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

Hong Luo[†] Jeffrey Macher[‡] Michael Wahlen[§]

January 10, 2019

(preliminary draft)

Abstract

This paper studies judgment aggregation from a large number of industry experts in the selection and development of early-stage ideas. Our context is the movie industry, where customer interest is notoriously unpredictable and investment costs are relatively high. The Black List, an annual publication, ranks unproduced scripts based on anonymous nominations from 250-300 films executives. We find that listed scripts are twice as likely to be released as unlisted scripts, with evidence to support a causal interpretation, and that the aggregated judgment is able to discern good ideas from poor ones that are ex-ante similar. We also find that scripts from less-experienced writers are more likely to be listed and to rank higher if listed. Yet, in terms of release outcomes, even though being listed has a large and positive effect on less-experienced writers, the discrepancy relative to experienced writers remains large, even for the top-ranked scripts. These results reveal a simple, albeit important, determinant of the data-generating process of such aggregations—the visibility of an idea among potential evaluators.

[*]We thank Heski Bar-Issac, Dan Gross, Aseem Kaul, Ramana Nanda and seminar participants at Harvard Business School, Strategic Research Forum, the University of Toronto for helpful comments. Tejas Ramdas and Esther Yan provided excellent research assistance. All errors and omissions are ours.

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

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

[§]MIT Sloan; email: jmwahlen@mit.edu.

1 Introduction

Successfully selecting and evaluating early-stage ideas is immensely important and yet extremely difficult. This is especially the case in many domains—such as creative industries and commercial technology startups—in which there is a lack of concrete, objective quality signals (Scott et al. (2018)).[1] Even worse, the number of ideas in these markets tend to be large with the majority of them being of little value. To address these challenges, one often focuses on ideas from those with strong, observable track records or relies on reputable gatekeepers—such as venture capital firms and talent agencies—to source and evaluate ideas (Caves (2000); Stuart and Sorensen (2010); Hellmann and Puri (2002)). Research on this topic generally suggests that funding decisions under such regimes are often subject to the various frictions faced by the small number of decision makers and exhibit high barriers to entry for those lacking prior experience or industry connections (Caves (2000); Hall and Lerner (2010)).[2]

An alternative approach, demonstrated by the increasing popularity of crowd-funding platforms, is to leverage the judgment of crowds (Mollick (2014); Agrawal et al. (2015)). This approach bears the promise of breaking the funding decisions away from a small number of gatekeepers and, thereby, lowering the barriers to entry for outsiders (Mollick and Nanda (2015)).[3] So far, typical crowd-funded projects require only small budget; and, as argued by **?**, crowds tend to require the help from more-skilled investors or other experts in order to make larger investments.[4] It remains unclear whether and to what extent collective judgment based on amateur crowds is able to influence large-budget, mainstream funding and production decisions.

In this paper, we study a novel approach that combines large-scale aggregation with opinions from individuals whose judgment carries certain credibility—that is, they are industry experts

[1] Scott et al. (2018) find that experts can differentiate among early-stage ventures based on written summaries, but only in fields with high R&D intensity.

[2] Even with the aid of intermediaries, however, the barriers to entry for ideas from outsiders can remain high due to their relationship-based nature.

[3] In their examination of theatre projects, Mollick and Nanda (2015) find that crowds' evaluations generally coincide with those of trained experts and that crowds may lower the incidence of "false negatives" (that is, viable projects that may be turned down by experts).

[4] For example, among successfully funded projects on Kickstarter, the median raised amount is less than \$10K and only 0.21 percent more than \$1m. Typical contributions are at the level of tens of dollars. When it comes to equity crowd-funding, the minimum investment scale is in the thousands of dollars, but even there, the average amount raised by successfully-funded ventures, on AngelList.com in 2015 for example, was \$369K, which is substantially smaller than the type of projects we study in this paper. Sources: https://www.kickstarter.com/help/stats; https://angel.co/2015/fundraising.

rather than amateur crowds. Our empirical context is the movie industry, in which the appeal to customers is notoriously difficult to predict (the famous claim of 'nobody knows anything' by Goldman (1983); and Waldfogel (2017)) and whose investment costs are very high (the median production budget for released movies is $25m in our sample). The Black List—the information producer we study—is an annual, anonymous survey of film executives that began in 2005. Every year, 250-300 film executives nominate up to ten of their favorite scripts that have not yet been made into movies. Scripts with more than a certain number of nominations are published in a single list and ranked by the number of nominations received.

Two features of the Black List are worth noting. First, the eligible voters—film and studio executives with at least one recent major release—are relevant decision makers in the industry. Second, in contrast to a structured and standardized process (used by grant-making institutions, for example), the Black List takes a light-touch approach, which may partly explain the high participation rate by busy executives. For example, instead of centrally allocating scripts to judges, the survey leverages the existing sales and development processes of movie projects. As a result, the set of voters who are at risk of nominating a particular script are those who have actually read the script when it is shopped around. Furthermore, the voters simply provide the names of the scripts and do not need to comply with a uniform or an extensive set of rating criteria.

With this setting, we focus on two questions. First, we examine whether such an aggregation system actually influences the allocation of industry resources and, hence, the likelihood of a given script being produced. Even though the wide participation of business executives may lend credibility to the list, the answer is not obvious given its lack of structure and uniform criteria. Second, we seek to understand whether it helps lower entry barriers for less-experienced writers—specifically, to what extent the list surfaces new talent or simply reinforces established insiders; and, importantly, whether being listed and a top place on the list helps close the gap for less-experienced writers relative to established ones in terms of the final outcomes.

Answering the above questions is challenging, however, due to the lack of exogenous variation and credible control industries. We address these difficulties by employing a number of different methods and by interpreting a holistic set of evidence based on multiple aspects of the data. In particular, to aid us in understanding the data, we describe a stylized model of judgment aggregation. The model focuses on two factors: (i) the ex-ante quality of a script (an assessment based on the characteristics of the script and its writers observable without reading it); and (ii) the script's visibility among potential voters for the Black List (i.e., the number of film executives who have actually read a particular script before the survey). This simple model yields predictions on two stages of the data—how the Black List nominations depend on the above two factors; and how decision makers may update their beliefs of a script's quality given the realized nominations. While the model by no means captures all the important factors, it does provide a useful guide for our understanding of key data patterns and exploration of alternative mechanisms.

Our empirical analysis is based on two samples that capture different segments of the nomination distribution. The first sample includes all Black-Listed scripts over 2007-2014. This sample provides the exact number of nominations and a detailed picture of the top-end distribution (i.e., movie scripts above the cut-off threshold). The second sample uses Done Deal Pro (DDP), a comprehensive Internet database tracking early-stage movie projects, and includes completed scripts that had already been sold over the same 2007-2014 period (of these, 13 percent were Black-Listed). By including unlisted scripts, the second sample extends to a lower part of the nomination distribution, though is still conditional on scripts that were good enough to be sold. By complementing these data with additional data from the Internet Movie Database (IMDb) and The Numbers database, we are also able to observe the writer's industry experience and the outcome of a project (e.g., whether it is eventually released in theaters and, if so, its box office revenues). These performance data provide valuable information that sheds light on potential mechanisms.

Our first main finding is that, on average, Black-Listed scripts are twice as likely to be released as unlisted scripts; and there is evidence favoring a causal interpretation of this difference. There are at least two challenges in identifying a causal effect: (i) scripts with a higher *ex-ante* quality (observable to decision makers but not to us) are both more likely to be listed and more likely to be released; and (ii) decision makers may already know how others view a script by other means before the list's publication. Our main method assumes that projects purchased by the same buyer in the same year (a portfolio)—once conditional on observable characteristics—are likely to have similar expected values prior to the Black List publication. This assumption seems reasonable given the reputed unpredictability of early-stage movie projects. The estimated effect is robust across this and additional methods—using only marginally listed scripts and coarsened exact matching—that restrict the comparisons to scripts that are likely to be ex-ante similar. Furthermore, we show that the estimated effect is driven exclusively by scripts that did not receive much media attention before the list's publication, consistent with the Black List containing a greater amount of *new* information that helps reduce uncertainty for these scripts than for scripts that were reported.

It is important to note that the above estimated difference between listed and unlisted scripts includes the potential crowding-out effect on the latter due to resources being allocated away from them. Drawing on performance data that are available only for released movies, we show that, on average, listed movies generate a significantly higher box office revenues than movies with the same budget and, hence, a higher return over investment. This result is consistent with the notion that the Black List's aggregation of individual judgment helps separate good ideas from poor ones. Additional evidence further supports the interpretation that the estimated effect is not simply a coordination effect—that is, simply by providing a focal point that attracts attention, the Black List may help coordinate industry resources and set priorities, even if it is common knowledge that listed scripts are not necessarily better.

Our second set of findings contrast scripts by experienced and less-experienced writers, both in terms of the probability of being listed and the effects of the Black List on a script's probability of being released as a film. The Black List appears quite effective at highlighting scripts from relatively novice writers—sold scripts from less-experienced writers are significantly more likely to be listed and, if listed, ranked higher. However, when it comes to the probability of being released, even though being listed has a large and positive effect for less-experienced writers, the discrepancy relative to experienced writers remains large, even for the top-ranked scripts.

Our conceptual model can rationalize both of the above patterns if scripts from less-experienced writers are on average more visible among potential survey respondents. This explanation is plausible because unlike experienced writers, who can sell their ideas without talking to many buyers, less-experienced writers need to offer their scripts to a wider audience in order to find interested buyers and the best match. Thus, nominations obtained by less-experienced writers may be systematically inflated relative to experienced writers in terms of the script's inherent quality, explaining the negative correlation between writer experience and Black List nominations. Consequently, decision makers would infer an inferior posterior quality for scripts from less-experienced writers, believing that they have been shopped around widely. This adjustment can, therefore, explain the lower release likelihood for inexperienced writers, even when their scripts receive the same number of nominations as those by experienced writers. Additional results on the roles of agencies and different sales methods in determining Black List nominations are also consistent with the importance of a script's visibility among potential voters in its data-generating process.

An alternative explanation proposed by the model is that scripts from less-experienced writers are more likely to be exceptionally good (e.g., if the quality distribution of their scripts has a greater variance), which could also explain the negative correlation between writer experience and Black List nominations. This explanation, however, requires factors outside the model to account for the large discrepancy in the release probability despite a higher posterior belief of script quality for scripts from less-experienced writers. We discuss two such factors: the first is risk aversion, as producers may be skeptical of scripts by inexperienced writers even though they were listed; second, it possible is that scripts from less-experienced writers are 'stuck' with either less-experienced or inadequately-financed producers, because much of the matching between producers and scripts has already been determined by the time the annual lists are published. We conduct further analysis on both of these possibilities and do not find evidence supporting these arguments.

This paper contributes to a small but growing body of work that focuses on the role of supporting institutions and processes in the evaluation of early-stage ideas. As noted above, one stream of this reasearch examines the role of experts, often in the context of peer reviews of scientific research proposals: Li (2017) shows that reviewers for the National Institutes of Health are both better informed and more (positively) biased about the quality of projects in their own area; whereas Boudreau et al. (2016), using data from a field experiment at a leading university, finds that the evaluators systematically give lower scores to proposals that are closer to their own expertise or highly novel. As mentioned previously, experts' ability to differentiate among early-stage ventures appears to be limited when moving away from fields with high R&D intensity (Scott et al., 2018). Astebro and Elhedhli (2006) and Astebro and Koehler (2007) suggest that when expert panels are found to be effective in evaluating these types of ideas, structured and standardized processes and extensive training of evaluators appear to be necessary. A second stream of research on this topic, also mentioned above, focuses on crowd-based evaluation (Mollick (2014); Agrawal et al. (2015); Mollick and Nanda (2015), often set in creative or consumer technology projects that require relatively small budgets. More recently, using proprietary data from a large accelerator, Catalini et al. (2018) compare expert evaluation and machine learning algorithms based on market-wide data, suggesting that experts seem to overly rely on founder characteristics. As noted above, our paper documents a novel mechanism that combines large-scale aggregation with opinions from industry experts, which appears to influence the allocation of mainstream, large-budget production in a creative industry. The paper also highlights a simple, albeit important way an unstructured aggregation approach such as the Black List may systematically bias the aggregated judgment—by reflecting differences in the visibility of an idea among potential evaluators.

Our paper also relates to studies in both economics and sociology on the importance of observable quality signals—such as reputation, status, identity, or affiliation—when the quality of ideas or products is uncertain; and how mechanisms that reduce this uncertainty may reduce the importance of these observable signals. In the context of internet platforms, for example, Luca (2011) shows that chain restaurants have declined in market share as Yelp penetration has increased, as its review system reduces the uncertainty over restaurant quality. In the context of scientific evaluations, Simcoe and Waguespack (2011) find causal evidence for a large positive effect of name-based signals on the publication of a typical submission but not for a set of prescreened proposals that receive more scrutiny, suggesting that status signals are less important when quality of ideas is less uncertain. Similarly, Azoulay et al. (2013) show that the effect of a major status-conferring prize on the perceived quality of a scientist's past research is smaller when there is less uncertainty about the article's quality. From these works, one might expect the Black List to also lower the importance of the writers' observable characteristics as it isolates a small set of projects for which industry insiders reach greater consensus about their underlying quality. Yet, for reasons we unpack below, we find starkly different results.

