When to Sell Your Idea:
Theory and Evidence from the Movie Industry

December, 2012

Abstract

How completely should an innovator develop his idea before selling it? In the context of selling original movie ideas, I present a model that features the writer’s private information on the idea’s value, different protection levels associated with different development stages, as well as costly buyer participation. The empirical results are consistent with the model’s predictions: Inexperienced writers are excluded from the market for earlier-stage ideas, restricting their choices to developing the idea fully or abandoning it; writers who have a choice sell better ideas at a later stage and worse ideas at an earlier stage; and lastly, intermediaries facilitate earlier-stage sales for inexperienced writers.

1 Introduction

A fundamental dilemma in selling ideas is that it is difficult for a potential buyer to assess an idea’s value before disclosure, but once the idea is known, the buyer would have little incentive to pay the seller (Arrow (1962)). A theoretical literature has developed to study how the seller appropriates rents when there is little or no property-right protection (e.g., Anton and Yao (1994, 2002)). The basic ideas are either to influence the buyer’s expropriation incentives or to solve the evaluation problem through signaling. This literature typically takes an idea as given. If we take a step back and consider the innovator’s investment decisions before any selling activities, there is often huge uncertainty, as well as significant (and irreversible) investments at stake. This leads to the central question of this paper: How completely should an innovator develop his idea before selling it?

Another challenge an innovator faces, before the disclosure-expropriation dilemma kicks in, is how to get the buyer to even listen. The buyer’s participation incentives have gotten less attention in the literature, with a few exceptions (Anton and Yao (2008) and Hellmann (2007a)). When the innovator possesses private information, how does his choice of the sale stage affect the buyer’s willingness to listen?

This paper studies the choice of the sale stage in the context of scriptwriting, in which writers sell original movie ideas to studios in Hollywood. A writer typically does one of the following: sell the storyline and (if sold) get paid to write the script (pitch); or write up the idea and sell a
complete script (spec). Thus, in this context, the choice of the sale stage is a decision of whether or not to sell an idea with a script: to spec or to pitch.

I construct a simple model that links the sale stage to the writer’s experience and his private information. The former serves as the basis for the buyer’s assessment of the expected value of the idea before knowing the details, and the latter reflects how much this particular idea’s value might deviate from the writer’s average. I then test the model with a sample of 1,847 ideas sold between 1997 and 2005 in Hollywood. A nice feature of the industry is that its relatively short life cycle allows us to observe a sale’s final outcome, which provides valuable information on an idea’s quality.

Even though I study this question in a specific setup, the basic trade-offs that writers and studios face are relevant in a variety of contexts, particularly in industries in which established firms rely increasingly on external sources for innovation (Arora et al. (2001)). An important example is the biotech industry; fewer than one third of the drugs are marketed by firms that brought them into phase I FDA trials (Levine (2009)). Understanding these trade-offs helps us address questions such as: At what stage should a start-up license its technology or form a research alliance with a downstream firm? Should an entrepreneur go to the VC with a business plan or with a more developed product?

The model has the following key elements. A writer has a nascent idea and decides at which stage to sell it to a potential buyer or to simply drop it. Before the idea is disclosed, the buyer observes the writer’s past experience, but not the writer’s private knowledge about this particular idea’s value. The buyer, in turn, decides whether to meet with the writer and incurs a cost if she does meet. After the idea is revealed, the buyer provides extra information on its value; her knowledge might come from extensive downstream experience or from idiosyncratic demand that is unknown to the writer. Finally, disclosing the idea exposes the writer to expropriation; however, this risk decreases as the idea is further developed, largely because IP protection is more effective.

From the writer’s point of view, there is a trade-off between pitching and speccing. One advantage of pitching is the informational value of the buyer’s feedback on the idea’s demand; if the idea turns out not to be that interesting, the writer can save himself the cost of writing the full script. Another advantage of pitching is a better bargaining position because the writer has not sunk the costs yet. However, pitching implies a greater risk of expropriation, leaving the writer with a smaller share of the idea’s surplus.

---

1The Feb 2, 1999 issue of Variety, an industry trade magazine, reported that among 146 movies released by major studios in 1998, 12 percent originated from pitches, and 43 percent originated from specs. The rest were based on books, plays, comic books, sequels, TV shows, remakes, etc.

2In addition to actual costs and the opportunity cost of time, listening to an idea also exposes the buyer to potential legal disputes. Movie studios routinely refuse to listen to pitches and return unsolicited scripts to writers unopened. An anecdote is that producers David Zucker and Jerry Zucker have a sign posted on their office door at Columbia Pictures: “Thank you for not pitching us your idea.” In fact, it is normal for many firms (e.g., Microsoft, IBM, and venture capitalists) to refuse to examine an unsolicited idea unless the independent inventor signs a waiver of confidentiality rights.

3This “informational value” of pitching is reflected in a quote from The Artful Writer: “It takes about 1-2 weeks to develop a solid pitch. If they bite, you go into draft knowing you’re not wasting your time.”
When would we expect the writer to pitch his idea rather than spec it? When being met by the buyer is not a problem, given any experience level, the writer would follow a threshold strategy: spec better ideas and pitch worse ideas. The writer’s choice, therefore, signals his private information. But because it is costly for the buyer to participate, not all writers would get a chance even to disclose—especially inexperienced writers trying to sell pitches. An immediate prediction of the model is that the likelihood of a sale being a spec is non-monotonic with respect to the writer’s experience. The likelihood of a spec is high for inexperienced writers because of the difficulty of securing a pitch meeting, it drops once the writer is good enough to get his pitches heard, and it eventually increases again because top writers want to spec more often in order to capture more of their ideas’ surplus.

How does the quality compare between specs and pitches? It is particularly interesting to see in the data whether sellers select better ideas to sell at a later stage. Empirical evidence for sellers’ selection behaviors in an IP market has been scarce and mixed (e.g., Pisano (1997), Arora et al. (2009), and Danzon et al. (2005)). Speaking to this, the model shows a countervailing effect: Buyers screen pitches for an extra round at the idea stage. This extra-screening effect might offset the seller-selection effect when comparing the outcome of ideas sold at different stages. Without observing rejected ideas and, hence, their performance, it is hard to differentiate these two effects. Interestingly, however, the model shows that the relative importance of these two effects varies with the writer’s experience. The prediction is that conditional on release, the expected performance of a spec increases faster than that of a pitch; therefore, it is likely that specs outperform pitches for sufficiently experienced writers, while the opposite is true for inexperienced writers.

Lastly, part of the reason why inexperienced writers face barriers to selling pitches is information asymmetry, which provides a role for intermediaries. Intermediaries—talent and literary agencies in this case—are important players in this market. They represent writers through two-year contracts and earn a commission of ten percent of the writer’s income. The model implies that being affiliated with a reputable intermediary would make inexperienced writers more likely to sell at an earlier stage. Because pitching incorporates the buyer’s information about an idea’s potential demand before significant investments, intermediaries potentially reduce inefficiency.

