORK	Allan Loeb	CAA	Carin Sage
thirteen mentions			
CHARLIE WILSON'S WAR	Aaron Sorkin	Endeavor	Jason Spitz
ten mentions			
KITERUNNER, THE	David Benioff	CAA	Todd Feldman

Note: Figure C.1 shows the first six scripts appearing on the Black List in 2005.

C.1.2 Release Outcome Data

To determine script release outcomes, we undertake two steps: first, using Internet Movie Database (IMDb) information, we confirm whether a script has been produced; and second, using TheNumbers.com (which records movie US box office revenues) information, we determine whether a script has been 'theatrically released.'

Confirming whether or not a script was ultimately produced can be tricky, as script names can change during the production process. To address this issue, we used information from writers' IMDb pages to guide our search. We began with a manual search of each writer's IMDb page. If the writer had no film releases in or after the year the script appeared on the Black List, we coded the script as not produced. If the writer had film releases before October 2017, we coded the film as produced if a title match with the Black List script occurred. This process was aided by IMDb's "also known as" field, which often provides the original script name (e.g., at the time it was Black List eligible). If no IMDb title matched the Black List script name, but the writer had subsequent releases, we reviewed IMDb plot summaries and compared them to Black List loglines, sometimes using third parties (such as Script Shadow) if and when more-detailed summaries were required. Where sufficient plot overlap occurred, and the Black List date preceded the matching IMDb film release date, we coded the script as produced.

This approach identified 222 out of 701 Black List scripts (31.6 percent) over 2007-2014 as produced. For these produced scripts, we manually searched TheNumbers.com to collect box office information. This step was relatively straightforward, as IMDb titles closely match TheNumbers titles. We ultimately identified 184 (83 percent) of the 222 produced scripts as theatrically released; the remainder were released at film festivals or through other channels such as direct to DVD.

C.1.3 Writer and Producer Experience

We measured writer experience by parsing writers' IMDb pages to collect all films for which the writer obtained writing credits. Simply using all past writing credits as a writer's experience measure is misleading, however, as the majority of IMDb films recorded are small and independently produced. To address this concern, we further collected U.S. film distributor information and used it to define major- release films as those distributed by one of the top 30 movie studios (based on market share). The number of major-release writing credits prior to a given sale represents the writer's experience measure. When more than one writer was listed, we took the maximum of the writers' experience as the team's experience.

We measured producer experience similarly to writer experience: parsing information from producers' IMDb pages to collect all films for which the producer obtained producing credits. We again considered only past major-release films for which producers obtained producing credits.

C.1.4 Agency Size and Agent Experience

We defined agency size by leveraging the entire Deal Done Pro (DDP) database. Section C.2.5 describes this database in more detail. We calculated agency market share based on the script sales recorded by the agency relative to total industry script sales. An indicator of top-ten agency is then defined as those agencies with the ten largest market shares.

We defined agent experience by leveraging the DDP database and using an approach similar to agency size. We counted the number of sales associated with agents at particular points in time. An indicator of experienced agent is then defined as those agents with more than fifteen prior sales.

C.2 Sample DDP

As discussed in Section 3.1, Sample DDP included only transactions for which completed scripts were in place at the time of sale. This section describes how the entire Done Deal Pro (DDP) sample was generated, including all transaction types.

C.2.1 Initial Data from Done Deal Pro

Done Deal Pro ((DDP) provides an online database (http://www.donedealpro.com/default.aspx) of screenplay sales since 1997.[5] Each sales record includes the script title, a brief summary (called a "logline"), the writer(s), the firms and individuals involved, and a comments field (called "More") containing additional sales information. Figure C.2 provides a sales entry example from the DDP website.