The rest of the paper is organized as follows: Section 2 provides background information on the movie industry and the Black List; section 3 presents a simple conceptual model; section 4 describes the two samples, descriptive results, and our empirical approach; section 5 presents the regression results and discusses potential mechanisms; and section 6 concludes.

2 Background

2.1 The movie industry

Making a movie is a long, costly, and uncertain process. Our data show that only 15 percent of purchased scripts are eventually theatrically released. For released movies, the median time from the sale of the script to release is 2.1 years; the median production budget, not including distribution and marketing costs is $25m, and the 5th and 95th percentiles are, respectively, $5m and $150m. The complexity of this process is due partly to the fact that it requires convincing and coordinating multiple, separate parties: the producer needs to attract a suitable director and key acting talent; coordinate their schedules; and, importantly, convince the studio executives to finance the project. Today, movie studios such as Universal Pictures act mainly as financiers, distributors, and marketers, while the projects themselves are developed and overseen by independent or semi-independent producers, who are typically attached to a movie studio through a 'first-look' contract.[5]

The process begins with a screenplay or script, which we use synonymously in this paper. Broadly speaking, scripts come from two sources: producers may acquire (option or purchase) finished screenplays from screenwriters, or they may hire screenwriters to adapt ideas into screenplays (e.g., based on a novel for which the producer has obtained adaptation rights). With tens of thousands of new scripts entering the industry every year, producers rely on various mechanisms, such as the writer's reputation and track record, to weed out a large number of low-potential ideas.

Agents—intermediaries representing writers and other talent—have long been important industry gatekeepers.[6] When selling a script, the reputation, experience, and connections of the writer's agent (and agency) are critical to identifying potential buyers and generating interest. Industry observers describe this process as very much relationship-based and suggest that producers seldom read scripts from agents they do not trust. Agencies that represent both literary and acting talent may also leverage their other clients, assembling a director and acting team to sell an entire 'package.'

Obtaining an agent's representation, especially that of a reputable agent at a large agency, can

[5]Through the contract, the studio will provide a development budget in exchange for the right of first refusal over a project developed within the duration of the contract (typically two or three years).

[6]The industry norm is that an agent-talent contract lasts for two years, with the agent's commission pegged at 10 percent of the talent's gross income.

be challenging. Well-connected agents are busy and, therefore, rely on referrals and typically do not read unsolicited materials submitted by unknown writers. Scripts do get acquired without agent representation, however: this may happen if the writer is him/herself an industry insider; if he/she knows some producers or their assistants; or if the script has placed at the top in well-known screenwriting competitions. Finally, it is important to note that, like writers, producers also vary greatly in their experience, status, and connections. Many smaller producers face significant barriers to accessing materials controlled by large agencies and, therefore, are more willing to read materials by novice writers.

2.2 The Black List

Franklin Leonard founded the Black List in 2005 as a means of identifying high-quality, unproduced scripts.[7] The List is published annually in December and compiled from suggestions by respondents of a simple survey. Eligible voters are "executives at major studios, major financiers, or production companies that have had a film in major release in the past three years."[8] Each is asked to suggest "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." According to Leonard, the response rate is slightly less than 50 percent, resulting in 250-300 responses in most years. The published list ranks scripts by the number of nominations each receives, and only scripts whose nominations total above a certain threshold make the list.[9] Appendix C.1 provides a screen shot of the screenplays that received the most nominations in 2005.

Anecdotal evidence suggests that the Black List's stamp of approval can be crucial. When asked why being listed may help, Helen Estabrook, one of the producers 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.'" It "reassures financiers, executives, and producers that they are not going too far out on a limb." (*The Atlantic*, 2017). Additionally, Black List ranking can potentially

[7] A junior executive at a production company at the time, Leonard was supposedly frustrated by the inefficiencies and difficulties of finding good scripts, so he turned to his peers for input. "I know I'm going to spend a ton of time reading scripts, and I'd like for them not to be terrible," said Leonard in an interview with the *LA Times*. "Franklin Leonard's Black List can help green-light screenplays." Josh Rottenberg, *LA Times*, December, 2014.

[8] Source: interview by John Horn for *The Frame*, December 12, 2017; https://www.scpr.org/programs/the-frame/2017/12/12/60656/getting-on-the-black-list-won-t-guarantee-your-scr/

[9] The threshold number of nominations increased from one in 2005, two from 2006-2007, 4-5 between 2008-2010, and has remained at 6 since 2011.

attract attention from actors and directors. Meryl Streep, Ben Affleck, and Benedict Cumberbatch reportedly found the films *Hope Springs*, *Argo*, and *The Imitation Game*, respectively, directly from the Black List.[10]

Because the survey takes place only annually, the majority of the listed scripts (about 80 percent) are already associated with a production company and/or a movie studio by the time the List is published. A potential voter typically reads a script when it is being shopped around, either before it is acquired by a producer or as the producer tries to secure financing or to generate interest among directors and actors who may also qualify to vote because of their own production companies. According to an article in *The Atlantic*, "scripts have to find their way into the industry pipeline before they can make the list;" in other words, a script must get "into the hands of executives so that they may, in turn, like it and vote for it."[11] A *Slate* article also made the point that "for a script to top that list, it needs to have been read by many of those people."[12] It is, therefore, not clear to what extent the List surfaces new talent or simply reinforces existing selection processes.

Our data show that 28 percent of all listed screenplays are subsequently theatrically released, which is consistent with what Leonard reports. Black-Listed movies have also been successful in terms of recognition and awards, having been nominated for 241 Oscars and won 48 (McGee and Mcara, HBS Case). Thus, even though some are skeptical that such movies would not have been successful in the absence of the List (e.g., *Slate's Culture Blog*, 2011), it is no surprise when the *LA Times* declared in 2014 that "[t]he Black List has become a Hollywood institution," and *Fast Company* named Leonard as "one of Tinseltown's youngest kingmakers."

3 Conceptual Framework

Before moving on to the data and results, we describe a simple, stylized model of judgment aggregation to aid in the interpretation of the data and the exploration of potential mechanisms. With this framework, we intend to capture some of the basic institutional details about how the Black List works described in the previous subsection. We discuss important factors left outside this simple model in section 5 as we unfold the results.

[10]Sources: "The Black List," HBS Case #9-317-027, by Henry McGee and Sarah Mcara; Hollywood 'Black List' pushing Oscar-worthy scripts to spotlight," and *CBS News*, January 13, 2016.

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

[12]"The Mostly Dull-Sounding Screenplays on This Year's 'Black List'," David Haglund, *Slate's Culture Blog*, 2011.

Assume that a movie script's inherent quality is either high or low: $q \in \{q_H, q_L\}$. Generally speaking, for creative projects, quality could mean different things to different people. Given that the outcome variables this paper focuses on are the probability of theatrical release and box office performance, we use the word 'quality' to mean a script's likelihood of commercial success. The common prior belief is that a script is high-quality with probability π. This probability is based on the script's (or the writer's) observable characteristics *before* one actually reads it, including the writer's track record; general script characteristics (e.g., genre); and a brief description (e.g., a poor man and an upper-class woman fall in love on an ill-fated ship).

After reading a script, each voter obtains a binary quality signal $s \in \{s_H, s_L\}$, such that $\mathrm{P}(s_H|q_H) = \mathrm{P}(s_L|q_L) = p$ and $\mathrm{P}(s_L|q_H) = \mathrm{P}(s_H|q_L) = 1 - p$. We assume that $p > 1/2$; that is, the signals are informative because they are more likely to reflect the true quality of the script. It is certainly possible that voters, by and large, receive uninformative signals as if they flip a coin (i.e., $p = 1/2$). We discuss this scenario later in section 5.2.1. For simplicity, assume that the voters' signals are independent of each other; and that a voter will nominate a script as long as he or she receives a good signal (s_H).

Denote by n the number of voters surveyed by the Black List who have actually read a particular script (i.e., the script's visibility). The notion of visibility is important because the probability of obtaining a nomination from a voter who has not yet read a script should, in principle, be zero. As explained in the previous section, potential voters (production-company and studio executives) typically read scripts when they are shopped around. Thus, the visibility of a script among potential voters is likely to vary widely depending on how widely the script is circulated and how likely it is to be read if received.

With this simple setup, we obtain two sets of predictions (see appendix B for the proofs). The first set concern the determinants of Black List nominations, and the results are straightforward:

Prediction 1 (Expected number of nominations)**.** *Other things being equal, the number of nominations that a script is expected to receive (1) increases with its prior quality, π; and (2) increases with its visibility, n.*

The second set of predictions concerns the posterior quality of a script, updated upon observing the realized number of nominations. Assume that people follow Bayes' rule, we have the following:

Prediction 2 (Belief updates of script quality). *Other things being equal, the posterior probability that the script is of high quality (1) increases with the number of nominations, m; (2) increases with the script's prior quality, π; and (3) decreases with the script's visibility among the voters, n.*

In summary, Prediction 2.1 shows that for scripts that are ex-ante homogeneous (i.e., the same prior quality and the same effective number of voters), with informative signals, the posterior quality increases with the number of nominations. In other words, one can simply use the Black List rankings to infer the relative merits of two different scripts that are, otherwise, similar. Predictions 2.2 and 2.3 compare scripts that differ in their prior quality or in their visibility among survey respondents. They show that, other things being equal, an inferior posterior quality would be inferred for the script that is believed to have been read by a wider set of survey respondents or to have a lower prior quality.

4 Data and Methods

4.1 Two samples

Our data come primarily from two sources. Our first sample—*Sample BL*—is based on the annual publications of the Black List and includes the 701 listed scripts between 2007 and 2014. We exclude the first two years (2005 and 2006) due to the lack of buyer-side information (the production companies and/or movie studios that have either purchased these scripts or commissioned them), which we use to determine whether or not a script was sold prior to publication and to control for the buyer's characteristics. We also cut off the data after 2014 to provide a sufficient window for scripts to be produced before our outcome data collection date (October 2017).

In order to compare scripts on and off the list, we use Done Deal Pro (DDP), an internet database that tracks script transactions on a daily basis, to create *Sample DDP*, our second sample. DDP is recommended by various industry organizations (e.g., the Writers Guild of America) and professionals as one of the top information sources on movie projects under development. The database by no means captures all projects developed by the industry, but it does seem to be quite comprehensive.[13] We manually match the records based on their titles and writers to the Black

[13]To confirm this, we performed a manual check of all movies released by the major studios in 2017 and were able to match about 80 percent back to records in DDP.

List publications to determine whether or not a script has ever been Black-Listed.

We exclude DDP records in which the script idea originated with the studio and a writer was hired to simply draft the screenplay. For this type of project, we cannot systematically observe whether a script is later completed and whether the completed script is good enough to advance to the next stage (i.e., the stage for the majority of the Black-Listed scripts). We, thus, exclude these projects to avoid confounding the effect of the Black List and different degrees of uncertainty.[14] Additionally, we exclude rewrite contracts because the listed screenwriters are typically not the original writers, as well as turnarounds (i.e., projects that were originally developed by a different studio and, often, a while ago). After these exclusions, Sample DDP includes 1,587 observations between 2007 and 2014, 13 percent of which were Black-Listed and 87 percent of which were not.[15]

A useful way to think about the two samples is that they capture different segments of the nomination distribution. Sample BL provides a detailed picture of the top end of the distribution (above the cut-off threshold). The key variation is the number of nominations received by a script (and, hence, its ranking on the list). Sample DDP, by also including unlisted scripts, extends to a lower part of the distribution, albeit still truncated from the left, because it includes only scripts that are good enough to be sold. Our analysis of this latter sample focuses on the differences between listed and unlisted scripts, and the results complement those obtained from Sample BL; together, they provide a more holistic interpretation of the effects and mechanisms.

Finally, we collect additional data from two complementary sources: (1) the Internet Movie Database (IMDb), which provides industry experience for writers and producers; and (2) The Numbers, which supplies outcome information, including whether scripts have been theatrically released and the box-office revenues generated.

4.2 Variables and summary statistics

Table 1a provides summary statistics for Black-Listed scripts in Sample BL. The number of survey respondents varies from 150 in 2007 to 300 in 2011, with an average of 250 per year. Black-Listed scripts, on average, receive 11.62 nominations, with the maximum being 133. Because the number

[14] Note that these excluded scripts tend to be based on existing property rights (e.g., the Avengers series). Consequently, Sample DDP is less representative of scripts that are not based on original ideas and movies that require the highest budgets, but this makes the sample more comparable to scripts on the Black List.