I empirically test these predictions on a novel sample of original movie ideas sold between 1997 and 2005. The primary data source is Done Deal Pro, an internet database that tracks idea and script transactions in Hollywood on a daily basis. I observe the sale stage, the names of the writer and the buyer, price (for half of the sample), and other information about the idea. Matching the sales data to IMDb and TheNumbers, I also observe the writer’s industry experience and the sale’s outcome (e.g., whether it is eventually released in theaters, and, if so, its box office revenues).

Empirically, the writer’s experience level is measured by the number of the writer’s major writing credits in the previous five years. First, I find that, other things being equal, the predicted likelihood of a spec sale is 0.58 for writers with zero credits; it drops to about 0.4 for writers with one or two credits, and increases again to above 0.5 for writers with three or four credits. Both the decline and the increase are economically and statistically significant. This result is consistent with
the model’s prediction; in particular, it provides evidence for costly buyer participation because otherwise, the model suggests a monotone increasing relationship between the likelihood of a spec and the writer’s experience, which is what I find for writers with some major credits.

Second, conditional on release, the predicted box office revenue of a spec is not different from that of a pitch for inexperienced writers, but it increases faster with the writer’s experience, and eventually performs significantly better. These patterns are consistent with the model’s prediction on how the two countervailing effects (writer-selection effect and extra-screening effect) play against each other for writers of different experience levels. We can reasonably conclude that the data show evidence consistent with the seller’s selection behavior.

Finally, the data show that inexperienced writers affiliated one of the five biggest agencies in Hollywood are significantly more likely to sell pitches, and for relatively experienced writers, such affiliation generally makes no significant difference in the sale stage. These results are consistent with the intermediaries’ role of reducing information symmetry between the seller and the buyer, which relaxes the barrier to an earlier-stage idea market for inexperienced sellers.

The empirical results are robust to alternative ways of measuring the writer’s experience. Data on the likelihood of release and purchasing price also suggest that potential mismeasurement of the writer’s experience should not affect the results significantly. But, as I will discuss later, it is important to recognize that this paper is interpreting the empirical observations through the lens of a particular model, which I believe is an important story. The data and institutional knowledge help to rule out a number of alternative explanations, but not all.

This paper is closely related to the literature on the sale of ideas (e.g., Bhattacharya and Ritter (1983), Anton and Yao (1994, 2002), and Biais and Perotti (2008)). When some protection exists, the choice involving the use of property rights could influence the buyer’s expropriation incentives and solve the evaluation problem through signaling (e.g., Gallini and Wright (1990), Bhattacharya and Guriev (2006), Anton and Yao (2005, 2008)). The model also shares common features with other applications of the Arrow problem. One category involves the employee’s choice of selling his invention to his employer or leaving the firm to form a start-up (Anton and Yao (1995), Anand and Galetovic (2000), Baccara and Razin (2006, 2012), Hellmann (2007b)). Another category is related to how concerns about employee expropriation (of the firm’s private knowledge) shape the organizational structure (Rajan and Zingales (2001) and Hellmann and Perotti (2011)).

Fewer papers share the concern of obtaining the buyer’s attention. Anton and Yao (2008) explain how waiving confidentiality rights could encourage the buyer to participate. Hellmann (2007a) studies the entrepreneur’s soliciting strategy when each buyer has the incentive to wait and free-ride on the others’ evaluation. In this paper, the development stage of an idea signals the idea’s value, and, thus, affects the buyer’s willingness to meet. Exploiting the relationship between the sale stage and the writer’s experience, I find evidence that is consistent with costly buyer participation and with the role of intermediaries in reducing information asymmetry. Elfenbein (2007) argues that researchers’ academic output and the likelihood of licensing are positively correlated because academic status attracts buyers’ attention.
This paper is also related to the literature on the commercialization strategy of research-focused firms—for example, whether to compete or cooperate with downstream firms (e.g., Teece (1986), Arora et al. (2001), and Gans et al. (2002)).\(^4\) Within the cooperation mode, a number of papers have also studied the stage/timing of sales. See Allain et al. (2012) for impacts of downstream market structure on the stage of licensing in the biotech industry; Jensen et al. (2003) on faculty’s choice of disclosing their inventions at the proof-of-concept versus prototype stage and its effects on the terms of licensing; and Gans et al. (2008) on how resolving uncertainties over the scope of IP rights affects the timing of licensing. Complementing these studies of technology markets, this paper provides some fresh evidence on a market for creative ideas. It also has a different focus: the relationship between the sale stage and the seller’s observable quality and experience. The paper also goes a step further to examine the outcome data, thus, making inferences about the quality of ideas associated with different sale stages.

The paper also touches on the certification role of intermediaries described in Biglaiser (1993), and Lizzeri (1999). Our results are consistent with a recent study by Stanton and Thomas (2012) on outsourcing intermediaries for the online hiring marketplace. They find that intermediary affiliation signals that affiliated workers are of a higher quality than independent contractors, and that such affiliation mitigates inefficiently low rates of inexperienced-worker hiring.

Finally, this paper contributes to the study of the organization of production in creative industries (e.g., Caves (2000); Gil and Spiller (2007)). There have been a large number of studies on various aspects of the movie industry:\(^5\) for example, on estimating the consumer demand (e.g., De Vany (1999); Eliashberg et al. (2006); Einav (2007)), the choice of financing (e.g., Fee (2002); Goettler and Leslie (2004)), strategic choice over release dates (e.g., Chisholm (2000); Corts (2001)); firm boundaries and contracts (e.g., Weinstein (1998); Gil (2007, 2009); Natividad (2012)). Most prior studies, however, typically use data on movies that have been released. This paper provides new evidence on the supply of creative ideas at the early stage of the production process. Closely related papers are Harris et al. (2012) and Goetzmann et al. (2012), which use similar data but focus on specs only. They study the pricing and contractual design for intellectual properties when there exist soft information and differences in opinions.

## 2 Theory

A writer, \(W\), has a nascent idea that he wants to sell to a buyer, \(B\). Both players are risk-neutral. The idea’s ultimate value, \(V\), is the sum of four parts:

\[
V = w + \theta + \epsilon_i + \epsilon_s.
\]

\(w\) represents the writer’s experience, which is observable to the buyer. It helps the buyer to assess the idea’s average value before learning about its specifics. \(\theta\) measures how much this particular

\(^4\)Cooperation includes technology licensing, research alliance, and outright purchase of the IP.