Figure C2: Sample Sale Entry via Deal Done Pro Website

Title:	**Things We Lost in the Fire**
Logline:	A woman is widowed when her husband dies suddenly, leaving her alone with two children. She decides to invite her husband's troubled best friend to live with them and, as the friend turns his life around, he helps the fractured family confront the emotional void left by the loss.
Writer:	Allan Loeb
Agent:	Jon Levin
Agency:	Creative Artists Agency
Studio:	DreamWorks SKG
Prod. Co:	Scamp
Genre:	Drama
Logged:	7/8/2005
More:	Spec. Sam Mendes' Scamp Films will produce.

Note: Figure C.2 shows a DDP sales entry screen shot.

These data were used to create several key variables, including the record date (defined as the sale date), writer(s), title, agent, agency, production company, movie studio, and genre. By parsing text in the "More" field, we developed several additional variables:

- **Original:** We identified whether a script is original (versus an adaptation) by creating a list of terms indicating whether the idea came from a previous source, such as 'book,' 'comic book,' 'tv series,' or 'short story.' The variable "original" equals zero if such a term is identified in the "More" field and one otherwise.

- **Turnaround/Rewrite:** 6.55 percent of the 6,294 records over 2007-2014 are identified as a turnaround or a rewrite, defined by the "More" field containing the keywords 'rewrite' and/or 'turnaround' (an industry term for projects that were originally developed by a different studio). We excluded these cases from the main analysis, but our results are robust to including them.

- **Complete:** For the remaining 5,882 scripts, we identified whether a completed script exists at DDP record capture. In particular, "complete" equaled zero if the "More" field contained terms such as 'pitch' (an industry term for an idea sale without a completed script), 'treatment' (a two- to three-page outline), 'to be based on,' 'to adapt,' 'will adapt,' 'to be adapted,' 'assignment,' etc. This category accounted for 40 percent of the 5,882 scripts. We coded "complete" as one if the "More" field contained keywords

[5]DDP also collects adaptation-right transactions identified by information in the 'Writer' field. These cases always include parentheses after the writer's name, such as (author) or (creator). We do not consider these transactions in our analysis, however, as they do not involve a screenwriter later hired to write the screenplay.

such as ‘spec,’ ‘script,’ ‘screenplay,’ ‘based on,’ and ‘adapted from’: 27 percent of the 5,882 scripts were in this category. We coded “complete” as missing if the “More” field did not contain sufficient information to determine whether a completed script existed, and 29 percent of the 5,882 scripts were in this category (these records typically contain information only about producers managing projects). We took a conservative approach and excluded these unclear cases from the main analysis, but our results were robust to including these cases in Sample DDP.

C.2.2 Whether Black-Listed

For each Black-Listed script, we manually searched for a match in the DDP database. We found matches for 59 percent of the scripts in Sample BL. There are two reasons why Black-Listed scripts do not appear on DDP: (1) these scripts may never be sold; and (2) even if these scripts are sold, DDP may not record the transaction.

C.2.3 Release Outcome Data

We used the same two-step procedure explained in Section C.1.2 to identify whether or not a project was produced and theatrically released. Given the DDP database size, we implemented an algorithmic approach instead of a manual approach.

We first cleaned DDP writer names as listed to remove any special characters and to split out multiple authors. We then created a search algorithm for writers’ IMDb pages via Google (e.g., ‘Paul Thomas Anderson’ and ‘IMDb’) to select the first IMDb link. We tested approximately 50 different writers of varying experience to confirm that this search did not produce spurious results and to verify that the site returns correct writers. We then saved writers’ html pages along with their unique IMDb identifiers.

We next created a list of all unique writer-DDP sale pairs (i.e., one pair for each writer) using all DDP projects. Using the IMDb information parsed from the step above, we created a corresponding writer-IMDb film pair list for writers’ released films prior to October 2017. We then matched the writer-IMDb film pair list and the writer-DDP sale pair list by writer name and script title using the Stata function matchit. For exact matches, we recorded the DDP project as produced into this particular IMDb film. For inexact matches, we manually reviewed the closest matches (i.e., matchit function similarity scores above 0.5) and determined whether a match could be found. Separately, there are about 1,000 DDP projects containing ‘untitled’ in name (e.g., ‘Untitled Facebook Project’). Because matching based on movie titles here is unlikely to generate high similarity scores, we manually matched these titles based on plot information, using the same procedure outlined in Section C.1.2 above. Finally, we matched produced films to The Numbers database to obtain box office information and determine whether a produced film had been theatrically released.