[15] See Appendix C.2 for more details on the construction of this sample.

of nominations varies substantially across years due to different numbers of respondents, we create a discrete rank group variable in the analysis to indicate where the script is placed on the list in a survey year: 4 = top 5 (5.71 percent of the scripts); 3 = top 6 to top 20 (19.40 percent); 2 = top 21 to just above the bottom (53.35 percent); and 1 = bottom scripts that receive the cut-off number of nominations used in a given year (21.54 percent).

Writer Major Credits is the number of writing credits the writer obtained in the previous ten years for movies that were distributed by the top 30 movie studios.[16] The ten-year restriction captures writers' industry experience and current status but avoids simply measuring industry tenure. Restricting the count to movies distributed by the top studios prevents inflating the writer's experience with small-budget independent movies. Whenever there is more than one writer, the maximum of the writers' credits is taken as that of the team.[17] Eighty-one percent of the listed scripts are from writers with zero major credits (defined as less-experienced writers in our analysis); for the remaining 19 percent (experienced writers), the average number of credits is 1.93.

Top-10 Agency indicates whether the writer's agency ranks in the top ten in terms of script-transaction market shares captured by DDP;[18] and *Experienced agent* indicates that the specific agent representing the writer has more than 15 sales prior to this particular one. The vast majority (91 percent) of writers of Black-listed scripts are affiliated with a top-10 agency, and 34.38 percent are represented by an experienced agent.

Sold prior to publication equals one if the script was associated with a production company or a movie studio at the time the list was published. 79 percent of the listed scripts are sold prior to publication. Similar to the measure of writer experience, *Producer Major Credits* is the number of producing credits the producer (which we use to mean the production company in this paper) obtained in the previous ten years for movies that were distributed by the top 30 movie studios.[19]

[16]Top-30 studios are defined in terms of their market shares of the total box-office revenues generated by movies they distributed between 2000 and 2017.

[17]For both samples, we censor the value at five (writers with more than five major credits are associated with about one percent of both samples).

[18]The top ten agencies are Creative Artists Agency, Endeavor Agency, International Creative Management, United Talent Agency, William Morris Agency, WME (after merging William Morris and Endeavor), Paradigm, Gersh, Agency for the Performing Arts, and Verve Talent and Literary Agency.

[19]The Black List publications and the DDP database always list the production companies and often list the individual producers (partners and high-level executives of the company) who are in charge of this particular project. When possible, we define the producer's experience as the maximum of individual producers' industry experience. When only a production company is listed and not the individuals, we use the past producing record of the company instead. This variable is censored from the top at 20, and defined as zero for unsold scripts in the regressions.

Movie Studio Buyer equals one if a movie studio is listed among the buyers, and 36.66 percent of the listed scripts have a studio buyer.[20] Projects managed by producers not associated with a movie studio when the List is published may be financed and/or distributed by a studio later.

Theatrically Released is defined as a script achieving positive U.S. box-office revenues. Among Black-Listed scripts, 26.25 percent are theatrically released. The median and the mean U.S. box office for released movies are, respectively, \$24.80m and \$45.33m.

Table 1b describes Sample DDP, in which 13 percent of scripts are Black-Listed. The writers in Sample DDP have, on average, obtained more major writing credits (0.62 versus 0.36) than writers in Sample BL. They are, however, less likely to be affiliated with a top-10 agency (55 versus 91 percent) or an experienced agent (17 versus 34 percent). Thirty-six percent of the scripts have a movie studio among the buyers, and 59 percent of the scripts are based on original content (versus an adaptation of an existing property).[21] For three percent of the observations, the record indicates that the script is purchased through a bidding process.[22] Some talent (a director, actor, or actress) is attached to 56 percent of the projects at the time of the record, possibly due to the agency's packaging efforts. The largest genres are drama (32 percent); comedy (17 percent); and thriller or suspense (17 percent).[23] Fifteen percent of the scripts are theatrically released, which is substantially lower than the release rate in Sample BL. Conditioned on release, the median and mean U.S. box-office revenues are \$14.60m and \$47.96m, respectively.

4.3 Descriptive results

In this section, we first present a number of stylized facts based on the raw data on (i) how the Black List nominations (that is, whether or not a script is listed and its ranking) correlate with the release outcomes; and (ii) how the nominations and outcomes vary by writer experience.[24] We then interpret these descriptive patterns in light of the conceptual model developed in section 3.

[20]Movie studios include the six major studios—Warner Bros., 21st Century Fox, Paramount, Sony (Columbia), Universal, and Disney—and mini majors such as Miramax, Focus Features, New Line, and Lionsgate. This variable is defined as zero for unsold scripts.

[21]Among the adaptations, 48 percent are based on a book; 14.8 percent franchise (such as sequels); 9.5 percent short stories; 5.7 percent comic books; and the remaining categories include musicals and plays and fairy tales.

[22]If the keyword 'bid'—as in bidding war or competitive bidding—is contained in the notes associated with a transaction in the DDP database.

[23]The other genres are science fiction (7 percent), adventure (6 percent), and romantic comedy (6 percent). Ee group the remaining genres into "other."

[24]Note that the following facts are not about the population of scripts but are conditional on those that are sufficiently good, as Sample DDP captures only sold scripts and Sample BL only Black-Listed ones.

Stylized fact 1: scripts from less-experienced writers are more likely to be listed than those from experienced ones (and ranked higher if listed). Figure 1a shows that the probability of being Black-Listed for less-experienced writers (with zero major writing credits in the previous ten years) is 15.1 percent, which is twice as much as that for experienced writers (p-value < 0.001). Figure 1b shows that, if listed, the likelihood of being among the top 20 is 27.6 percent for less-experienced writers and 14.4 percent for experienced ones, a difference of 92 percent. This result also holds for the probability of being ranked among the top 5 (6.7 vs. 1.5 percent, and p-value is 0.02).

Stylized fact 2: Black List nominations are positively related to release probabilities. Figure 2a shows that Black-Listed scripts are 14 percentage points more likely to be released than unlisted scripts (a difference of 100 percent, and p-value < 0.001). Excluding top-5 listed scripts, the difference is smaller at 6.3 percentage points, or 42 percent (p-value is 0.02). Figure 2b shows a monotonic increasing relationship between how high the script is placed on the list and the release probability, even though only the difference between top-five scripts (at 47.5 percent) and lower-ranked ones (25.0 percent, on average) is statistically significant.

Stylized fact 3: the box office data for released movies are consistent with the Black List being effective at weeding out low-potential ideas. Figure 3a shows that, relative to unlisted (released) scripts, the density distribution of log(U.S. box office) for Black-Listed (released) scripts is more concentrated on the higher end of the distribution and is associated with a lower variance. Testing for the equality of distributions yields a p-value of 0.001, and the mean of log(U.S. box office) is significantly higher for Black-Listed scripts than for unlisted scripts (p-value = 0.003). Figure 3b shows that within Black-Listed movies, however, the density distribution of log(U.S. box office) is statistically similar across the rank groups.

Stylized fact 4: scripts by less-experienced writers are substantially less likely to be released than scripts from experienced ones, and being listed (or a higher place on the list) does not help close the gap. Figure 4b shows that for any given Black List rank group, the gap is 29 to 55 percentage points, amounting to differences of more than 100 percent. The gap does not decrease for higher-ranked scripts; and if anything, it seems to be increasing. We later confirm this large discrepancy in regressions that compare scripts receiving exactly the same number of nominations in the same year. This pattern also extends to the comparison between scripts on vs. off the list (Figure 4a), even though this comparison is not as clean, because each group—listed and unlisted—pools together scripts with a wide range of different nominations.

In the model described in section 3, among ex-ante similar scripts, those with more nominations are associated with a higher quality (prediction 2.1). The data show that the difference in the release outcomes is substantial and significant when comparing Black-Listed to unlisted scripts but not across the Black List rank groups except for the probability of release for top-5 scripts (stylized facts 2 and 3). These results suggest that voters may be better at discerning bad scripts from good ones than telling good scripts apart from each other. Note that the descriptive relationships are only correlational; and it is possible that the Black List simply lists high-quality scripts that are more likely to be produced regardless. In the next subsection, we discuss the challenges to obtain causal evidence and the methods we use to address them.

Our model can also rationalize the remaining stylized facts described above if scripts from less-experienced writers are, on average, more visible among potential voters (i.e., a higher n). According to predictions 1.2 and 2.3, this 'visibility' mechanism can simultaneously explain: (i) the negative correlation between writer experience and Black List nominations (stylized fact 1); and (ii) a lower release likelihood for less-experienced writers, even when their scripts have received the same number of nominations as those by experienced writers (stylized fact 4), because an inferior posterior quality is inferred due to a wider visibility. This explanation is plausible considering the likely difference in the sales scope for writers with different experience levels. Given their track record, experienced writers may not need to market their scripts widely, as information about their scripts' quality and about which potential buyers are likely to be a good match is more complete to all involved.[25] In contrast, less-experienced writers may need to market their scripts to a much wider audience in order to find an interested buyer or the best match. Simply put, if an experienced writer markets a script to four production companies, the script can receive at most four nominations. In contrast, a less-experienced writer marketing a script to twenty production companies has the chance of obtaining more than four nominations, especially if the script is very good.

An alternative mechanism offered by the model—scripts from less-experienced writers are more likely to be exceptionally good (i.e., a higher π)—can also explain stylized fact 1 (prediction 1.1).

[25]Industry accounts also suggest that sellers have genuine incentives to limit sales scope due to concerns over information leakage and time costs.

This assumption is possible because, even though the average quality of scripts from novice writers may be lower, the variance could be greater. This explanation cannot explain stylized fact 4, however, as prediction 2.2 claims that for scripts with the same nominations we should observe a higher, not a lower, release likelihood for less-experienced writers. Of course, there may be factors outside the model that could explain why Black List nominations for less-experienced writers—despite the belief that they are of better quality—translate less well to eventual success than for experienced ones. We discuss these alternative explanations in section 5.3.1.

4.4 Empirical specification

In this section, we focus on the methods we use to address the challenges for identifying the effects of the Black List; and we report the results based on these methods in section 5.2. The other regression results in section 5 are based on straightforward OLS regressions.

Consider the following OLS regression for script i that is managed by buyer j in year t:

$$\text{Release}_i = \delta \text{Black-Listed}_i + \beta X_i + c_{jt} + \epsilon_i, \tag{1}$$

where Release_i indicates whether or not a script is theatrically released; Black-Listed_i indicates whether or not a script is listed; X_i are observable factors associated with the script; and c_{jt} is the buyer×year fixed effects. Notice that separate buyer and year fixed effects drop out of the estimation because equation (1) uses a more stringent set of fixed effects. We define the buyer as the movie studio if one is among the buyers at the time of the DDP record, or as the production company, if no movie studio is involved.

According to prediction 2.1, being listed should have a positive effect, because for ex-ante similar scripts, decision makers are likely to adjust upward their beliefs about the quality of listed scripts (relative to unlisted ones). Relevant decisions makers include the managing producer, who may adjust the prioritization over projects in her portfolio, as well as talent (directors and actors) and studio executives, who may change their scheduling priorities or their resource allocation decisions across competing projects. It is, thus, important to note that our estimate of the effect of the Black List inevitably includes the potential crowding out effect on unlisted scripts due to resources allocated away from them.

Additionally, it is important to note that even if the Black List does not change how one thinks about a script's underlying quality (e.g., if it is common knowledge that voters' signals are uninformative and, hence, nominated scripts are not necessarily better than ex-ante similar but unlisted scripts), being listed may still have a positive causal effect. This is because, by providing a focal point that attracts people's attention, the Black List may help coordinate industry resources and set priorities. In section 5.2.1, we draw on available data to shed light on how important this pure 'coordination effect' might be.

The fact that we are unlikely to observe all the information available to industry players creates two types of empirical challenges to $\hat{\delta}$ being an unbiased estimate of the treatment effect of being listed. First, there may be factors unobservable to the researchers but observable to decision makers that are both systematically correlated with the likelihood of being Black-Listed and the likelihood of release. For example, scripts with a higher unobservable (ex-ante) quality are more likely to be Black-Listed, which may lead to an over-estimation of the effect of being listed because the belief about the listed projects is already quite high, even without seeing them listed.

Apart from including as many controls as we can, we exploit the idea that projects purchased by the same buyer in the same year (defined as a portfolio, c_{jt})—once conditional on observable characteristics—are likely to be similar to each other in their expected value *prior to* the List's publication. Admittedly, even projects set up at the same studio in the same year (presumably above a uniform quality threshold) may be different in ways observable to relevant decision makers but unobservable to us. That being said, this assumption is not unreasonable given the reputed unpredictability of early-stage movie projects. We also employ additional methods that help restrict the ex-ante qualities of the scripts to be as similar as possible, including the coarsened exact matching method and using only scripts that are marginally listed (e.g., just making the cut).

Second, if a listed script has already been widely talked about in the industry, decision makers may already know *what others think about a script* before the publication of the Black List; in other words, the amount of new information provided by the Black List can be quite small. As explained in detail later in section 5.2, we collected additional data on the media coverage of a given script before the List's publication. We use this variable both as a control and to derive differential effects of being Black-Listed, depending on the informational environment surrounding a given script.