\(^5\)Please see Eliashberg et al. (2007) for a comprehensive review.
idea’s value deviates from the average. It is privately observed by the writer before the details are disclosed. Ex-ante, the writer is still uncertain about the idea’s value: $\varepsilon_i$ and $\varepsilon_s$ are, respectively, the uncertain quality of the idea and of the script. The realization of these uncertainties happens after the buyer’s evaluation. The buyer could provide valuable information on an idea’s demand because of her extensive experience in producing and marketing movies, or because of her idiosyncratic demand that is unknown to the writer.\textsuperscript{6} This assumption allows us to capture the potential gain from obtaining the buyer’s feedback before making a substantial investment. For simplicity, assume that the support of $w$ and the random variables are $\mathbb{R}$, and they are independent of each other.\textsuperscript{7}

At the beginning of the game, $w$ and the distributions of the random variables are common knowledge; and the writer privately observes $\theta$. The game proceeds as follows:

Stage 1. Given $w$ and $\theta$, the writer decides to spec, to pitch, or to drop the idea. If spec, the writer pays the writing cost, $c_s$; if pitch, the writer pays no costs; and if drop, the game ends.

Stage 2. Given $w$ and the writer’s choice, the buyer decides whether to meet.\textsuperscript{8} If they meet, the buyer pays a meeting cost, $c_m$; and the game ends otherwise. In addition to actual costs and opportunity cost of time, the meeting cost also reflects the legal risk from the exposure to an idea.\textsuperscript{9}

Stage 3. Given a meeting, the storyline is disclosed if a pitch, and the script is disclosed if a spec. The buyer now also observes $\theta$. In addition, part or all of the uncertainties are realized: $\varepsilon_i$ is realized for a pitch, and both $\varepsilon_i$ and $\varepsilon_s$ are realized for a spec. They are observed by both players.

Stage 4. The idea is dropped if its expected surplus is negative. Given positive surplus, the writer makes a take-it-or-leave-it offer.\textsuperscript{11} For a spec, the buyer transfers an amount to the writer in exchange of the script. For a pitch, the buyer pays the writer; the writer pays the writing cost; and, given realized $\varepsilon_s$, the buyer decides whether to continue.

The writer makes the offer under the shadow of buyer expropriation. For example, the buyer may pass on the content to another writer and commission a script. Assume that the buyer’s expected payoff from expropriation is $(1 - \lambda_s)$ (resp. $(1 - \lambda_p)$) proportion of the spec’s (resp. pitch’s) expected surplus, where $\lambda_s$ and $\lambda_p$ measure the strengths of appropriability conditions for the seller (e.g., the likelihood that the buyer is caught). More importantly,

\textsuperscript{6}E.g., the buyer might already be developing a similar idea, which would decrease the value of the writer’s idea.

\textsuperscript{7}The independence between the variables is not necessary. It is sufficient that the expected value of an idea given $(w, \theta, \varepsilon_i)$ has properties roughly consistent with what is described in Lemmas 1 and 2 (in the Appendix). For example, $\varepsilon_s$ can be positively correlated with $w$ (i.e., writers with better $w$ are likely to write better scripts).

\textsuperscript{8}If a pitch, meeting usually means an actual meeting, in which the writer presents the storyline to the buyer. If a spec, I use meeting to mean that the buyer decides to read the script. An actual face-to-face meeting might or might not happen before the purchasing decision.

\textsuperscript{9}After an idea is rejected, for whatever reason, the writer might find a later-released movie by the studio similar to his rejected idea and take legal actions. Because proof of access needs to be met in an infringement or an idea-theft claim, avoiding the direct contact to an idea is the first layer of defense against such legal risks.

\textsuperscript{10}Assuming that the buyer volunteers her assessment truthfully avoids modeling bargaining under asymmetric information. However, this assumption is not particularly consistent with a later assumption that the writer makes a take-it-or-leave-it offer because the buyer is leaving information rent on the table. My conjecture is that modeling bargaining under asymmetric information would not change the results qualitatively if the buyer’s incentives to capture information rent through manipulating her assessment is not too different between a pitch and a spec.

\textsuperscript{11}The qualitative results do not change when allowing both players to have some bargaining power—e.g., with a generalized Nash bargaining solution where the writer’s bargaining power is an arbitrary number between 0 and 1.
**Assumption 1.** $\lambda_s > \lambda_p$.

First, an obvious motivation for this assumption is that copyright protection is more effective when there is a complete script. Copyright does not protect abstract ideas; what it protects is the particular way an idea is expressed in a written work. Even though many writers prepare a treatment (a written account of the storyline, often no more than a couple of pages) when they pitch, the plots, dialogues, and characters in a complete script are much more concrete than those in a treatment. Therefore, the probability of winning an infringement suit is, in general, greater with a complete script. Second, that a spec is better protected can also be motivated by non-legal theories. Anton and Yao (1994) argue that the threat of selling the idea to a rival buyer (hence, undermining the current buyer’s monopoly position) allows the seller to capture a sizable share of the surplus. Applying this intuition, a spec imposes a greater threat to the current buyer than a pitch does because a rival buyer can bring a spec to the next stage—and eventually to the market—much faster than she could a pitch.\(^{12}\)

The following assumption on relative costs simplifies the analysis and is broadly consistent with the reality in the movie industry. Relatively speaking, for a spec, the writer’s upfront cost (i.e., the writing cost) is greater than the buyer’s (i.e., the meeting cost); that is,

**Assumption 2.** $\frac{c_s}{\lambda_s} > \frac{c_m}{1 - \lambda_s}$.

Technically, the assumption bounds from above the writer’s share of the idea’s surplus, $\lambda_s$, with respect to other parameters. It helps rule out a scenario that is not supported by the data. Also, having an upper bound on the protection level is not an unreasonable assumption given the general difficulty of enforcing IP rights.

Lastly, assume that the random variables have the following standard properties.\(^{13}\)

**Assumption 3.** The probability distributions of $\theta$, $\epsilon_i$, $\epsilon_s$, and $\epsilon_i + \epsilon_s$ have 1) a monotone increasing hazard rate, and 2) a monotone decreasing reversed hazard rate.

### 2.1 Equilibrium

I solve the game for a semi-separating perfect Bayesian equilibrium. By backward induction, I start with stage 4, from which the equilibrium payoffs enter the writer’s and the buyer’s problems.

**A. Equilibrium payoffs at stage 4**

Given a spec, all uncertainties are realized. The idea is dropped if $V$ is negative. Given positive value, the writer’s and the buyer’s payoffs are $\lambda_s V$ and $(1 - \lambda_s)V$, respectively. Recall that the buyer’s expected payoff from expropriating the idea is $(1 - \lambda_s)V$, which is what the buyer gets when the writer has all the bargaining power.

\(^{12}\)This argument is similar to “lead time,” which is found to be the most effective appropriation mechanism in many industries (Levin et al. (1988) and Cohen et al. (2000)).

\(^{13}\)Log-concave distributions have such properties (see Bagnoli and Bergstrom (2005) for a list of examples).
Given a pitch, part of the uncertainties, \( \epsilon_i \), is realized. The buyer has a chance to terminate the project once the script is finished (i.e., after \( \epsilon_s \) is realized); the expected value of a pitch is then

\[
v(w, \theta, \epsilon_i) = \mathbb{P}(V \geq 0)\mathbb{E}_{\epsilon_s}[V|V \geq 0].
\]

(1)

A pitch is sold if and only if it is worth writing; i.e., \( v(w, \theta, \epsilon_i) \geq c_s \). Under the threat of buyer expropriation, the writer’s and the buyer’s equilibrium payoffs are, respectively, \( \lambda_p(v(w, \theta, \epsilon_i) - c_s) \) and \( (1 - \lambda_p)(v(w, \theta, \epsilon_i) - c_s) \).