It is possible that the algorithm approach undercounted the number of listed scripts in the DDP database that were released, relative to those in Sample BL. In order to ensure that there were no systematic differences in the way that we determined listed versus unlisted script release outcomes within Sample DDP, we used the same algorithmic method for both (i.e., we did not impute the manually collected film release information from Sample BL for Black-Listed scripts).

C.2.4 Other Variables

We defined writer experience, producer experience, agency size, and agent experience in the same way as for Sample BL.

C.2.5 Logline based novelty measure

This section describes the process used to create our logline-based novelty measure of the scripts. To build our measure, we used logline data (short summaries of scripts) and textual analysis. We built our measure on the premise that a script that is dissimilar to scripts developed before it must be accompanied by a logline that is dissimilar in meaning to the loglines of the preceding scripts. To measure the semantic similarity of loglines, we employed a vector embedding approach. Similar methods have recently been used to study changes in political discourse (Ash and Labzina, 2019), judicial text (Galletta et al., 2019), and US culture (Kozlowski et al., 2018).

Our data contain a sample of about 12K loglines from Sample DDP, each corresponding to a single script and containing more than six words (those with fewer than six words have been excluded from our analysis).[6] Each logline L_i can be represented as a set of its constituent words w_{ij}. Using the Word2Vec approach (Mikolov et al., 2013), each word w_{ij} is represented as a vector $\mathbf{W_{ij}}$ with dimension $1 \times n$, where $n = 300$. These word vectors are built such that words that appear in the same context in our training corpus are located close to each other in our vector space. These vector embeddings were obtained from the archives of the Natural Language Processing Group at the University of Oslo (Kutuzov et al., 2017). Word associations were learned from the fifth edition of the English Gigaword Corpus (Parker et al., 2011). In this vector space, words similar in meaning are represented as being closer to each other in semantic space. In addition, these vector embeddings allow for algebraic operations to be performed on words through their vector representations. For example, "Queen - Woman + Man = King," or "Moscow - Russia + USA = Washington." This allowed us to execute a variety of tasks involving analogies, similarities, and classification.

We created two novelty measures that differ in how the 'logline vector' ($\mathbf{Z_i}$)—the vector that represents a hypothetical average word that could be used to summarize logline L_i—was calculated. The first measure takes a simple average; that is, for each logline i, $\mathbf{Z_i}$ was calculated by taking the average of $\mathbf{W_{ij}}$'s of the logline's constituent words. We then calculated the vector $\mathbf{Y_y}$—taking the average of all logline vectors $\mathbf{Z_i}$ developed before year y. This vector represents the (hypothetical) average logline of all scripts developed before year y. Our novelty measure of each movie i (developed in year y) is the *cosine distance* between the logline vector $\mathbf{Z_i}$ and $\mathbf{Y_y}$. *cosine distance* $= 1 -$ *cosine similarity*, where *cosine similarity* $= \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}}$. Thus, the higher the cosine distance, the more 'novel' the script is.

To create the second measure, we followed the same steps as above, with the following changes to the calculation of our logline vector $\mathbf{Z_i}$. Following Arora et al. (2017), this measure recognizes the fact that some words appear out of context, and some more-common words appear in all sentences, regardless of context. To this end, we made two adjustments. First, we weighted each word in a logline by how frequently it appears in the logline corpus (all the loglines in our database). In particular, instead of a simple average, each word vector $\mathbf{W_{ij}}$ is weighted by weight k_{ij} in calculating $\mathbf{Z_i}$. The weight k_{ij} is equal to $\alpha/(\alpha + P(w))$, where $\alpha = 0.001$ is a smoothing parameter following the value used in Arora et al. (2017); and $P(w)$ is the probability of finding word w_{ij} in the logline corpus (that is, the number of times word w_{ij} occurred in the logline corpus divided by the total number of words in the logline corpus). Thus, words used more frequently are weighted less. Second,