5 Regression Results

In this section, we first present the regression results on the determinants of Black List nominations. Then, in section 5.2, we report the results on the average effects of being listed on release outcomes. We provide evidence based on a number of methods that favors a causal interpretation of the estimated differences. Finally, in section 5.3, we present results on how the relationship between Black List nominations and release outcomes varies by writer experience and discuss potential interpretations of these results in conjunction with the results presented in section 5.1 on factors that are important for determining the Black List nominations.

5.1 Determinants of Black List nominations

The OLS regression results in table 2 first confirm the negative correlation between writer experience and Black List nominations exhibited in the raw data (stylized fact 1 in figure 1). Columns 1 and 2 show that, among listed scripts, those from experienced writers are 5.7 percentage points less likely to be among the top five and 12.9 percentage points less likely to be among the top 20; column 3 further confirms this negative correlation using the rank-group variable. In addition to the writer's experience, the regressions also include the characteristics of the writer's agent, the buyer, and the survey-year fixed effects. Columns 4 and 5 in table 2, using Sample DDP, show that sold scripts from experienced writers are about ten percentage points less likely to be Black-Listed. The regressions include additional variables available for this sample, including whether an idea is original (versus an adaptation), and two groups of dummies indicating the genre and the source of the underlying content, such as a book or a short story if the idea is not original. The number of observations drops significantly in column 5 because it controls for buyer fixed effects and excludes scripts by buyers with only one project in the sample.[26]

Apart from writer experience, we highlight two other factors that are also significantly correlated with Black List nominations. First, the results in table 1 show that the experience of the writer's agent and the size of the agency are, overall, positively correlated with Black List nominations. In the context of our framework, this could be because large agencies and more-experienced agents are associated with both a higher prior quality of the script (π) and/or a greater visibility among the

[26] Appendix table A1 shows that this negative correlation is robust to alternative measures of writer experience and to including the writer's other experience, such as major directing or acting experience.

voters (n). The positive association with prior quality could be either due to positive assortative matching (better writers are more likely to be represented by larger agencies and more-established agents) or because these agents have a greater reputation at stake and tend to do a better job of screening out bad ideas. The positive association with visibility is also quite intuitive, given that larger intermediaries have more connections, greater resources, and a better reputation both to promote a script more widely and to convince relevant parties to read it.

In table 3a, we investigate the relationship between writer experience and Black List nominations in separate subsamples, each fixing the size of the agency and/or the experience of the agent. The idea is to restrict each subsample to a group of writers who are more likely to be similar in their overall quality (e.g., within writers good enough to be represented by a top agent at a large agency). Interestingly, the results show that the significant negative correlation between writer experience and Black List nominations holds only when the writer is associated with a large agency, and the contrast is largest in magnitude for writers who are also associated with an experienced agent.[27] Considering the baseline likelihood of being Black-Listed for less-experienced writers presented at the bottom of each panel, those who receive the most nominations tend to be less-experienced writers associated with very experienced agents at top agencies. Though subject to other interpretations, this result is consistent with the idea that more nominations may require scripts to be shopped widely—which critically depends on the resources and connections of the writer's agents.

The second factor we highlight relates to the different sales methods of scripts. Columns 4 and 5 of table 2 show that scripts identified as having gone through some type of bidding process are 20 percentage points more likely to be Black-Listed. This could be because the bidding method is more likely to be used for scripts with a higher quality; additionally, this sales method may result in a high visibility among potential voters as the script is often sent to many potential buyers at the same time. Interestingly, bidding is not differentially associated with experienced versus less-experienced writers (2.4 vs. 2.6 percent, p-value = 0.76). However, it is significantly more likely to be used by a top-10 agency (4.1 vs. 0.7 percent, and p-value $<$ 0.001); and within top-10 agencies, more by the most experienced agents than by less-experienced agents (6.9 vs. 2.9 percent, and

[27] Almost all experienced agents work for one of the ten largest agencies, so we do not further split the subsample not associated with a top-10 agency by agent experience.

p-value is 0.007).

Table 3b shows that the negative correlation between writer experience and Black List nominations holds for sales both with and without a bidding process. In fact, for sales identified with a bidding process, scripts by less-experienced writers are 37.9 percentage points more likely to be listed than experienced writers, which is a very large difference (note, however, this result is based on a very small sample). Even though it is possible that scripts from less-experienced writers chosen for this sales method are substantially better, it seems more plausible that the large discrepancy is because it is necessary to send scripts from less-experienced writers to many more potential buyers in order to generate sufficient interest and to find a good match.

5.2 Average effects of being Black-Listed on release outcomes

As explained in section 4.4, our main identification method assumes that projects by the same buyer in the same year (a portfolio)—once conditional on observable characteristics—are likely to be similar to each other in their expected value prior to the publication of the Black List. Column 1 in table 4 reports the regression results based on this method (i.e., equation 1); the dependent variable indicates whether or not a script is theatrically released, and the regression includes a complete set of buyer-year dummies. Even though Sample DDP includes 1,022 unique portfolios, 81 percent of them contain only one project and are omitted from the regression; that is, the (within-portfolio) identification variation comes from relatively large buyers. The results show that Black-Listed scripts are 16.2 percentage points more likely to be released than unlisted scripts, representing a difference of about 100 percent (the baseline release probability of unlisted scripts used in the regression is 0.153). As discussed previously, this estimated difference is inclusive of the potential crowding-out effect on unlisted projects. The assumption of a largely fixed budget is especially relevant given that we are comparing projects within a given portfolio.

The next set of results exclude listed scripts that are highly ranked: column 2 of table 4 excludes the top-20 listed scripts; and column 3 uses only scripts that received the threshold number of nominations in a given year. The idea is to zoom in on the subset of listed scripts that are most likely to have an (ex-ante) quality that is similar to unlisted scripts in a buyer's portfolio, as the latter may have received nominations not far below the threshold.[28] Excluding the top-20 listed

[28]Ideally, we want to observe scripts receiving a number of nominations just below the cut-off threshold and use a

scripts, the effect becomes slightly smaller, at 16 percentage points (p-value = 0.05). When using only scripts that just make the cut, the point estimate, at 20.6 percentage points, is economically large but statistically insignificant; this may be the case because there are only 13 listed scripts at the threshold level when controlling for portfolio fixed effects.

As discussed in section 4.4, another empirical challenge is that decision makers may already know *what others think about a script* before the publication of the Black List, leading to a minimal amount of new information contained in the publication. To address this issue, we used Google search to find the number of articles about particular scripts (listed and unlisted) published in the top-two trade magazines in Hollywood—Variety and Hollywood Reporter—that are dated about two years before the exact publication date of the Black List in the year the script was sold.[29]

Column 4 of table 4 replicates the regression in column 2 (i.e., excluding top-20 scripts and controlling for portfolio fixed effects) but includes the number of articles as a control variable. The idea is to compare listed and unlisted scripts with the same amount of industry exposure, and the coefficient of Black-Listed is only slightly smaller and remains highly significant. Column 5 includes the interaction term between Black-Listed and a dummy variable indicating scripts reported by the two magazines before the list's publication. The results show that the effect of being listed is driven exclusively by scripts for which we found zero articles.[30] These results are consistent with the idea that the amount of new information—i.e., what other people think about a script—contained in the Black List is likely to be greater for scripts that have not received much media attention than for scripts that have.

In table 5, we report the results based on an alternative method—coarsened exact matching (CEM, Iacus et al. (2012)). The CEM method generates a matched sample of listed and unlisted scripts that are similar to each other in a number of pre-specified dimensions.[31] For each specification, we include buyer and year fixed effects. The results are consistent with the results reported

regression-discontinuity approach. These data have, unfortunately, not yet been made available.

[29] We found zero articles for 79 percent of the 733 observations used in column 1; one article for 16 percent; and two or three articles for the remaining five percent. The basic comparisons seem sensible: the average number of articles we found is 0.38 for listed scripts and 0.23 for unlisted scripts (p-value is 0.006); and for unlisted scripts, the release probability is higher for those with some reporting than for those with none, though the difference is not statistically significant at the conventional level (18.2 vs. 13.7 percent, p-value = 0.161).

[30] For scripts that were reported by the magazines, the difference between listed and unlisted scripts is the sum of the coefficients of Black Listed and Black-Listed×Reporting, which is $0.209 - 0.213 = -.004$ (p-value = 0.96).

[31] Variables used to create the matched sample are writer major credits, top 10 agency, top agent, producer major credits, original, talent attached, reporting (# of links), and studio-buyer dummy.

above using portfolio fixed effects.

For movies that are released, we also have information on their box office performance. Column 1 in table 6 shows that, consistent with the raw data reported in figure 3a, released movies that have been Black-Listed generate a significantly higher box office revenue in the U.S. than unlisted ones. Furthermore, TheNumbers.com provides production-budget estimates for about two thirds of the released movies.[32] Column 2 shows that, for this subsample of released movies, the result in column 1 is robust to controlling for the budget information. Consequently, column 3 shows that listed movies generate a significantly greater return over investment, which is defined as U.S. box office revenue divided by production budget.

5.2.1 Potential mechanisms for the estimated effect

Results in the previous section show a large difference in the probability of release between listed and unlisted scripts, and we provided evidence that favors a causal interpretation of this difference. In the following, we discuss potential mechanisms that may underlie this positive effect.

In our conceptual model, the aggregation of individual judgment by the Black List reduces uncertainty over the quality of scripts that are ex-ante similar: according to prediction 2.1, one should adjust their beliefs about the quality of listed scripts upward relative to unlisted ones. Put differently, the list is wise, at least in terms of telling good ideas apart from poor ones (i.e., those that are off the list).[33] As illustrated by the anecdotal evidence cited in section 2.2, this quality signal may help convince studio executives about a project's merit and reduce the search and evaluation costs for talent.

Alternatively, as discussed in section 4.4, being listed could have a positive effect by providing a focal point that attracts people's attention, thereby aiding the coordination of resources. Because it is impossible to observe or measure a script's quality, it is hard to cleanly isolate this *pure coordination effect* (that is, facilitating coordination by simply picking the winner and not by differentiating qualities). The result on box office performance reported in the previous section—that is,

[32]The exact sources of these budget estimates are unclear. According to TheNumbers.com, "Budget numbers for movies can be both difficult to find and unreliable. The data we have is, to the best of our knowledge, accurate but there are gaps and disputed figures." https://www.the-numbers.com/movie/budgets/all.

[33]As discussed in section 4.3, the aggregated judgment by the Black List may only be able to distinguish quality up to a certain points; i.e., while it can seperate good ideas from poor ones, it may not be able to distinguish good from great ideas.

Black-Listed movies generate a significantly higher box office revenue (and, hence, a higher return to investment) than unlisted movies with the same budget—is not consistent with a pure coordination effect. Additional evidence, discussed in detail in the following paragraph, also suggests that the pure coordination effect, though likely to be present, does not seem to be a key driving force of the data.

First, column 1 in table 6 shows that, on average, the production budget of listed (released) movies is not greater than released movies that are not listed; if anything, the average budget is slightly lower, even though the difference is not statistically significant (p-value is 0.249). This result shows that conditional on movies that are eventually released, listed movies do not attract more resources than unlisted ones. Second, column 2 in table 6 shows that the average time taken from the script's sale to release is also similar between listed and unlisted movies, which does not support the idea that that Black List nominations influence the outcomes mainly by coordinating talent. Third, the last column of table 6 shows that the effect of being list is not greater for projects to which no talent has yet committed at the time of the DDP records—which, intuitively, could benefit more from a coordination effect of the Black List—than projects with talent already attached initially.

5.3 Release outcomes by writer experience

Table 7 uses Sample DDP and the same methods—portfolio fixed effects and coarsened exact matching—that are used in section 5.2. The regressions include an interaction term between Black-Listed and an indicator for experienced writers. The results are consistent across the specifications: (i) the coefficient of Black-Listed is large and statistically significant, suggesting that being listed does have a large and positive effect for less-experienced writers; and (ii) the coefficient of the interaction term, though not significant, is positive and economically large. The second result shows that being listed, though a selective event, does not help close the gap for less-experienced writers, as, otherwise, we would have observed a negative coefficient of the interaction term.

The next set of results uses Sample BL and shows that this pattern is also present at the top end of the nomination distribution. Column 2 in table 8 first confirms that there is a large discrepancy in the release probabilities between experienced and less-experienced writers for listed scripts. The difference is 18 percentage points, representing a 100-percent difference (the baseline release probability for scripts from less-experienced writers is 17.7 percent). Note that this regression controls for nomination×survey year fixed effects. The comparison is, therefore, among scripts receiving exactly the same number of nominations in the same survey year.

The remaining columns of table 8 show that this discrepancy does not shrink as the number of nominations increases. These regressions include the interaction terms between writer experience and dummies indicating the Black List rank group (we group all top-20 scripts together because it is rare to have multiple scripts in the top-five slot with the same nominations), and the omitted baseline is the bottom-listed group. The results show that relative to bottom-listed scripts, the gap in the release probabilities from writer experience does not become smaller as the rank group increases. In fact, the gap becomes economically larger for top-20 scripts, though the difference is not statistically significant.