B. Writer’s problem at stage 1

Given \( w \) and \( \theta \), the writer anticipates the buyer’s meeting decisions at stage 2, and decides to spec, to pitch, or to drop the idea accordingly. Dropping the idea yields zero payoff. Conditional on meeting with the buyer, the writer’s expected payoff from a spec is the probability that a spec is sold multiplied by his expected payoff conditional on sale and minus the writing cost; that is,

\[
S^W(w, \theta) = \mathbb{P}(V \geq 0)\mathbb{E}_{\epsilon_i, \epsilon_s}[\lambda_s V|V \geq 0] - c_s.
\]

His expected payoff from a pitch, conditional on meeting the buyer, is similarly defined, except that the writer incurs the writing cost only if the pitch is sold.

\[
P^W(w, \theta) = \mathbb{P}(v(w, \theta, \epsilon_i) \geq c_s)\mathbb{E}_{\epsilon_i}[(1 - \lambda_p)(v(w, \theta, \epsilon_i) - c_s)|v(w, \theta, \epsilon_i) \geq c_s].
\]

C. Buyer’s problem at stage 2

Observing the writer’s experience, \( w \), and his choice of the sale stage, the buyer decides whether to meet the writer. She updates her belief about \( \theta \) according to the writer’s strategy and Bayes’ rule. Let \( h(\theta|w, S) \) be the buyer’s posterior of \( \theta \) seeing a writer of \( w \) choosing to spec, and \( h(\theta|w, P) \) is that for a pitch. The buyer’s expected payoffs from meeting a spec and a pitch are the integrations of her payoffs over each possible value of \( \theta \) according to her respective posterior beliefs:

\[
S^B(w) = \int \{\mathbb{P}(V \geq 0)\mathbb{E}_{\epsilon_i, \epsilon_s}[(1 - \lambda_s)V|V \geq 0] - c_m\} h(\theta|w, S)d\theta,
\]

\[
P^B(w) = \int \{\mathbb{P}(v(w, \theta, \epsilon_i) \geq c_s)\mathbb{E}_{\epsilon_i}[(1 - \lambda_p)(v(w, \theta, \epsilon_i) - c_s)|v(w, \theta, \epsilon_i) \geq c_s] - c_m\} h(\theta|w, P)d\theta.
\]

D. Equilibrium

There exists a unique semi-separating equilibrium. The uniqueness is the sense of the players’ equilibrium strategies, which can be supported by multiple belief systems that are different off the equilibrium path. All proofs are in the appendix.

**Proposition 1.** There exists a unique semi-separating equilibrium, in which the buyer always meets the writer for a spec, and meets the writer for a pitch if and only if \( w \geq \bar{w} \). The writer’s strategy is such that:
(i) when \( w \geq \bar{w} \), he specs if \( \theta \geq r_0(w) \) and pitches otherwise;

(ii) when \( w < \bar{w} \), he specs if \( \theta \geq r_s(w) \) and drops the idea otherwise.

Furthermore, \( r'_0(w) = r'_s(w) = -1 \).

Figure 1: Writer’s Choice in Equilibrium

Figure 1 illustrates the writer’s choice in equilibrium. There are three notable features. First, relatively experienced writers (i.e., \( w \geq \bar{w} \)), follow a threshold strategy: they select better ideas to spec and worse ideas to pitch. To see this, write the writer’s trade-offs between a spec and a pitch,

\[
\Delta W(w, \theta) = S^W(w, \theta) - P^W(w, \theta),
\]

as follows:

\[
\Delta W(w, \theta) = \lambda_p P(v(w, \theta, \epsilon_i) < c_s) \mathbb{E}_{\epsilon_i} [v(w, \theta, \epsilon_i) - c_s | v(w, \theta, \epsilon_i) < c_s] \\
- (1 - \lambda_p)c_s + (\lambda_s - \lambda_p) \mathbb{E}_{\epsilon_i} [v(w, \theta, \epsilon_i)].
\]

(2)

Pitching is desirable for two reasons. One is the informational value of the buyer’s feedback, so that if the idea turns out not to be that interesting, the writer can save himself the writing cost. \( \lambda_p P(v(w, \theta, \epsilon_i) < c_s) \mathbb{E}[v(w, \theta, \epsilon_i) - c_s | v(w, \theta, \epsilon_i) < c_s] \) reflects the writer’s share of the improved efficiency relative to a spec. Pitching is also desirable because the writer has not yet sunk the writing cost; and \(-(1 - \lambda_p)c_s\) reflects this more favorable bargaining position of pitching. However, a spec is desirable because the writer is able to obtain a greater share of the idea’s surplus (due to a smaller risk of expropriation). This appropriation advantage of a spec is reflected by \((\lambda_s - \lambda_p) \mathbb{E}[v(w, \theta, \epsilon_i)]\).

For any given \( w \geq \bar{w} \), the writer follows a threshold strategy because a spec becomes more attractive as \( \theta \) increases (i.e., \( \Delta W(w, \theta) \) is monotone increasing in \( \theta \)). In particular, bigger \( \theta \) means a better expected value of the idea, which implies: a bigger incentive to capture a greater share of the surplus; and less of a need for interim feedback from the buyer.

\[\text{Notes: } \bar{w} \text{ is the buyer’s meeting threshold for pitches. When } w < \bar{w}, \text{ the writer is indifferent between spec'ing and dropping the idea when } \theta = r_s(w); \text{ and when } w \geq \bar{w}, \text{ the writer is indifferent between spec'ing and pitching the idea when } \theta = r_0(w).\]
The second feature of Figure 1 is that \( r_0(w) \) decreases with \( w \), implying that writers with better \( w \) spec more often. This is because \( w \) and \( \theta \) both contribute positively to the idea’s expected value. Therefore, ideas from writers with better \( w \) need to have lower \( \theta \) to be above the threshold.\(^{15}\)

Lastly, the buyer also follows a threshold strategy in meeting the writer, but more interestingly, she is stricter about meeting a pitch than a spec. As a result, inexperienced writers (i.e., \( w < \bar{w} \)) either have to develop a full script or drop the idea entirely. To see this, consider the marginal writer, \( \bar{w} \), with whom the buyer is indifferent between meeting or not for a pitch. However, the buyer might still be happy to meet him for a spec because: 1) a spec signals a better posterior distribution of \( \theta \); and 2) the writer has already sunk the writing cost. But on the flip side, the buyer is able to appropriate less when the writer sells a spec (i.e., \( 1 - \lambda_s < 1 - \lambda_p \)). Therefore, the buyer would be happy to meet \( \bar{w} \) for a spec as long as she does not need to give up too much more of the surplus (i.e., \( \lambda_s \) is not too much higher than \( \lambda_p \)). Assumption 2 implies that \( \lambda_s < \frac{c_p}{c_m + c_s} \) and is a sufficient condition to guarantee that this is the case.\(^{16}\)

### 2.2 Empirical Implications

Immediately from Figure 1, we see that conditional on sale, the likelihood of a spec is non-monotonic with respect to the writer’s experience. For inexperienced writers (i.e., \( w < \bar{w} \)), the likelihood of a spec is one because the buyer would not meet them for pitches. Once the writer is good enough to get his pitches heard (i.e., \( w \geq \bar{w} \)), the likelihood of a spec drops. However, as the writer’s experience level increases, it is in his own interest to spec more often. Thus, the likelihood of a spec increases again.\(^{17}\)

**Hypothesis 1.** Conditional on sale, the likelihood of a spec is high for both inexperienced and highly experienced writers, and low for writers in the middle.