[6] As explained in the paper, not all DDP records have a complete script at the time of the record, but they all have loglines.

we boost the co-occurrence probability of words that have a high component along the common discourse of the logline corpus in order to ensure that infrequent words that are not closely related to the main theme of the loglines are deemphasized. Specifically, we form a matrix $\mathbf{X}$ whose columns are weighted logline vectors $\mathbf{Z_i}$ and let $\mathbf{u^T}$ be its first singular vector. For each logline vector $\mathbf{Z_i}$ in our matrix $\mathbf{X}$, we further define $\mathbf{Z_i'}$ as the eventual logline vector to be used as $\mathbf{Z_i} - \mathbf{uu^T Z_i}$, which intuitively removes words not closely related to the main themes of our logline corpus.

References

Arora, Sanjeev, Yingyu Liang, and Tengyu Ma (2017). "A Simple But Tough-to-beat Baseline for Sentence Embeddings" *Conference paper at ICLR 2017* https://openreview.net/pdf?id=SyK00v5xx.

Ash, E. and E. Labzina (2019). "Does Politicized Media Distort Political Discourse? Evidence from U. S. Cable News Channels," *Working Paper*.

Galletta, S., E. Ash, and D. L. Chen (2019). "Do judicial sentiments affect social attitudes?" *Working Paper*, Available at SSRN 3415393.

Kozlowski, A. C., M. Taddy, and J. A. Evans (2018). "The geometry of culture: Analyzing meaning through word embeddings." arXiv preprint arXiv:1803.09288.

Kutuzov, A., M. Fares, S. Oepen, and E. Velldal (2017). "Word vectors, reuse, and replicability: Towards a community repository of large-text resources, " *Proceedings of the 58th Conference on Simulation and Modeling* 271-276. Linköping University Electronic Press.

Mikolov, T., K. Chen, G. Corrado, and J. Dean (2013). "Efficient Estimation of Word Representations in Vector Space, " arXiv e-prints, arXiv:1301.3781.

Parker, R., D. Graff, J. Kong, K. Chen, and K. Maeda (2011). "English gigaword." fifth edition, June, Data Consortium, LDC2011T07 12.

Online Appendix D. Supplementary Results for Conceptual Framework

D.1 A Micro-foundation of Assumption 2

In this section, we describe a very simple writer's maximization problem that provides one out of potentially multiple ways to micro-found Assumption 2 in the paper that the breadth of circulation may be greater for less-experienced writers.

Consistent with the model setup in the paper, the inherent quality of a movie script is either high or low: $q \in \{q_H, q_L\}$. The prior probability that a script is high-quality is a positive function of the writer's experience level, $w \in \mathbb{R}$; that is, $\pi'(w) > 0$ (Assumption 1 in the paper). For simplicity, we assume that the writer does not possess any private information about the quality of his script; the writer and the buyers share a common prior. A potential buyer, having read a given script, obtains a binary quality signal $s \in \{s_H, s_L\}$, where $\mathrm{P}(s_H|q_H) = p_1$ and $\mathrm{P}(s_H|q_L) = p_2$. Assume that $p_1 > p_2$; that is, the likelihood of obtaining a positive signal (s_H) is greater when the true quality is high than when the true quality is low. For simplicity, we also assume that buyers (and, hence, their signal structures) are homogeneous and that any two buyers' signals are conditionally independent given the true script quality.