5.3.1 Potential explanations for results by writer experience

Taking the results on writer experience at both stages together presents an interesting and important contrast. On the one hand, section 5.1 shows that the Black List seems quite effective at highlighting scripts from relatively novice writers; they are both more likely listed and to rank higher if listed. On the other hand, section 5.3 shows that when it comes to the release outcomes, even though being listed has a large and positive effect within less-experience writers, the large discrepancy relative to experienced writers remains even for the top-ranked scripts.

As discussed in Section 4.3, our model can rationalize the above results under the premise that scripts from less-experienced writers are, on average, more visible among potential survey respondents (likely due to their need to shop their scripts around more than experienced writers). Put differently, on average, the number of nominations obtained by less-experienced writers appears to be inflated relative to experienced writers in terms of the inherent quality of a script. This explanation highlights the importance of a script's visibility among potential voters as a systematic determinant of the data-generating process of the Black List. Other results in section 5.1 on the roles of large agencies, experienced agents, and different sales methods are also consistent with this explanation. Even though decision makers appear to have accounted for this systematic bias by penalizing ideas from less-experienced writers (i.e., the lower release likelihood of their scripts compared to those from experienced writers that receive the same nominations), it is unclear whether such adjustments are optimal given the limited information available on the actual visibility of a particular script and the characteristics of its effective voters.

Alternatively, as noted above, our model suggests one other explanation for the negative correlation between writer experience and Black List nominations: that scripts from less-experienced writers are more likely to be exceptionally good (e.g., if the quality distribution of their scripts has a greater variance).[34] This explanation requires additional factors to account for the large discrepancy in release probabilities, however, considering that it implies a posterior belief of higher script quality for scripts from less-experienced writers. We discuss two such factors. The first is risk aversion: given that having a script is only the start of a complex and uncertain process, scripts from more-experienced writers are safer bets. If risk aversion is important, conditional on released movies, we should observe a better box office performance for movies from less-experienced writers than experienced ones (as the decision threshold for the former is likely to be stricter to compensate for the greater risk). The data, however, do not support this prediction. Table 9a shows no significant difference between released movies from experienced writers and those from less-experienced ones in terms of production budget, box office performance, or return over investment.

The second possible factor arises from the fact that the Black List surveys happen only once a year and the majority of the scripts (80 percent) are already associated with a buyer by the time the list is published. Given that rematching based on the new information revealed by the Black List would be costly, it is possible that scripts from less-experienced writers are often 'stuck' with less-experienced producers or producers without easy access to financing (e.g., not already associated with a major studio).[35] Factors corresponding to producer characteristics, according to unreported regression coefficients in table 8, are all significantly correlated with the likelihood of release. In table 9b, we report subsample regression results, each including scripts that are relatively homogeneous in these pre-determined dimensions. Furthermore, all the subsamples are also conditional on listed scripts that are sold prior to publication so as to avoid confounding the

[34] Appendix section D.1 discusses two additional factors outside our conceptual model that might also help explain the negative correlation between writer experience and nominations:(i) experienced writers' scripts progress into the production stage faster; and (ii) voters have preferences for scripts from inexperienced writers that are unrelated to a script's commercial appeal. The negative correlation remains significant after controlling for these alternative explanations.

[35] Indeed, listed scripts from less-experienced writers are significantly less likely than those from experienced writers to be sold prior to publication (76 versus 92 percent, and p-value < 0.001); and conditional on being sold, they are also less likely to be managed by producers associated with a movie studio (39 versus 73 percent, and p-value $<$ 0.001), though the managing producers' experience does not differ by writer experience in Sample BL.

additional uncertainty associated with having to first sell the scripts afterwards.[36] The results show that, even after removing some of the pre-determined differences in relevant decision makers' ability to utilize the Black List nominations, the discrepancy by writer experience in terms of the eventual release probability remains large and statistically significant.

6 Conclusion

This paper studies the aggregation of expert judgment in the selection and development of early-stage ideas in the context of the movie industry, where investment costs are high and consumer taste is notoriously unpredictable. To do so, we use The Black List, an annual survey that ranks scripts not yet produced based on anonymous nominations from 250-300 films executives. We find that listed scripts are over one hundred percent more likely to be released than unlisted scripts and that they generate a significantly greater return on investment. We also find that scripts from less-experienced writers are more likely to be listed and to rank higher if listed. Yet, while being listed has a large and positive effect on the probability of release for both experienced and less-experienced writers, the gap between the two remains large (even for the top-ranked scripts). These two results by writer experience are consistent with the idea that the number of scripts listed by less-experienced writers is inflated relative to experienced writers, as the former may need to shop their scripts around more widely, and therefore, have greater visibility among potential voters.

Ultimately, our results suggest that aggregating the individual judgment of experts can succeed in reducing uncertainty over the quality of ideas. This ability is important, particularly where there are few concrete signals of quality, such as in markets where idea's appeals to consumers are largely subjective and/or it is costly to experiment before committing to large investments. While the movie industry is one such setting, many other contexts experience these challenges, including other creative industries, consumer technology start-ups, and product development teams, to name a few. Though not definitive, the fact that the Black List can influence main-stream, large-budget production seems to have resulted from the credibility of its voters and its large scale. These characteristics may, in turn, be partly due to the Black List's light-touch approach that keeps the participation cost low for busy executives.

[36] Appendix section D.2 takes a closer look at the likely effects of the Black List for 20 percent of the listed scripts that are not sold prior to the list's publication.

Our results also reveal some of the limitations of this approach In particular, the visibility of an idea among potential evaluators appears to be a systematic determinant of the number of nominations. This visibility may be influenced by factors or frictions that are unrelated to an idea's inherent quality, and decision makers may not be able to make optimal adjustments given the limited information available about the actual visibility of a particular idea and the characteristics of its effective evaluators.[37] Even though such biases may be mitigated by improving the structure or transparency of the process, one needs to be mindful that such modifications may also be costly and subject to trade-offs implied by the necessity to keep the participation cost low.

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.

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

Astebro, T. and D. J. Koehler (2007). Calibration accuracy of a judgmental process that predicts the commercial success of new product ideas. *Journal of Behavioral Decision Making 20*, 381–403.

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

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.

Catalini, C., C. Foster, and R. Nanda (2018). Machine intelligence vs. human judgement in new venture finance.

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

Goldman, W. (1983). *Adventures in the Screen Trade: A Personal View of Hollywood and Screenwriting*. Warner Books.

Hall, B. and J. Lerner (2010). The financing of r&d and innovation. In B. Hall and N. Rosenberg (Eds.), *Elsevier Handbook of Economics of Innovation*, pp. 609639. Elsevier.

Hellmann, T. and M. Puri (2002). Venture capital and the professionalizationof start-up firms: Empirical evidence. *Journal of Finance 57*(1), 169–197.

[37]This problem is not unique to the Black List. An example in our own profession is that academic conferences often provide awards based on votes for the the best paper presented without allowing for the fact that some time slots have much greater visibility than others (e.g., spots just after lunch on the main day versus the final afternoon of the conference).

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

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

Luca, M. (2011). Reviews, reputation, and revenue: The case of yelp.com. Working paper.

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.

Scott, E., P. Shu, and R. Lubynsky. Entrepreneurial uncertainty and expert evaluation: An empirical analysis. Management Science (forthcoming).

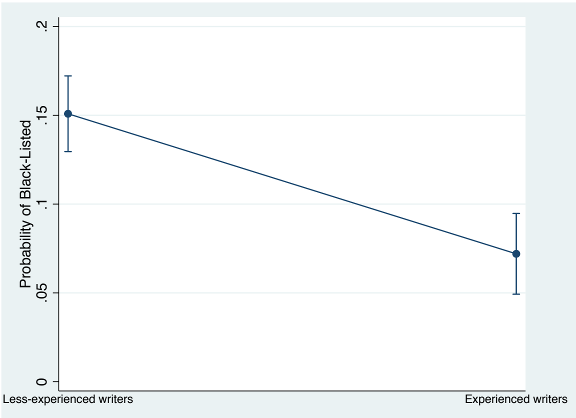
Simcoe, T. S. and D. M. Waguespack (2011). Status, quality, and attention: What's in a (missing) name? *Management Science 57*, 274–290.

Stuart, T. E. and O. Sorensen (2010). Syndication networks and the spatial distribution of venture capital investments. *American Journal of Sociology 106*, 15461588.

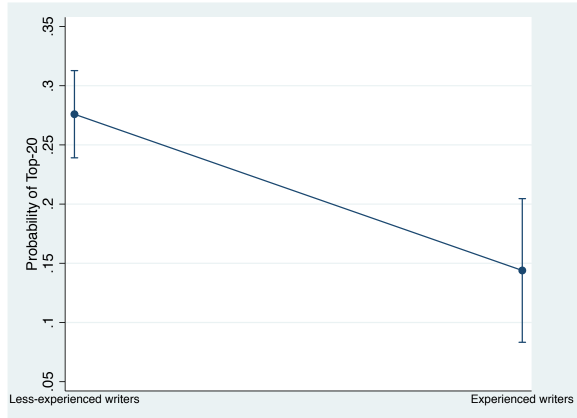
Waldfogel, J. (2017). How digitization has created a golden age of music, movies, books, and television. *Journal of Economic Perspectives 31*(3), 195–214.

Figure 1: Probability of being Black-Listed (or among top-20 if listed) by writer experience

(a) Sample DDP



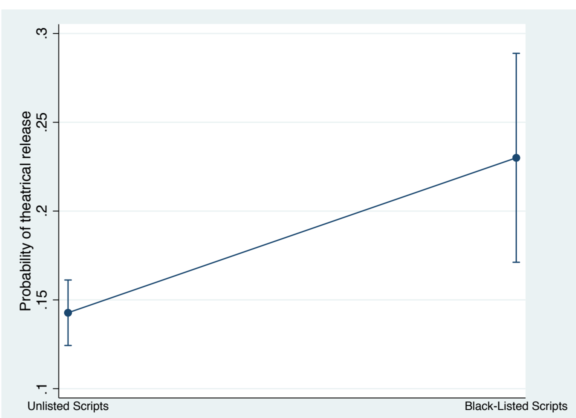
(b) Sample BL



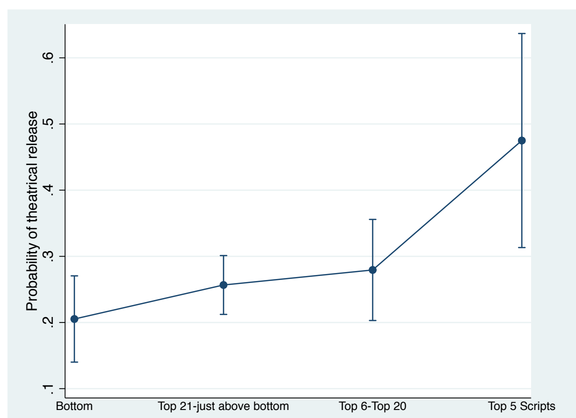
Note: (a) uses Sample DDP and summarizes the likelihood of being Black-Listed for less-experienced writers (i.e., those with zero major writing credits in the previous ten years) and for experienced writers (i.e., those with one or more major writing credits in the previous ten years). (b) uses Sample BL and summarizes the likelihood of ranking among the top 20.

Figure 2: Probability of theatrical release by whether Black-Listed (or by ranking if listed)

(a) Sample DDP



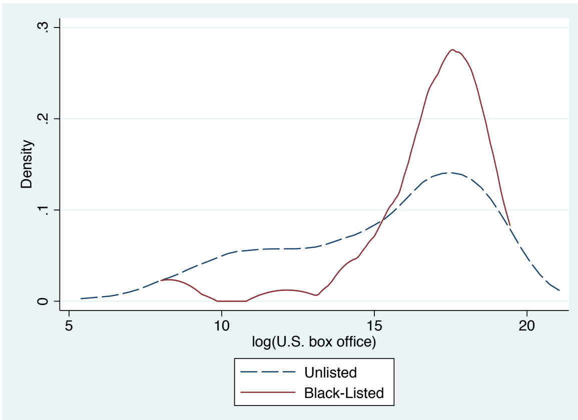
(b) Sample BL



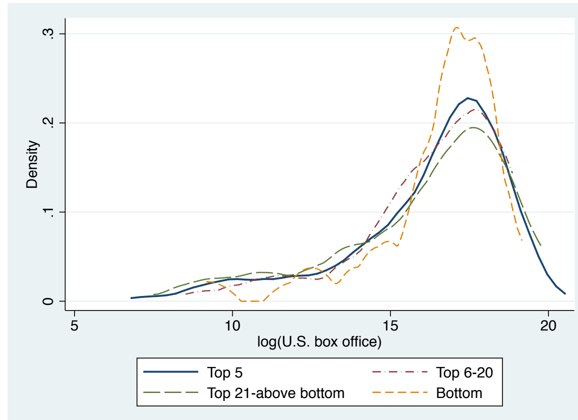
Note: (a) uses Sample DDP and summarizes the probability of theatrical release by whether or not the script is Black-Listed. (b) uses Sample BL and summarizes the probability of theatrical release by Black List rank group: top 5; top 6-20; top 21 to just above the threshold number of nominations used in a survey year; and the bottom (i.e., the threshold number).