Also clear from Figure 1 is that part of the reason why inexperienced writers have difficulty in selling pitches is information asymmetry. In particular, the buyer does not observe \( \theta \) and, therefore, cannot base her meeting decision on it. This provides a role for intermediaries. Intermediaries (talent and literary agencies in this case) are important players in this market. They represent writers through two-year contracts and earn a commission of ten percent of the writer’s income.

If a reputable intermediary can credibly convey the value of \( \theta \) to the buyer,\(^{18}\) writers with \( w < \bar{w} \) would have a chance to pitch, as well. The intermediary could also help relax the barriers

---

\(^{15}\)Note that the linearity of \( V \) with respect to \( w \) and \( \theta \) is not necessary. It is sufficient that \( V \) is an increasing function of \( w \) and \( \theta \), which is a reasonable property.

\(^{16}\)The equilibrium of the game depends critically on the fact that the buyer is more selective in accepting a pitch than a spec. Suppose that \( \lambda_s \) is close to 1 (i.e., scripts are extremely well protected) and \( c_m \) is sufficiently high. Then, a buyer would not accept a spec if \( w \) is sufficiently low, while she might still want to meet a pitch. This possibility is ruled out by Assumption 2. As is discussed after the assumption, this possibility is not quite consistent with what is generally true in the industry, and it is also rejected by the data.

\(^{17}\)For \( w \geq \bar{w} \), the model suggests that the writer specs more often as \( w \) increases. Whether this increasing relationship holds once conditional on sale depends on the distribution assumptions. It holds for normal distributions, but, more generally, for distributions satisfying Assumption 3, it holds when \( w \) is sufficiently high.

\(^{18}\)To protect both parties, it is conventional that no details are disclosed until the buyer agrees to meet. But a reputable agent might be able to convince the buyer to meet because she trusts the agent’s taste.
to selling a pitch through signaling the value of $w$. There is likely positive sorting between writers and agencies; that is, better agencies are likely to represent better writers (see Sorensen (2007) and Stanton and Thomas (2012)). Then, observing two writers who are otherwise similar, the buyer would infer that the writer affiliated with a better agency has a higher $w$ and, hence, would be more likely to agree to a pitch meeting. Because pitching incorporates the buyer’s information about an idea’s potential demand before significant investments, intermediaries potentially reduce inefficiency.

The following hypothesis describes the relationship we expect to see in the data between the affiliation with a reputable intermediary and the choice of the sale stage.

**Hypothesis 2.** Sales from writers affiliated with a reputable intermediary are less likely to be spec—especially for inexperienced writers.

Lastly, how does the quality of specs and pitches compare? Of particular interest is the writer’s selection behavior suggested by the model. It is hard to observe an idea’s quality directly. Previous studies typically use ex-post measures such as the likelihood of success (e.g., obtaining the FDA approval of a drug) to infer a project’s quality. Here, however, it is not appropriate to directly compare the likelihood of success (e.g., whether the movie is released) between spec and pitch sales because pitches are intrinsically more uncertain (i.e., $\epsilon_s$ is not realized yet). What we would like to observe is the set of pitch sales that are continued to the next stage after the first drafts are written. Unfortunately, we do not observe this intermediate step. The only point in the data at which the outcome is comparable between specs and pitches is when the movie is released.

The model suggests two differences between specs and pitches at the sourcing stage that have countervailing effects on their relative performance. On the one hand, the writer’s selection behavior implies that the distribution of $\theta$ of a spec is better than that of a pitch. This *writer-selection effect* makes specs better, on average, regardless of $w$. On the other hand, the buyer screens pitches for an extra round; that is, pitches are purchased only when $\epsilon_i$ is above a certain threshold. This *extra-screening effect* makes pitches better, on average, regardless of $w$.

Interestingly, the relative importance of these two effects varies with $w$. In particular, the extra-screening effect is the biggest for low $w$, and diminishes as $w$ gets better. Because pitches from top writers are most likely to be taken anyway, an extra round of evaluation has minimal effect. In contrast, the writer-selection effect is the biggest for writers of high $w$. This is because the expected value of pitches are bounded from above by a threshold; as a result, even though the expected performance of a spec and a pitch both increase with $w$, the former is theoretically unbounded, while the latter is not. Combining these two effects, we have the following hypothesis.

**Hypothesis 3.** Conditional on release, the movie’s expected performance increases with $w$ for both specs and pitches, with the former increasing faster than the latter. Thus, it is likely that specs perform better than pitches for relatively experienced writers, while the opposite is true for relatively inexperienced writers.
When taking these implications to the data, it is important to note that what is observable to the buyer is not all observable to the researcher. Suppose that \( w = w^o + w^u \), where \( w^o \) is observable to the researcher, but \( w^u \) is not. It is not unreasonable to assume that \( w^u \) is independent of \( w^o \). Then, it is straightforward to show that implications derived with respect to \( w \) also carry through to \( w^o \), which is the empirical measure of \( w \). I will discuss the validity of this assumption after the main results.

3 Data and Empirical Strategy

The primary data source for this paper is *Done Deal Pro*, an internet database that tracks transactions of movie ideas in Hollywood since 1997.\(^{19}\) It covers a significant portion of projects purchased by major studios and big production companies; for example, by manually checking movies released by the major studios in 2008, I can track down about 70% of them in the sales database.

Because this paper studies original-idea sales by writers, I exclude the following cases: 1) transactions of movie rights for literary materials (e.g., a book); 2) commissions from the buyer to adapt material or other people’s ideas into scripts, or to rewrite an existing script; 3) authors adapting their own work (e.g., a novel) into a script; and 4) specs or pitches that are based on existing materials. These cases, together, account for about 55% of the projects set up at the studios. In terms of the time span, I use ideas sold by 2005 to leave enough time to observe the final outcome of a sale by 2009. Finally, not all sales indicate whether they are specs or pitches. Using two complementary sources, *Hollywood Literary Sales* and *Who’s Buying What*, I complete this information for about 70% of the sales. Thus, the final sample for analysis contains 1,834 sales.