The writer's problem is to choose the number of buyers, n, to whom he/she will simultaneously market a script to in order to maximize the expected payoff.[7] Assume that the writer will make a sale if at least one out of the n buyers gets a good signal, s_H.[8] The sale probability is, thus, $(\pi(w)(1-(1-p_1)^n) + (1-\pi(w))(1-(1-p_2)^n))$. In each state of the world (i.e., the true quality being either q_H or q_L), the probability that at least one out of n buyers draws a good signal is one minus the probability that all n buyers draw bad signals. This probability is $(1-(1-p_1)^n)$ when $q = q_H$ and $(1-(1-p_2)^n)$ when $q = q_L$. The probability of a sale, thus, weights the two states by their respective prior probabilities. For tractability, we assume that the writer captures a fixed share (α) of q_H, which is the value of a high-quality project, if a script is sold. Finally, marketing to an additional buyer is costly; and the marginal cost is c. As discussed in the paper, apart from monetary costs, disutility of effort, and the opportunity cost of time, industry accounts also suggest that sellers have incentives to avoid an unnecessarily wide sales scope due to concerns of information leakage.

Given the above setup, the writer maximizes his expected payoff:

$$\max_n [\pi(w)(1-(1-p_1)^n) + (1-\pi(w))(1-(1-p_2)^n)]\alpha q_H - nc. \tag{A5}$$

The first derivative of the writer's expected payoff in equation (A5) with respect to n is

$$F(n;w) = [-\pi(w)(1-p_1)^n \ln(1-p_1) - (1-\pi(w))(1-p_2)^n \ln(1-p_2)]\alpha q_H - c.$$

For a writer of experience level w, the optimal number of buyers to approach is determined by the first-order condition $F(n;w) = 0$. Note that because $1-p_1 < 1$ and $1-p_2 < 1$, the sum of the two terms in the brackets

[7] Note that the model describes a single decision-maker problem; the buyers are not actual players in a game-theoretic sense, as their signals and buying decisions are mechanical.

[8] Note that this assumption greatly simplifies the sale probability. Alternatively, the buyers, upon drawing a signal, may update their belief about the probability that the script is high-quality based on the prior. Incorporating such updating will introduce an additional reason that less-experienced writers face a lower sale probability, as, given the same signal, the buyer would infer a lower probability of high quality for less-experienced writers. Our results do not change qualitatively upon incorporating such updating.

is positive.

Notice that the second derivative of the expected payoff with respect to n is

$$\frac{\partial F(n;w)}{\partial n} = [-\pi(w)(1-p_1)^n(\ln(1-p_1))^2 - (1-\pi(w))(1-p_2)^n(\ln(1-p_2))^2]\alpha q_H < 0. \qquad \text{(A6)}$$

Therefore, for a given w, there exists a unique n^* that satisfies the first-order condition of $F(n^*;w) = 0$.

Denote $n(w)$ as the function of the optimal number of buyers to market to for a writer of experience level w. We are interested in deriving the comparative statics of how $n(w)$ varies with w. By the implicit function theorem,

$$n'(w) = -\frac{\dfrac{\partial F(n;w)}{\partial w}}{\dfrac{\partial F(n;w)}{\partial n}}.$$

Equation (A6) shows that the denominator in the above equation is negative. The numerator (i.e., the derivative of $F(n;w)$ with respect to w) is:

$$\frac{\partial F(n;w)}{\partial w} = \frac{\partial F(n;w)}{\partial \pi(w)}\pi'(w) = (-(1-p_1)^n\ln(1-p_1) + (1-p_2)^n\ln(1-p_2))\alpha q_H \pi'(w). \qquad \text{(A7)}$$

Under Assumption 1, $\pi'(w) > 0$. Thus, the sign of $n'(w)$ is consistent with the sign of $-(1-p_1)^n\ln(1-p_1) + (1-p_2)^n\ln(1-p_2)$, which is negative if and only if the following condition holds:

$$\begin{aligned} -(1-p_1)^n\ln(1-p_1) + (1-p_2)^n\ln(1-p_2) < 0 \quad &\Leftrightarrow \quad (1-p_2)^n\ln(1-p_2) < (1-p_1)^n\ln(1-p_1) \\ &\Leftrightarrow \quad \left(\tfrac{1-p_2}{1-p_1}\right)^n > \tfrac{\ln(1-p_1)}{\ln(1-p_2)} \\ &\Leftrightarrow \quad n\ln\left(\tfrac{1-p_2}{1-p_1}\right) > \ln\left(\tfrac{\ln(1-p_1)}{\ln(1-p_2)}\right) \\ &\Leftrightarrow \quad n > \ln\left(\tfrac{\ln(1-p_1)}{\ln(1-p_2)}\right)/\ln\left(\tfrac{1-p_2}{1-p_1}\right). \end{aligned}$$