Figure 3: Density distributions of log(U.S. box office) for released movies

(a) Sample DDP



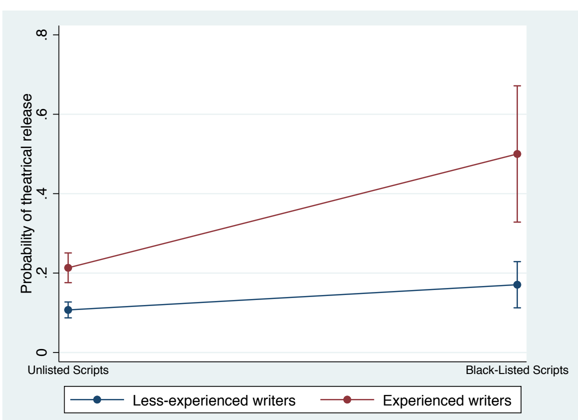
(b) Sample BL



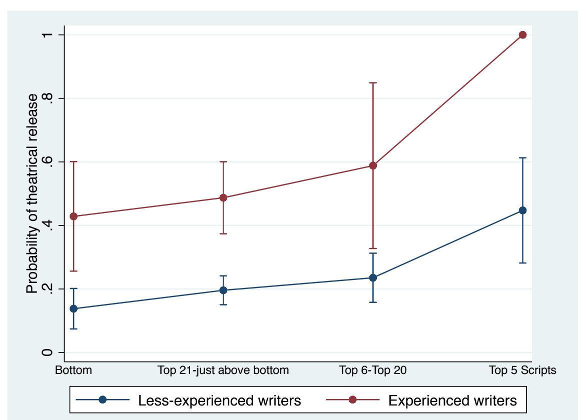
Note: (a) uses Sample DDP and plots the density distributions of log(U.S. box office) for released movies by whether or not the script is Black-Listed. (b) uses Sample BL and plots the density distribution of log(U.S. box office) for released movies by the Black List rank group: top 5; top 6-20; top 21 to just above the threshold number of nominations used in a survey year; and the bottom (i.e., the threshold number).

Figure 4: Release probability and Black List nominations, by writer experience

(a) Sample DDP



(b) Sample BL



Note: (a) uses Sample DDP and summarizes the probability of theatrical release by whether or not the script is Black-Listed for experienced and less-experienced writers, respectively. Figure (b) uses Sample BL and summarizes the probability of theatrical release by Black List rank group. Experienced writers are defined as writers with one or more major writing credits in the previous ten years; and less-experienced writers are defined as writers with zero major writing credits in the previous ten years. Black list rank groups are top 5; top 6-20; top 21 to just above the threshold number of nominations used in a survey year; and the bottom (i.e., the threshold number).

Table 1: Summary statistics

(a) *Sample BL*

Variables	Obs	Mean	Median	S.D.	Min	Max
Executives responded (per year)	701	252.51	250	52.87	150	300
Threshold # of nominations (per year)	701	4.71	5	1.47	2	6
Nominations received	701	11.62	8	11.28	2	133
Rank group	701	2.91	3	0.79	1	4
Writer major credits	701	0.36	0	0.93	0	5
Experienced writer (major credits > 0)	701	0.19	0	0.39	0	1
Top-10 agency	701	0.91	1	0.28	0	1
Experienced agent (prior sales > 15)	701	0.34	0	0.48	0	1
Sold prior to publication	701	0.79	1	0.41	0	1
Producers major credits	494	7.24	6.00	6.27	0	20
Movie studio buyer	701	0.37	0	0.48	0	1
Theatrically released	701	0.26	0	0.44	0	1
US box office ($m)	183	45.33	24.80	61.66	0.00	408.01

(b) *Sample DDP*

Variables	Obs	Mean	Median	S.D.	Min	Max
Black-Listed	1,587	0.13	0	0.33	0	1
Writer major credits	1,587	0.62	0	1.11	0	4
Experienced writer (major credits > 0)	1,587	0.32	0	0.46	0	1
Top-10 agency	1,587	0.55	1	0.50	0	1
Experienced agent (prior sales > 15)	1,587	0.17	0	0.37	0	1
Producer major credits	1,587	5.73	4	6.14	0	20
Movie studio buyer	1,587	0.36	0	0.48	0	1
Original idea	1,587	0.59	1	0.49	0	1
Bidding process	1,587	0.03	0	0.16	0	1
Talent attached	1,587	0.56	1	0.20	0	1
Theatrically released	1,587	0.15	0	0.36	0	1
US box office ($m)	244	47.96	14.60	80.22	0	504.01

Table 2: Determinants of Black List nominations

	Sample BL			Sample DDP	
Dependent Variable	Top 5	Top 20	Rank Group	Black-Listed	Black-Listed
	(1)	(2)	(3)	(4)	(5)
Experienced writer	-0.057**	-0.129***	-0.212***	-0.093***	-0.115***
	(0.019)	(0.024)	(0.060)	(0.018)	(0.027)
Top-10 agency	0.037	0.110**	0.225**	0.085***	0.037
	(0.020)	(0.044)	(0.084)	(0.019)	(0.031)
Experienced agent	-0.002	0.003	0.05	0.119***	0.121***
	(0.021)	(0.030)	(0.044)	(0.024)	(0.040)
Producer major credits	0.001	0.001	0.008	0.003**	0.003
	(0.002)	(0.002)	(0.004)	(0.001)	(0.002)
Sold prior to publication	0.008	0.108*	0.180		
	(0.019)	(0.050)	(0.119)		
Movie studio buyer	0.025	-0.003	-0.034	0.041**	
	(0.023)	(0.028)	(0.046)	(0.019)	
Bidding				0.230***	0.204**
				(0.049)	(0.087)
Original				0.119***	0.163***
				(0.042)	(0.049)
Talent attached				-0.033**	-0.029
				(0.017)	(0.019)
Genre	N	N	N	Y	Y
Content source	N	N	N	Y	Y
Survey (sale) year FE	Y	Y	Y	Y	Y
Buyer FE	N	N	N	N	Y
N	701	701	701	1587	981
R-squared	0.017	0.042	0.071	0.170	0.294
# Unique buyers					161

Note: OLS (linear probability) regressions. Columns 1-3 use Sample BL, and columns 4-5 use Sample DDP. The dependent variables of columns 1 and 2 are dummy variables indicating whether a script is listed among the top 5 or the top 20 in a given survey year; and that of column 3 is the group rank variable: 4 = top 5; 3 = top 6-20; 2 = top 21 to just above the threshold number of nominations used in a given survey year; and 1 = the bottom. The dependent variable of columns 4 and 5 is a dummy variable indicating whether a script is Black-Listed. Experienced writers are defined as writers with one or more major writing credits in the previous ten years. Experienced agents are defined as agents associated with more than fifteen prior sales captured in the entire DDP database. Standard errors (in parentheses) of columns 1-4 are clustered at the survey (or sale) year level, and standard errors of column (5) are clustered at the producer level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Black List nominations and writer experience (subsample results)

(a) Subsamples by agency size and agent experience

Sample	Top-10 agency = 0	Top-10 agency = 1	
		Experienced agent = 0	Experienced agent = 1
Dependent Variable	Rank group	Rank group	Rank group
	(1)	(2)	(3)
Panel a: Sample BL			
Experienced writer	-0.107	-0.190***	-0.321**
	(0.188)	(0.053)	(0.111)
N	62	406	233
R-squared	0.169	0.065	0.087
Mean (DV| Exp writer = 0)	1.81	2.12	2.26
Panel b: Sample DDP			
Experienced writer	0.007	-0.100***	-0.260***
	(0.045)	(0.034)	(0.072)
N	268	416	169
R-squared	0.409	0.235	0.563
# unique buyers	87	53	32
Mean (DV| Exp writer = 0)	0.09	0.22	0.45

(b) Subsamples by sales mechanisms

Sample	Sample DDP	
Sales mechanism	Bid = 0	Bid = 1
DV	Black-Listed	Black-Listed
	(1)	(2)
Experienced writer	-0.087***	-0.379**
	(0.018)	(0.165)
N	1546	41
R-squared	0.141	0.389
Mean (DV| Exp writer = 0)	0.14	0.55

Note: OLS (linear probability) regressions. Panel (a) uses Sample BL. The dependent variable in all columns is the Black List rank group variable: 4 = top 5; 3 = top 6-20; 2 = top 21 to just above the threshold number of nominations used in a given survey year; and 1 = the bottom. Experienced writers are defined as writers with one or more major writing credits in the previous ten years. Experienced agents are defined as agents associated with more than fifteen prior sales captured in the entire DDP database. Column 1 uses the subsample of scripts for which the writer is not associated with a top-10 agency, columns 2 and 3 use the subsamples of cases for which the writer is associated with a top-10 agency but with or without an experienced agent. All regressions include the same controls as in the first three columns of Table 2. Panel (b) conducts the same analysis as in Panel (a), using Sample DDP. The dependent variable in all columns is a dummy variable indicating whether a script is Black-Listed. All regressions include the same controls as in column 5 of Table 2, including buyer fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Average effects of being Black-Listed on the probability of release (portfolio fixed effects)

Sample		Exclude top-20	Include only bottom-listed	Exclude top-20	
DV	Release	Release	Release	Release	Release
	(1)	(2)	(3)	(4)	(5)
Black-Listed	0.162***	0.161***	0.206	0.150***	0.209***
	(0.043)	(0.052)	(0.135)	(0.053)	(0.060)
Writer major credits	0.063***	0.066***	0.063***	0.065***	0.064***
	(0.016)	(0.016)	(0.017)	(0.016)	(0.016)
Top 10 agency	0.051	0.061	0.069*	0.065*	0.066*
	(0.038)	(0.039)	(0.041)	(0.039)	(0.039)
Experienced agent	-0.015	-0.019	-0.012	-0.020	-0.014
	(0.041)	(0.045)	(0.051)	(0.044)	(0.045)
Producer major credits	0.002	0.001	0.002	0.001	0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Original	0.041	0.024	0.042	0.026	0.031
	(0.057)	(0.061)	(0.062)	(0.060)	(0.060)
Talent attached	0.068**	0.061*	0.081**	0.062*	0.057*
	(0.032)	(0.033)	(0.036)	(0.033)	(0.034)
Bidding	0.011	-0.022	-0.028	-0.029	-0.012
	(0.078)	(0.076)	(0.112)	(0.074)	(0.079)
Reporting (# links)				0.050	0.083
				(0.035)	(0.051)
Black-listed ×					-0.213**
Reporting (dummy)					(0.095)
Reporting (dummy)					-0.015
					(0.063)
Genre dummies	Y	Y	Y	Y	Y
Content source dummies	Y	Y	Y	Y	Y
Portfolio (buyer×year) dummies	Y	Y	Y	Y	Y
N	733	684	606	684	684
R-squared	0.337	0.347	0.346	0.350	0.356
Unique # of portfolios	191	183	172	183	183

Note: These OLS regressions estimate equation (1) using Sample DDP, controlling for portfolio (i.e., buyer*year) fixed effects. The dependent variable in all columns is a dummy variable indicating whether the movie is theatrically released. Column 2 excludes top-20 Black-Listed scripts, and column 3 keeps only listed scripts that receive the threshold number of nominations. Column 4 replicates column 2, controlling for the number of articles reported on the script by the top two trade magazines in Hollywood two years before the publication date of the Black List, and column 5 includes an interaction term between Black-Listed and whether reported and a dummy indicating whether reported. The standard errors (in parentheses) of all regressions are clustered at the portfolio level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Average effects of being Black-Listed on the probability of release (CEM)

Sample			Exclude top-20	
		Not reported		Not reported
DV	Release	Release	Release	Release
	(1)	(2)	(3)	(4)
Black-Listed	0.119**	0.180***	0.094	0.173**
	(0.046)	(0.055)	(0.063)	(0.073)
Writer major credits	0.143***	0.118***	0.132***	0.106**
	(0.035)	(0.035)	(0.041)	(0.041)
Top-10 agency	0.025	0.029	-0.048	0.159
	(0.056)	(0.064)	(0.096)	(0.113)
Experienced writer	-0.026	0.061	-0.013	0.096
	(0.061)	(0.081)	(0.086)	(0.111)
Producer major credits	-0.002	0.002	-0.006	-0.000
	(0.003)	(0.004)	(0.004)	(0.005)
Original	0.248	-0.006	0.347	-0.142
	(0.165)	(0.202)	(0.301)	(0.174)
Talent attached	0.035	-0.147**	-0.024	-0.166*
	(0.051)	(0.070)	(0.072)	(0.088)
Reporting (# of links)	-0.043		-0.048	
	(0.041)		(0.052)	
Genre dummies	Y	Y	Y	Y
Source dummies	Y	Y	Y	Y
Buyer dummies	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y
N	319	215	205	128
R-squared	0.350	0.339	0.322	0.404
Unique # of buyers	55	33	36	21