The sales data are matched to *IMDb*, which has comprehensive information on a person’s resume in the industry; and *TheNumbers*, which has performance data for released movies. The online appendix describes the data sources and the matching procedures across them.

The data provide a novel setup, in which the sale stage can be easily identified. In addition, the industry’s relatively short life cycle allows us to observe not only whether the project reaches the market, but also its sales. However, it is important to note a number of limitations, and how they constrain the empirical tests. First, I do not observe ideas that are rejected and, hence, cannot test the writer’s choice directly. Such data are hard to find in general. It is comforting that Hypothesis 1 shows that the prediction conditional on being sold is roughly consistent with predictions not conditional on the sale outcome. Second, I do not observe which pitches are terminated after the first draft. As discussed in the previous section, I cannot use the likelihood of release to test how qualities of specs and pitches compare. Hypothesis 3 suggests comparing the performance for movies that are released. The downside is that this dramatically reduces the number of observations and introduces more noise (i.e., post-release uncertainty). Finally, the number of observations at the higher end of the writer’s experience is small, which is also typical for such studies. As a result, for some tests, I resort to certain parametric specifications.

\(^{19}\)The database is obtained from www.donedealpro.com. It is recommended by various industry organizations, including the Writers Guild of America, as a valuable resource to stay up to date on projects developed.
3.1 Empirical Strategy

The empirical method is straightforward: I look at relationships between variables in the data and see whether they are consistent with what the model predicts. Note that, in the end, I am interpreting the data through the lens of the model, which is only one plausible story. After the main results, I will discuss caveats of these interpretations and alternative explanations. The unit of analysis is a sale. The analysis is cross-sectional because 80% of the writers have only one sale.

Hypothesis 1 shows that the likelihood of a spec sale has a non-monotonic relationship with respect to the writer’s experience. The following probit model regresses a dummy indicating whether the sale is a spec or a pitch on WEXP, which is a discrete measure of the writer’s experience in scriptwriting. A group of dummies indicating each value of WEXP is used.

\[
P(SPEC_i = 1) = P(\beta_0 + \sum_{k=1}^{K} \beta_k 1_{\{WEXP_i = k\}} + \beta_X X_i + u_i \geq 0).
\]

(3)

\(X_i\) are controls, including the writer’s non-scriptwriting experience and characteristics of the idea and the buyer; and \(u_i\) captures unobservable factors that might also affect the idea’s sale stage.

Hypothesis 2 predicts how affiliating with a reputable intermediary is associated with the sale stage. Because the number of observations is small at the higher end of WEXP after splitting the sample, I use a quadratic specification and interact it with a dummy indicating the agency affiliation.

\[
P(SPEC_i = 1) = P(\beta_0 + \beta_1 WEXP_i + \beta_2 WEXP_i^2 + \beta_3 \text{BIGAGENT} + \beta_4 \text{BIGAGENT} \cdot WEXP_i + \beta_5 \text{BIGAGENT} \cdot WEXP_i^2 + \beta_X X_i + u_i \geq 0).
\]

(4)

Hypothesis 3 predicts that, conditional on release, the expected performance for both specs and pitches increases with the writer’s experience. However, the rate of increase for specs is faster than that for pitches. I estimate the following linear model for movies that are released:

\[
\text{Performance}_i = \beta_0 + \beta_1 SPEC_i + \beta_2 WEXP_i + \beta_3 SPEC_i \cdot WEXP_i + \beta_4 M_i + u_i,
\]

(5)

where performance is measured by a movie’s box office revenues and its return to investment; and \(M_i\) are control variables that might affect performance.

3.2 Variables

A. Main variables

SPEC is a dummy variable that equals one if the sale is a spec. 54% of the sales are specs.

WEXP measures the writer’s experience: the number of writing credits the writer has obtained in the previous five years for movies that are produced or distributed by a major studio (hereafter, major writing credits).\(^{20}\) The restriction to the previous five years captures the writer’s current

\(^{20}\)I include the following ten studios and their divisions: Walt Disney, Warner Bros., Paramount, Universal, Fox,
Table 1: Descriptive Statistics for Selected Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPEC</td>
<td>Dummy, 1 if sale is a spec</td>
<td>1,834</td>
<td>0.54</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>WEXP</td>
<td>Number of writer’s major writing credits in the previous five years</td>
<td>1,834</td>
<td>0.33</td>
<td>0.73</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>BIGAGENT</td>
<td>Dummy, 1 if writer is affiliated with one of the five biggest agencies</td>
<td>1,834</td>
<td>0.40</td>
<td>0.49</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>WRITER_TENURE</td>
<td>Number of years since writer’s first writing credit</td>
<td>1,834</td>
<td>6.02</td>
<td>8.46</td>
<td>0</td>
<td>60</td>
</tr>
<tr>
<td>WRITER_TV</td>
<td>Dummy, 1 if writer has any major TV writing credit</td>
<td>1,834</td>
<td>0.05</td>
<td>0.23</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>WRITER_DIRECTOR</td>
<td>Dummy, 1 if writer has any major directing credit</td>
<td>1,834</td>
<td>0.07</td>
<td>0.25</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>WRITER_ACTOR</td>
<td>Dummy, 1 if writer has any major acting credit</td>
<td>1,834</td>
<td>0.15</td>
<td>0.36</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>WRITER_PRODUCER</td>
<td>Dummy, 1 if writer has any major producing credit</td>
<td>1,834</td>
<td>0.14</td>
<td>0.35</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>NUM_WRITER</td>
<td>Number of writers in the team</td>
<td>1,834</td>
<td>1.35</td>
<td>0.51</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>ATTACH_STAR</td>
<td>Dummy, 1 if there are stars attached</td>
<td>1,834</td>
<td>0.20</td>
<td>0.40</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ATTACH_DIRECTOR</td>
<td>Dummy, 1 if there is a directo attached</td>
<td>1,834</td>
<td>0.13</td>
<td>0.33</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>RELEASE</td>
<td>Dummy, 1 if the movie is released in theater</td>
<td>1,834</td>
<td>0.12</td>
<td>0.32</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PRICE</td>
<td>Price of two drafts and a polish ($000)</td>
<td>1,010</td>
<td>319.20</td>
<td>345.57</td>
<td>1</td>
<td>5000</td>
</tr>
<tr>
<td>US_BO</td>
<td>U.S. box office revenue in $million</td>
<td>217</td>
<td>45.66</td>
<td>47.72</td>
<td>0.00</td>
<td>242.71</td>
</tr>
<tr>
<td>WORLD_BO</td>
<td>Worldwide box office revenue in $million</td>
<td>176</td>
<td>95.27</td>
<td>108.62</td>
<td>0.30</td>
<td>624.35</td>
</tr>
<tr>
<td>ROI</td>
<td>U.S. box office/production budget</td>
<td>191</td>
<td>1.57</td>
<td>1.44</td>
<td>0.02</td>
<td>8.67</td>
</tr>
<tr>
<td>PROD_BUDGET</td>
<td>Estimated production budget in $million</td>
<td>191</td>
<td>36.83</td>
<td>26.05</td>
<td>1.5</td>
<td>150</td>
</tr>
<tr>
<td>NUM_SCREEN</td>
<td># screens during 1st weekend of the movie’s release</td>
<td>217</td>
<td>2181.14</td>
<td>1054.62</td>
<td>1</td>
<td>3965</td>
</tr>
<tr>
<td>NUM_STAR</td>
<td>Dummy, 1 if there are highly-paid stars</td>
<td>217</td>
<td>0.20</td>
<td>0.40</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>DIRECTOR_EXP</td>
<td>Number of directing credits of the movie’s director</td>
<td>217</td>
<td>4.37</td>
<td>5.15</td>
<td>0</td>
<td>21</td>
</tr>
</tbody>
</table>