Note that we can preserve the direction of the inequality sign in the fourth line because $\ln\left(\frac{1-p_2}{1-p_1}\right) > 0$ when $p_1 > p_2$. Let $\bar{n} = \ln\left(\frac{\ln(1-p_1)}{\ln(1-p_2)}\right)/\ln\left(\frac{1-p_2}{1-p_1}\right)$. Thus, we have $n'(w) < 0$ if the optimal number of buyers to market to is greater than $\bar{n}$. $\bar{n}$ depends on the values of p_1 and p_2. Numerical examples provide a sense of how stringent this requirement is. For example, suppose that $p_1 = 0.6$, which implies that if the idea is good, 60% of the time the signal is good, and 40% of the time the signal is bad. Numerical simulation shows that if p_2 takes any value between the range [0.1,0.6) such that $p_2 < p_1$, the maximum of $\bar{n}$ is less than 3. Similarly, if $p_1 = 0.7$, the maximum of $\bar{n}$ for $p_2 \in [0.1, 0.7)$ is also less than 3. This suggests that at these values for $p1$ and $p2$, we need the optimal number of buyers that the writer needs to approach to only be greater than three in order be consistent with Assumption 2. It is certainly conceivable that a writer would need to market to more than three buyers to sell a script.

In summary, the above model provides a simple, yet intuitive, way to derive the breadth of circulation of a script as the writer's endogenous choice as he/she attempts to sell the script. We show that the optimal number of buyers to market to decreases with the writer's experience level under plausible conditions. This result, thus, provides a micro-founded support for Assumption 2. It is also quite intuitive: less-experienced writers want to approach more buyers, even though it is more costly, because the chances of any individual buyers receiving a positive signal about their scripts are lower. It is, however, important to note that this model makes several simplifying assumptions. In addition, it misses other potential mechanisms through which Assumption 2 may hold. For example, less-experienced writers may face a lower opportunity cost of time, and they may be worse at targeting buyers that may like their scripts.

D.2 The Studio's Decision to Produce a Movie

In this appendix, we explicitly model the studio's decision to produce a movie, with the objective of deriving the expected revenues of movies conditional on production. In particular, the following result provides a more-formal foundation that motivates the null hypothesis stated in the paper—*If the Black List is not predictive, conditional on released movies, listed scripts and unlisted scripts should generate similar box office revenues, given the same physical and human capital allocated to them*. This null hypothesis helps to better understand whether the Black List is purely causal without being predictive of quality.

A simple way to model the production decision is that the studio will produce a movie if its expected revenue is greater than the production cost. Let the expected revenue of a project be $V(m,w)-\varepsilon$, where $V(m,w)=\mathrm{P}(q_H|m,w)q_H$: that is, the updated probability that the idea is of high quality, $\mathrm{P}(q_H|m,w)$ as is defined in equation (2) in the paper, multiplied by the value of a high-quality project, q_H. We add a random factor $\varepsilon \sim F$ to generate a probability of production. The project's expected revenue is, thus, a function of the number of nominations, m, and the writer's experience level w. Note that $\frac{\partial V(m,w)}{\partial m}=\frac{\partial \mathrm{P}(q_H|m,w)}{\partial m}q_H>0$, according to Proposition 1, and $\frac{\partial V(m,w)}{\partial w}=\frac{\partial \mathrm{P}(q_H|m,w)}{\partial w}q_H>0$, according to Proposition 3. Finally, let C be the production cost.