Note: OLS regressions using Sample DDP and a coarsened exact matching method. Columns 3-4 exclude top-20 listed scripts; and columns 2 and 4 use only scripts for which we have not found any reporting by the top two trade magazines in Hollywood two years before the publication date of the Black List. The variables used to create the matched sample are writer major credits, top 10 agency, top agent, producer major credits, original, talent attached, and whether there was any reporting prior to publication. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Box office performance and other outcomes

Sample	Released movies					All
DV	log(US BO)	log(US BO)	log(ROI)	log(Budget)	Time to release	Release
	(1)	(2)	(3)	(4)	(5)	(6)
Black-Listed	1.122*	0.649**	0.659**	-0.294	-0.040	0.141***
	(0.599)	(0.322)	(0.318)	(0.253)	(0.120)	(0.040)
log(Budget)		0.965***				
		(0.132)				
Talent attached						0.034
×Black-Listed						(0.061)
Talent attached	0.258	0.227	0.229	-0.051	-0.272***	0.088***
	(0.499)	(0.282)	(0.281)	(0.224)	(0.100)	(0.022)
Other controls	Y	Y	Y	Y	Y	Y
Genre dummies	Y	Y	Y	Y	Y	Y
Source dummies	Y	Y	Y	Y	Y	Y
Buyer dummies	Y	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y
N	243	163	163	163	243	1563
R-squared	0.632	0.794	0.582	0.670	0.346	0.205
Unique # of producers	64	46	46	46	64	162

Note: OLS regressions. Columns 1-5 use only released movies; and column 6 uses all projects. Production budget estimate is available for two thirds of the released movies and, hence, columns 2-4 have fewer number of observations than columns 1 and 5. Time to release is the number of years between the time when the DDP database captures the record and the release date of the movie. ROI is defined as the ratio between US Box office and production budget. The regressions control for the same set of variables as table 5. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Effects of being Black-Listed by writer experience

Method	Portfolio fixed effects		Coarsened exact matching	
Sample		Exclude top-20		Exclude top-20
DV	Release	Release	Release	Release
	(1)	(2)	(3)	(4)
Black-Listed	0.123**	0.139**	0.092**	0.110**
	(0.047)	(0.054)	(0.044)	(0.054)
Black-Listed × Experienced writer	0.198	0.101	0.180	0.115
	(0.123)	(0.137)	(0.127)	(0.164)
Experienced writer	0.151***	0.164***	0.244***	0.157
	(0.044)	(0.044)	(0.083)	(0.104)
Other controls	Y	Y	Y	Y
Genre dummies	Y	Y	Y	Y
Source dummies	Y	Y	Y	Y
Portfolio (buyer×year) dummies	Y	Y	N	N
Year dummies	N	N	Y	Y
Buyer dummies	N	N	Y	Y
N	733	684	412	273
R-squared	0.347	0.352	0.315	0.292
Unique # of portfolios	191	183		
Unique # of buyers			53	44

Note: Sample DDP. The dependent variable in all columns is a dummy variable indicating whether the movie is theatrically released. Experienced writers are defined as writers with one or more major writing credits in the previous ten years. Columns 1-2 include the same set of controls as column 4 in table 4; and columns 3-4 include the same controls as table 5. The standard errors (in parentheses) of columns 1-2 are clustered at the portfolio level; and that of columns 3-4 are clustered at the buyer level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Black List rankings and release probability

DV	Release (1)	Release (2)	Release (3)	Release (4)	Release (5)
Experienced writer	0.212***	0.187***	0.322**	0.333**	0.406
	(0.046)	(0.048)	(0.139)	(0.131)	(0.243)
Top 5	0.309***			0.301***	
	(0.088)			(0.099)	
Top 6-20	0.102			-0.026	
	(0.102)			(0.075)	
Top 21-above bottom	0.069			0.070	
	(0.064)			(0.076)	
Top 20 × Experienced write			0.117	0.193	0.165
			(0.148)	(0.142)	(0.260)
Top 21-above bottom × Experienced write			-0.040	0.123	0.012
			(0.077)	(0.155)	(0.107)
Other controls	Y	Y	Y	Y	Y
Year×Nominations dummies	N	Y	Y	N	Y
Year dummies	Y	N	N	Y	N
Buyer dummies	Y	N	N	Y	Y
N	701	611	611	339	239
R-squared	0.183	0.317	0.318	0.289	0.460
Unique # of nominations		101	101		56
Unique # of buyers				53	37

Note: OLS regressions use Sample BL, and the dependent variable in all columns is a dummy variable indicating whether the movie is theatrically released. The reference rank group excluded is bottom-listed scripts. For columns 3-5, we group top-5 and top 6-20 scripts together as top-20 scripts. Other controls include top 10 agency, experienced agent, major producer credits, sold prior to publication, and major studio buyer (dummy). The standard errors (in parentheses) of columns 2, 3, and 5 are clustered at the Year×Nominations level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Black List rankings and release outcomes: alternative explanations

(a) Budget and performance for released movies

DV	log(US BO)	log(Budget)	log(ROI)
	(1)	(2)	(3)
Experienced writer	0.006	0.098	0.059
	(0.537)	(0.238)	(0.394)
Other controls	Y	Y	Y
Year×Nominations dummies	Y	Y	Y
N	123	95	85
R-squared	0.498	0.422	0.444
Unique # nominations	35	29	24

(b) Subsample split by buyer characteristics

		Sold prior to publication		
		No studio buyer	Studio buyer	Studio buyer & top producer
DV	Release	Release	Release	Release
	(1)	(2)	(3)	(4)
Experienced writer	0.203***	0.279***	0.224***	0.350**
	(0.057)	(0.099)	(0.076)	(0.150)
Other controls	Y	Y	Y	Y
Year×Nominations dummies	Y	Y	Y	Y
N	465	226	188	82
R-squared	0.326	0.376	0.338	0.433
Unique # nominations	90	60	45	28

Note: OLS regressions using Sample BL. Regressions in both (a) and (b) include the same controls as column 2 in table 8. The standard errors (in parentheses) are clustered at the Year×Nominations level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Online Appendices (not for publication)

A. Appendix Tables and Figures

Table A1: Writer experience and Black List nominations: alternative experience measures

(a) Sample BL

Dependent Variable	Rank Group	Rank Group	Rank Group	Rank Group	Rank Group
	(1)	(2)	(3)	(4)	(5)
Writer's writing experience	-0.072***	-0.069**	-0.029***	-0.013***	-0.115
	(0.021)	(0.021)	(0.006)	(0.003)	(0.088)
Writer's directing experience					-0.184***
					(0.052)
Writer's producing experience					-0.189
					(0.143)
Writer's acting experience					-0.016
					(0.078)
N	701	701	701	701	701
R-squared	0.080	0.079	0.086	0.080	0.090

(b) Sample DDP

Dependent Variable	Black-Listed	Black-Listed	Black-Listed	Black-Listed	Black-Listed
	(1)	(2)	(3)	(4)	(5)
Writer's writing experience	-0.045***	-0.036***	-0.009***	-0.007***	-0.108***
	(0.009)	(0.007)	(0.002)	(0.001)	(0.033)
Writer's directing experience					-0.024
					(0.035)
Writer's producing experience					0.017
					(0.033)
Writer's acting experience					-0.050
					(0.036)
N	981	981	981	981	981
R-squared	0.298	0.294	0.295	0.302	0.300
Unique # of buyers	161	161	161	161	161

Note: OLS regressions. Different columns use different measures of writer experience. In both panels, column 1 uses the number of writing credits for movies distributed by top-30 distributors in the previous ten years ('writer major credits' defined in Section 4.2); 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 the survey (sale) year without restrictions of time period or distributor size; column 4 uses the length of a writer's writing career since the year of his/her first movie writing credit; and column 5 uses a dummy variable indicating that the writer has one or more major writing credits (the same as used in the paper) and includes indicators for the writer's other experience—whether he/she has any directing, producing, or acting credits for movies distributed by top-30 distributors in the previous ten years. Panel (a) uses Sample BL, and the dependent variable of all columns is the Black List rank group variable: 4 = top 5; 3 = top 6-20; 2 = top 21 to just above the threshold number of nominations used in a given survey year; and 1 = the bottom. The regressions use the same control variables as in the first three columns of Table 2. Panel (b) uses Sample DDP; the dependent variable of all columns is a dummy variable indicating whether a script is Black-Listed, and all regressions include the same control variables as column 5 in Table 2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A2: Comparison of sold and unsold Black-Listed 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 equals one if a script is associated with a production company or a studio and zero otherwise.

Appendix B. Proof of the results

Proof of Result 1

Proof. Notice that 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. Thus, given each state, the expected number of nominations is the number of voters, n, multiplied by the probability that each voter draws a positive signal s_H given this state. Weighting the two states by their prior probabilities, the expected number of nominations for a script is:

$$\mathrm{E}[m|n,\pi] = \pi np + (1-\pi)n(1-p). \tag{A1}$$

Thus, $\frac{\partial \mathrm{E}[m|n,\pi]}{\partial \pi} = np - n(1-p) > 0$, because $p > \frac{1}{2}$; and $\frac{\partial \mathrm{E}[m|n,\pi]}{\partial n} = \pi p + \pi(1-p) > 0$. □

Proof of Result 2

Proof. To derive the posterior quality, first notice that the probability of receiving m nominations for a script with n and π is:

$$\mathrm{P}(m|n,\pi) = \pi\binom{n}{m}p^m(1-p)^{(n-m)} + (1-\pi)\binom{n}{m}(1-p)^m p^{(n-m)}. \tag{A2}$$

In words, when the true quality of the script is q_H, there are $\binom{n}{m}$ ways to have m evaluators receive positive signals s_H, each with a probability of $p^m(1-p)^{(n-m)}$. When the true quality is q_L, there are also $\binom{n}{m}$ ways to have m evaluators receive positive signals, each with a probability of $(1-p)^m p^{(n-m)}$. $\mathrm{P}(m|n,\pi)$ weights these two states by their respective prior probabilities.

Based on Bayes' rule, the posterior probability that the script is of high quality after observing the actual number of nominations, m, is:

$$\mathrm{P}(q_H|m,n,\pi) = \frac{\pi \mathrm{P}(m|n,q_H)}{\mathrm{P}(m|n,\pi)} = \frac{\pi p^m(1-p)^{(n-m)}}{\pi p^m(1-p)^{(n-m)} + (1-\pi)(1-p)^m p^{(n-m)}}. \tag{A3}$$

Taking derivatives with respect to its three components, we have the following result:

$$\begin{aligned}
\frac{\partial \mathrm{P}(q_H|m,n,\pi)}{\partial m} &= \frac{\partial \pi p^m(1-p)^{(n-m)}}{\partial m}(1-\pi)(1-p)^m p^{(n-m)} - \pi p^m(1-p)^{(n-m)}\frac{\partial(1-\pi)(1-p)^m p^{(n-m)}}{\partial m} \\
&= \pi(1-\pi)p^n(1-p)^n 2(\log p - \log(1-p)) > 0,
\end{aligned}$$

because $p > \frac{1}{2}$.

$$\begin{aligned}
\frac{\partial \mathrm{P}(q_H|m,n,\pi)}{\partial \pi} &= \frac{\partial \pi p^m(1-p)^{(n-m)}}{\partial \pi}(1-\pi)(1-p)^m p^{(n-m)} - \pi p^m(1-p)^{(n-m)}\frac{\partial(1-\pi)(1-p)^m p^{(n-m)}}{\partial \pi} \\
&= p^n(1-p)^n > 0.
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \mathrm{P}(q_H|m,n,\pi)}{\partial n} &= \frac{\partial \pi p^m(1-p)^{(n-m)}}{\partial n}(1-\pi)(1-p)^m p^{(n-m)} - \pi p^m(1-p)^{(n-m)}\frac{\partial(1-\pi)(1-p)^m p^{(n-m)}}{\partial n} \\
&= \pi(1-\pi)p^n(1-p)^n(\log(1-p) + \log p) < 0,
\end{aligned}$$

because $p > \frac{1}{2}$.

□

Appendix C: Data Appendix

As discussed in Section 4.1 above, this study uses two different samples in its analysis. In this Appendix, we further explain how each was created.

C.1 Creation of the two samples

C.1.1 Initial data from the Black List website

Each year, the results of the Black List are available on the Black List's website (https://blcklst.com/lists/) in PDF format. As shown in Figure C.1 below, this provided us with the name of the script, the writer(s), the year that the script was on the Black List, the writer's agent, the agent's agency, and the number of nominations the script received. Starting in 2007, the publications also have contained information about the production company or movie studio associated with the script. A total of 1,224 scripts were listed between 2005 and 2016. As explained in the paper, we use the 701 scripts listed between 2007 and 2014 to take advantage of the buyer-side information and to leave the scripts sufficient time to get released before our outcome data collection time (October 2017).