Notes: The unit of analysis is a sale. Whenever there are more than one writer, I take the maximum of the writers’ experience as that of the team. Descriptive statistics of dummies for genre, buying studio, creative type, MPAA rating, sale year, release year, and release week are omitted.

status and avoids simply measuring the length of his industry tenure. Restricting movies to those by major studios avoids inflating the number with small-budget independent projects. Lastly, 67% of the sales are from single writers, and the rest are from a team of two or more. Whenever there are more than one writer, I take the maximum of the writers’ experience as that of the team. 79% of the sales are from writers with zero credits, and for the other 21%, the average is 1.52.21

BIGAGENT is a dummy variable indicating whether the writer is affiliated with one of the five biggest literary and talent agencies in Hollywood.22 I focus on the big five mainly because having an intermediary is more or less the norm in Hollywood. In terms of reputation capital and market access, however, the difference between being represented by one of the big five and not is greater than the difference between being represented by any intermediary and not. In the sample, 40% of the sales are through these five agencies. On average, sales from writers with a big agency are significantly more likely to be a pitch than are sales from writers without a big agency (59.9% v.s. 42.9%). As robustness checks, I also define BIGAGENT differently to include medium-sized agencies, and the results do not change much.

I define a movie as released if it is theatrically released in the U.S. and has a positive box

---

21 The raw measure ranges from zero to seven. For writers with more than four credits, I group them at four to obtain a decent number of observations for the top cell.

22 During the time period that I study, the big five were Creative Artists Agency, United Talent Agency, William Morris, International Creative Management, and Endeavor. In 2009, Endeavor and William Morris merged.
office revenue.\textsuperscript{23} 12\% of the sales are released. The rate of release is 15.6\% for specs and 8.4\% for pitches. For movies that are released, US\_BO and WORLD\_BO are, respectively, the movie’s box office revenues in the U.S. and worldwide. I also define ROI (return on investment) as the ratio between U.S. box office revenues and the production budget. The sample averages of these performance measures are not significantly different between specs and pitches.

\textbf{B. Control Variables}

\textit{Other writer characteristics.} This set of variables includes the writer’s tenure in movie writing (the number of years since the writer’s first movie-writing credit); a dummy indicating whether he has written for major TV networks; three dummies indicating whether he has ever obtained a directing, acting (top five listed actors/actresses in a movie, by importance), or producing credit for movies by the major studios. The number of writers in the team is also controlled for.

\textit{Buyer.} Ten dummies indicating that one of the ten major studios is listed as a buyer are included. Studio buyers are different from independent production companies in their capacity, resources, and marketing capabilities. The major studios buy 62.3\% of the ideas.

\textit{Sale characteristics.} Dummies for genre are included, and an idea can have multiple genres.\textsuperscript{24} In addition to capturing market size and competition, the writing of an idea is said to be more critical for some genres (e.g., comedy) than for others (e.g., action). Agents sometimes attach talents to a project to attract buyers, which is called packaging. Two dummies indicating whether there are stars or a director attached are included. Finally, dummies for the year of sale are also included.

\textit{Characteristics of movies that are released.} This set of variables includes production budget, the number of screens in the first weekend of the movie’s release, whether there are highly-paid stars,\textsuperscript{25} the director’s experience, dummies for genre, MPAA rating, year of release, week of release, and creative type.\textsuperscript{26} These variables control for the nature of the movie (e.g., a mainstream high-concept movie versus an art-house movie); the studio’s marketing strategy (e.g., a wide versus a limited release); the size of the market (e.g., holiday versus non-holiday releases); and competition conditions from movies outside the sample (e.g., the year and week of release).

4 Estimation Results

4.1 Sale stage and the writer’s experience

\textsuperscript{23}Because specs are more developed, the release time for a spec is cut by July 2009 and that for a pitch by November 2009. The four-month difference is because a typical pitch contract gives the writer three months to finish the first draft and the buyer two weeks to a month to decide whether to continue. There are six cases in which the movies were released at film festivals, and the box office revenues are recorded as $0. I categorize them as not released.

\textsuperscript{24}There are 22 genres in total; the biggest categories are drama, comedy, action, romantic comedy, and thriller.

\textsuperscript{25}STAR indicates whether the movie is on the list of top 1,000 “Highest Combined Star Gross” defined by TheNumbers. 44 movies are on the list.

\textsuperscript{26}MPAA (Motion Picture Association of America) rating reflects a film’s thematic and content suitability for certain audiences. The ratings are G, PG, PG-13, R, and NC-17, in an increasing order of inappropriateness for a younger audience. Creative type is a unique categorization used by TheNumbers, including contemporary fiction, kids’ fiction, dramatization, factual, fantasy, history fiction, science fiction, and super hero.
Column (1) in Table 2 reports the Probit estimates for the relationship between the likelihood of a spec sale and the writer’s experience. Figure 2 plots the relationship, keeping the control variables at their sample means. The predicted probability of a spec sale is 0.58 for writers with zero major credits in the previous five years; it drops to 0.41 and 0.36 for writers with one and two credits; and it increases back to 0.50 and 0.68 for writers with three and four credits. The drop in the likelihood from zero to one/two credits is economically significant (by 20 percentage points), and statistically significant at the 1% level. The likelihood for writers with four credits almost doubles that for writers with one/two credits, and the difference is different from zero at the 5% level. Column (2) uses a quadratic form of WEXP. The results are qualitatively similar, and statistically stronger with the help of the functional form.

![Figure 2: Predicted Probability of a Spec Sale](image)

Notes: The dots are predicted probabilities of a spec sale, keeping the control variables at their sample means. The dashed lines are the 95% confidence interval. The plot is based on results in Column (1) of Table 2.

These results are consistent with Hypothesis 1, and, in particular, inexperienced writers face difficulties in obtaining the buyer’s attention with earlier-stage ideas. Note that the non-monotonicity comes from the interaction between the buyer’s costly participation and the writer’s incentives. Supposing that it is costless for the buyer to participate (i.e., she is willing to meet every idea), the model suggests a monotone increasing relationship between the likelihood of a spec and the writer’s experience, which is what I find for writers with some major credits.