The studio will produce a project if and only if the expected revenue is greater than the production cost:

$$\mathrm{P}(\text{Produce}|m,w)=\mathrm{P}(V(m,w)-\varepsilon\geq C)=\mathrm{P}(\varepsilon\leq V(m,w)-C)=F(V(m,w)-C).$$

The expected revenue, conditional on being produced, is then

$$\mathrm{E}[V(m,w)-\varepsilon|V(m,w)-\varepsilon\geq C]=\int_{-\infty}^{V(m,w)-C}(V(m,w)-\varepsilon)\frac{f(\varepsilon)}{F(V(m,w)-C)}d\varepsilon. \tag{A8}$$

To see how the expected revenue compares between listed and unlisted movies, we take the derivative of the above equation with respect to m. By Leibniz's rule, we have

$$\begin{aligned}
&\frac{\partial \mathrm{E}[V(m,w)-\varepsilon|V(m,w)-\varepsilon\geq C]}{\partial m}\\
=\;&\int_{-\infty}^{V(m,w)-C}[\frac{\partial V(m,w)}{\partial m}\frac{f(\varepsilon)}{F(V(m,w)-C)}-(V(m,w)-\varepsilon)\frac{f(V(m,w)-C)f(\varepsilon)}{F^2(V(m,w)-C)}\frac{\partial V(m,w)}{\partial m}]d\varepsilon+C\frac{f(V(m,w)-C)}{F(V(m,w)-C)}\frac{\partial V(m,w)}{\partial m}\\
=\;&\frac{\partial V(m,w)}{\partial m}\{\int_{-\infty}^{V(m,w)-C}[\frac{1}{F(V(m,w)-C)}-(V(m,w)-\varepsilon)\frac{f(V(m,w)-C)}{F^2(V(m,w)-C)}]f(\varepsilon)d\varepsilon+C\frac{f(V(m,w)-C)}{F(V(m,w)-C)}\}\\
=\;&\frac{\partial V(m,w)}{\partial m}\{1-\frac{f(V(m,w)-C)}{F(V(m,w)-C)}(V(m,w)-E[\varepsilon|\varepsilon\leq V(m,w)-C]-C)\}.
\end{aligned} \tag{A9}$$

In the following, we discuss two scenarios: (i) if the Black List is predictive of quality (that is, $\frac{\partial V(m,w)}{\partial m}>0$); and (ii) if the Black List is not predictive of quality (that is, $\frac{\partial V(m,w)}{\partial m}=0$).

D.2.1 If the Black List is predictive of quality; i.e., $\frac{\partial V(m,w)}{\partial m} > 0$

Under the premise of $\frac{\partial V(m,w)}{\partial m} > 0$, the sign of equation (A9) will depend on the sign of the terms in the large brackets. In general, it will *not* be zero except in special cases (i.e., if $1 = \frac{f(V(m,w)-C)}{F(V(m,w)-C)}(V(m,w) - E[\varepsilon|\varepsilon \leq V(m,w) - C] - C)$).

As m increases, even though the marginal project will have the same expected revenue (which is the production cost), how the expected revenue compares between listed and unlisted projects will also depend on *infra-marginal* projects, whose expected revenue will depend on the shape of the conditional distribution, which itself depends on F and the values of other parameters. Intuitively, the sign of equation (A9) will be positive if the conditional distribution for scripts that receive more nominations is statistically more favorable. The opposite may also be possible if the conditional distribution for scripts that receive more nominations is statistically less favorable.

To get some intuition, Figure C3 provides some numerical examples, assuming that ε follows a normal distribution. For simplicity, we do not consider different writer experience; in our regression results on box-office performance, we control for writer experience. The different panels of Figure C3 differ by the variance of the normal distribution and the production cost. For these examples, at least, conditional on production, listed movies (those with nominations above a certain threshold) will have a higher expected revenue than unlisted movies (those with nominations below a certain threshold).