Figure C.1: 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 on the Black List in 2005.

C.1.2 Release outcome data

To determine the release outcome of a script, we took two steps: first, using information from the Internet Movie Database (IMDb), we confirmed whether a script had been produced; second, we determined whether it had been 'theatrically released,' based on whether TheNumbers.com recorded positive US box office revenues for the movie.

Confirming whether or not a script was ultimately produced can be tricky, as the names of scripts can change during the production process. To address this issue, we used information from the writer's IMDb page to guide our search. We manually searched for each writer's IMDb page.

If the script's writer had no film releases in or after the year the script was on the Black List, we coded the script as not produced. If the writer had films released before October 2017, we coded the film as produced if a film matched the title of the Black List script exactly. This process was aided by IMDb's "also known as" field which often provided the original name of the script, as it would have been called at the time it was eligible for the Black List. If no IMDb title matched the script's name on the Black List script, but the writer had subsequent releases, we reviewed the plot summaries. By comparing the IMDb summaries to the log-lines in the Black List—and using third parties such as Script Shadow when more detailed summaries were required—we were able to learn the plots of both the Black List scripts and IMDb titles. Where these overlapped and the date of the Black List preceded the matching film's release date, we coded the script as produced.

This approach led us to identify 222 out of 601 Black List scripts (31.6 percent) between 2007 and 2014 as produced. For these produced scripts, we manually searched TheNumbers.com to collect the box office information. This step is straightforward, as IMDb titles closely match The Numbers titles. Ultimately, we identified 184 (83 percent) Black List scripts as theatrically released, with the remaining produced titles released at film festivals or sent straight to DVD.

C.1.3 Writer's and producer's experience

To measure the writer's experience, we parsed the writer's IMDb pages to collect all the films for which the writer obtained a writing credit. Simply using all of the writer's past writing credits as a measure of the writer's experience would be misleading, however, as the majority of films recorded on the IMDb are small and independently produced. To address this, we further collected the U.S. distributor information of the films and defined a film as a major release if it was distributed by one of the top 30 movie studios (based on their market shares in terms of quantity in the distribution market). We used the number of major releases prior to a given sale as the writer's experience. When there was more than one writer in the team, we took the maximum of the writers' experience as the team's experience.

We collected and defined the producer's experience similarly, based on information on the person's IMDb page about the past films for which he/she obtained a producing credit.

C.1.4 Agency size and agent's experience

To define agency size, we used the data collected from the entire Deal Done Pro (DDP) database, explained in detail below in section C.2.5, to calculate the market share for each agency. We identified the names of the top-ten agencies and coded an indicator variable accordingly, based on the agency information in the Black List data. Similarly, we use the number of sales associated with a particular agent in the DDP database before a particular point in time to define an agent's experience.

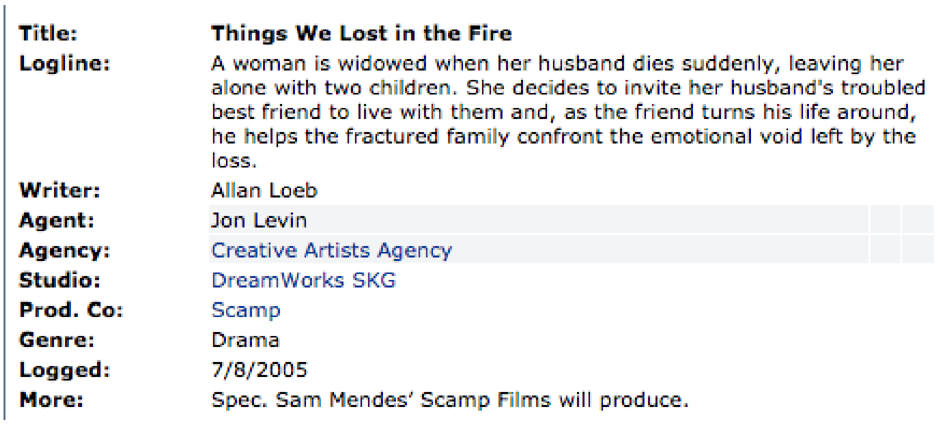
C.2 Sample DDP

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

C.2.1 Initial data from Done Deal Pro

DDP has an online database (http://www.donedealpro.com/default.aspx) of screenplay sales dating back to 1997.[38] Each of these sales contains the title of the script, a brief summary (called a "Logline"), the writer(s), the firms and individuals involved in the sale, and a comments field (called "More") containing information on the sale, as Figure C.2 shows.

Figure C.2: Sample sale entry on the 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 screen shot of a DDP entry.

From these data, we created several of our key variables, including the date of the record (we define as the sale date), writer(s), title, agent, agency, production company, movie studio, and genre. Additionally, by parsing the text in the "More" field, we were able to develop a few other variables:

- **Original:** To identify whether a script was original (as opposed to an adaption), we created a list of terms indicating whether the idea came from a previous source, such as 'book,' 'comic book,' 'tv series,' or 'short story.' We coded the variable "original" as zero if such a term was identified in the "More" field.

- **Complete:** Using a similar approach, we were able to identify whether there was a completed script when the record was captured by the DDP database. In particular, we coded the variable "complete" as zero if the "More" field contained terms such as 'pitch' (an industry term for a sale of an idea without a completed script), 'treatment' (a two-three page outline), 'to be based on,' 'to adapt,' 'will adapt,' 'to be adapted,' 'assignment,' etc. Out of the 6,294 records between 2007 and 2014, 40 percent fall into this category. We coded "complete" as one if the "More" field contained keywords such as 'spec,' 'script,' 'screenplay,' 'based on,' and 'adapted from.' 25 percent of the 6,294 records belong to this category. For another 27 percent of the records, the "More" field did not contain sufficient information for us to determine whether there was a completed script; and these records typically contained only information about the producers managing the project. As explained in the paper, to be

[38]DDP also collects adaptation-right transactions, which can be 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, as they do not yet involve a screenwriter who will be later hired to write the screenplay.

conservative, we excluded these unclear cases from the main analysis and provided robustness checks of our results including them.

- **Turnaround/Rewrite:** For the final 6.5 percent of the records, the "More" field contained the keywords 'rewrite' and/or 'turnaround' (an industry term for projects that were originally developed by a different studio). As explained in the paper, we excluded these cases from the main analysis and provided robustness checks including them in the sample.

C.2.2 Whether Black-Listed

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

C.2.3 Release outcome data

We used the same two-step procedure explained in the previous section (C.1.2) to identify whether or not a project was produced and theatrically released. Because of the size of the DDP database, we used an algorithmic rather than a manual approach.

Specifically, after cleaning the writers' names as listed on DDP to remove any special characters and to split out multiple authors, we first created an algorithm to search for the writer's IMDb page via Google (e.g., 'Paul Thomas Anderson' and 'IMDb') and to select the first IMDb link. To confirm that this would not produce spurious results, we tested approximately 50 different writers of varying experience levels and verified that the site returned the correct writer. We then saved the writer's html page along with the unique IMDb identifier for the writer.

Next, we created a list of all the unique writer-DDP sale pairs (i.e., one pair for each writer, so a DDP sale with three writers would appear three times), using all DDP projects. Using the IMDb information parsed from the step above, we created a corresponding list of writer-IMDb film pairs for the writer's released films before October 2017. We then matched the list of writer-IMDb film pairs and the list of writer-DDP sale pairs by writer name and script title using the matchit function in Stata. For those that were an exact match, we recorded the DDP project as produced into this particular IMDb film. For those that did not match exactly, we manually reviewed the closest matches (that is, where the similarity score generated by the matchit function was above 0.5) and determined whether a match could be found. Separately, there were about 1,000 DDP projects containing 'untitled' in their names (e.g., 'Untitled Facebook Project'). Because matching based on movie titles was unlikely to generate high similarity scores, we manually matched all these titles based on information about the plots, using the same procedure outlined in C.1.2, above. Finally, we matched the produced films to The Numbers database in order to obtain the box office information and determine whether a produced film had been theatrically released.

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

C.2.4 Other variables

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

Appendix D. Additional Results

Appendix D.1 Additional explanations for the negative correlation between writer's experience and Black List nominations

Our conceptual model focuses on two factors: prior quality of the script and its visibility among voters for the Black List. Here, we discuss two additional factors that may also help explain the negative correlation between writer experience and nominations.

First, it is possible that scripts by more-experienced writer are produced more quickly, such that the voters automatically disqualify them.[39] Two exercises confirm that the negative correlation between writer experience and Black List nominations is robust to subsamples for which this concern is minimal. In Sample BL, 73 percent of the listed scripts are not associated a movie studio (the major financing source) at the time of publication and, hence, are far from getting to the stage of principal photography. In this subsample, 11.6 percent of the scripts from experienced writers and 26.7 percent of the scripts from less-experienced writers are ranked in the top 20 (p-value of the difference is 0.031). Furthermore, because the Black List is published annually, in the second week of December, the survey is likely to be conducted in November. The significant negative correlation between writer experience and the likelihood of being listed is robust to using scripts sold close to the survey time (e.g., sold in October and November) , when the likelihood of being in production is very low.

Second, recall that the nomination criterion for the Black List is 'favorite' scripts. Survey respondents may nominate scripts that appeal to them for reasons other than potential box-office appeal, which may lead to an over-representation of less-experienced writers. It is hard to characterize script content; but at a broad genre level, this explanation does not seem to be the main driving force because the only genre in which scripts from experienced versus less-experienced writers differ significantly is adventure, which accounts for less than five percent of observations. Furthermore, regression results confirm the negative relationship between writer experience and being listed, conditional on genre fixed effects (columns 4-5 in table 2). Moreover, additional evidence suggests that experienced producers or producers associated with a movie studio are not less likely to purchase Black-Listed scripts. In fact, they maybe even more likely to do so when one compares Black-Listed and unlisted scripts.[40] Furthermore, box-office revenues of released movies based on Black-Listed scripts are significantly higher than those based on unlisted scripts. Thus, the overall evidence suggests that the nomination does not seem to be driven mainly by the voters' preferences for scripts by less-experienced writers that are divorced from box-office appeal.

Appendix D.2 Unsold Black-Listed scripts

Out of the 145 Black-Listed scripts that are unsold at the time of the list's publication, only eight (six percent) were eventually released. To further understand why so few eventually made it to the theater, we unpack the six-percent 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.

[39] The stated criteria for Black List inclusion did change from "will not be released in theaters during this calendar year" over 2005-2010 to "will not have begun principal photography during this calendar year" after 2011. With the former criteria, the above concern would likely be unimportant because the data show that the median time from sale to release is about 2.5 years and the 5th percentile is one year. Even with the latter criteria, it still takes at least a few months to revise the script, assemble cast and crew, and secure financing before shooting begins.

[40] In particular, in Sample BL, the proportion of top-20 ranked scripts that are associated with a movie studio is exactly the same as that for lower-ranked scripts (both at about 36 percent); and in Sample DDP, 27 percent of Black-Listed scripts are associated with a movie studio, whereas the percentage is 33 percent for unlisted scripts (p-value < 0.001).

To obtain an estimate of the sale rate, we match these unsold scripts to the entire DDP database. We find a match for 26 percent of them, which is a lower-bound estimate 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 Listed scripts that are sold and unsold before publication, we obtain a point estimate of a 40-percent sale rate.[41] Though certainly 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 obtain a conditional release rate of 13.5 percent (8/59). Though obviously a very small sample, this release rate is far below the (conditional) release rate for Black-Listed scripts that were already sold (32 percent) and much nearer to that for unlisted sold scripts from Sample DDP (16 percent).

The above statistics show that unsold Black-Listed scripts do not have a low release rate because they do not get sold, but because these scripts, once sold, have a release rate on a par with or slightly lower than those that were never listed. There are, at least, three broad possible reasons that this could occur: (i) there may be a strong selection effect that results in these scripts having less-viable material than listed but sold scripts;[42] (ii) relatedly, due to the fact that the script is not yet sold, the market may infer that it is of poorer quality and, hence, be hesitant to put a great deal of resources into it; and (iii) finally, the Black List may carry momentum for only a short period of time, which would mean that in the time that it takes for these scripts to be sold, the previous year's Black List would already be essentially forgotten, leaving these scripts without the force necessary to overcome the hurdles inherent in the rest of the process. Unfortunately, the small sample size and incompleteness of the data restrict our ability to investigate these conjectures further.

[41] 356 out of 556 (64 percent) of Black-listed scripts that are sold before publication are captured by the DDP database.

[42] The data suggest some selection effect, at least on the observables. In particular, table A2 shows that among listed scripts, relative to sold ones, unsold scripts are ranked lower on the list, are much less likely to come from experienced writers; and are likely to be represented by smaller agencies and less-experienced agents.