D.2.2 If the Black List is not predictive of quality; i.e., $\frac{\partial V(m,w)}{\partial m} = 0$

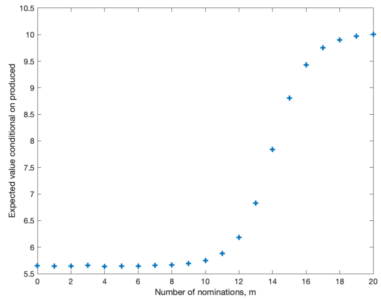
If $\frac{\partial V(m,w)}{\partial m} = 0$, equation (A9) would be zero, suggesting that the expected revenue, given the same physical and human capital allocated to the projects, should be similar between listed and unlisted scripts. As we discuss in the paper, results on the expected box-office performance for released movies (Section 4.1.1) reject this null hypothesis.

It is worth noting why we don't view the positive relationship between box-office performance and whether listed or not, conditional on production, to be consistent with the purely-causal (without being predictive of quality) interpretation, even though a positive correlation between the likelihood of release and whether listed might be. Recall that in the paragraph, "The third and final possibility ..." on page 9 of the paper, we explain that if the Black List is not predictive at all—that is $\frac{\partial \mathrm{P}(q_H|m,w)}{\partial m} = 0$—the List may still result in a higher likelihood of being produced. This may happen (i) if the studio (mistakenly) assumes that voter signals are informative of true quality; or (ii) if the Black List, by providing a focal point that attracts attention, helps to overcome any coordination failures that may occur as the producer assembles the resources.

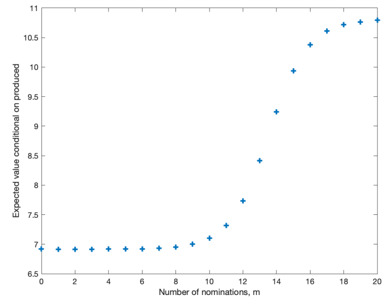
These two reasons are unlikely to hold for the relationship between box office performance and whether listed (conditional on production) for two reasons. First, a pure coordination effect would likely imply greater resources devoted to a project. In the regressions, we not only control for the production budget, but also explicitly show that the production budget for produced movies is not significantly greater for listed movies. Second, even though the mistaken belief that the List carries discriminating information may influence the decision of whether to produce, once one conditions on the movie being financed and the cast and crew being already assembled, it is unlikely that people will pay attention to the fact that it is listed. Furthermore, consumers are also unlikely to know that a movie is from a Black-Listed script. Thus, the differences between listed and unlisted movies, conditional on the same resources and talent, are likely to be rooted in the

Figure C3: Illustration of Conditional Means by Nominations

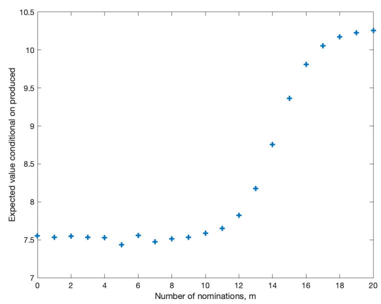
(a) $C = 5$ and $\sigma = 2$



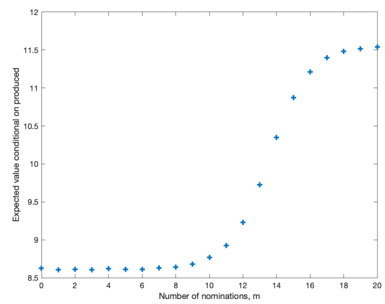
(b) $C = 5$ and $\sigma = 4$



(c) $C = 7$ and $\sigma = 2$



(d) $C = 7$ and $\sigma = 4$



Note: The horizontal axis is the number of nominations, and the vertical axis is the conditional expected revenue of projects that are produced. The common parameters are: n (the number of voters) = 20; $\pi = 0.5$, $p_1 = 0.6$, $p_2 = 0.4$, and $q_H = 10$; and ε is a normal distribution with mean 0. Different figures vary by the variance of the normal distribution and the production cost. The conditional mean for each m is calculated numerically with 1 million random draws from the specified normal distribution.

underlying quality of the script rather than in coordination failure or mistaken beliefs.