As for other variables, the writer’s writing tenure also suggests a similar U-shaped pattern. The sale is more likely to be a pitch if there are more than one writer in the team and if the writer is also an experienced actor. Dummies for major studios are jointly significant in that they are more likely than independent production companies to buy pitches. This might be because they are more experienced at evaluating earlier-stage ideas (hence a greater value of their feedback) or have a larger pipeline to fill (hence a lower opportunity cost of meeting). The dummies for GENRE are jointly significant (at the 5% level), and so are the dummies for YEAR_SALE (at the 10% level).

**B. Sale stage and the affiliation with a big agency**

Columns (1) and (2) in Table 2 show that, on average, writers affiliated with a big agency are
## Table 2: Probit Estimates for a Spec Sale (DV = SPEC)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WEXP = 1</td>
<td>-0.166***</td>
<td>-0.215***</td>
<td>-0.085</td>
</tr>
<tr>
<td>WEXP = 2</td>
<td></td>
<td>-0.215***</td>
<td></td>
</tr>
<tr>
<td>WEXP = 3</td>
<td></td>
<td></td>
<td>0.099</td>
</tr>
<tr>
<td>WEXP = 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEXP</td>
<td>-0.235***</td>
<td>-0.345***</td>
<td></td>
</tr>
<tr>
<td>WEXP^2</td>
<td>0.066***</td>
<td>0.101***</td>
<td></td>
</tr>
<tr>
<td>BIGAGENT</td>
<td>-0.118***</td>
<td>-0.118***</td>
<td>-0.146***</td>
</tr>
<tr>
<td>BIGAGENT × WEXP</td>
<td>0.201**</td>
<td>0.201**</td>
<td>-0.059*</td>
</tr>
<tr>
<td>BIGAGENT × WEXP^2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRITER_TENURE</td>
<td>-0.013***</td>
<td>-0.013***</td>
<td>-0.013***</td>
</tr>
<tr>
<td>WRITER_TENURE^2</td>
<td>0.000***</td>
<td>0.000***</td>
<td>0.000***</td>
</tr>
<tr>
<td>WRITER_TV</td>
<td>-0.092</td>
<td>-0.093</td>
<td>-0.094</td>
</tr>
<tr>
<td>WRITER_DIRECTOR</td>
<td>0.035</td>
<td>0.035</td>
<td>0.031</td>
</tr>
<tr>
<td>WRITER_ACTOR</td>
<td>-0.064*</td>
<td>-0.065*</td>
<td>-0.064*</td>
</tr>
<tr>
<td>WRITER_PRODUCER</td>
<td>0.058</td>
<td>0.058</td>
<td>0.057</td>
</tr>
<tr>
<td>NUM_WRITER</td>
<td>-0.055**</td>
<td>-0.055**</td>
<td>-0.054**</td>
</tr>
<tr>
<td>ATTACH_DIRECTOR</td>
<td>0.027</td>
<td>0.026</td>
<td>0.026</td>
</tr>
<tr>
<td>ATTACH_STAR</td>
<td>-0.062*</td>
<td>-0.062</td>
<td>-0.063*</td>
</tr>
<tr>
<td>YEAR_SALE dummies</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>MAJOR_STUDIO dummies</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>GENRE dummies</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Pseudo R^2</td>
<td>0.115</td>
<td>0.115</td>
<td>0.117</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-1119.21</td>
<td>-1119.28</td>
<td>-1116.51</td>
</tr>
<tr>
<td>N</td>
<td>1,834</td>
<td>1,834</td>
<td>1,834</td>
</tr>
</tbody>
</table>

Note: Numbers in parentheses are robust standard errors. Marginal effects are reported. (1) uses dummies for each value of WEXP (i.e., the number of the writer’s major writing credits in the past five years); (2) uses a quadratic specification of WEXP; (3) uses a quadratic specification of WEXP interacted with the dummy variable indicating whether the writer is affiliated with one of five biggest agencies. ***, **, and * are respectively significant levels of 1%, 5%, and 10%.

more likely to sell a pitch. Column (3) reports the relationship by WEXP, which is illustrated in Figure 3. For inexperienced writers, the difference in the likelihood of a spec is 14.6 percentage points and is statistically different from zero at the 1% level. For writers with some major credits,
the difference between these two groups is mostly statistically insignificant.

![Graph showing predicted probabilities of a spec sale for writers affiliated with a big agency, and the triangles are those for writers not affiliated with a big agency. The control variables are kept at their sample means. The dashed lines are the 95% confidence intervals. The plot is based on results in Column (3) of Table 2.](image)

**Figure 3: Predicted Probability of a Spec Sale by Big Agency Affiliation**

These results are consistent with the intermediary’s role of reducing information asymmetry between the writer and the buyer. This role is, intuitively, most effective for relatively inexperienced writers. As argued in Hypothesis 2, the affiliation with a big agency might help overcome the barrier to selling at an earlier stage because of the selection effect (the likely positive sorting between writer quality and agency size) or because of the treatment effect (credibly communicating the value of $\theta$ to the buyer). Cleanly separating these two effects is beyond the scope of the paper. It is also hard to find valid instrumental variables that affect only the writer’s agency affiliation and not the writer’s choice of the sale stage. A more structural approach, such as the matching model used in Sorensen (2007), might be necessary.

**C. Movie performance conditional on release**

Table 3 reports OLS results for movies’ performance (equation (5)). The production budget is not available for all released movies, though the number of screens in the first weekend is a reasonable proxy for the movie’s budget level (the correlation of these two variables is 0.52). Columns (1) and (2) use log(US_BO) as the dependent variable, with the latter using only the subsample for which the production budget is available; Column (3) uses ROI (i.e., U.S. box office/production budget); and Column (4) uses log(WORLD_BO) to measure performance.\(^{27}\)

Across different specifications, a similar pattern emerges: The expected performance of a spec is not different from a pitch to start with, but it increases faster with WEXP and eventually becomes higher. Figure 4 illustrates this using results from Column (1). For writers with one credit or more, specs perform better than pitches by 34, 51, 90 and 129%. The differences are all statistically different from zero at the 5% level.

\(^{27}\)41 movies have zero international box office revenue, perhaps because they were not theatrically released abroad. Including these 41 movies does not change the results much.
Notes: The dots (triangles) are the predicted log(US_BO) for specs (pitches). The control variables are kept at their sample means. The dashed lines are the 95% confidence intervals. The plot is based on results in Column (1) of Table 3.

Figure 4: Predicted Movie Performance Conditional on Release

Table 3: OLS Estimates for Movie Performance after Release

<table>
<thead>
<tr>
<th></th>
<th>(1) log(US_BO)</th>
<th>(2) log(US_BO)</th>
<th>(3) ROI</th>
<th>(4) log(WORLD_BO)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPEC</td>
<td>0.066 (0.178)</td>
<td>0.058 (0.183)</td>
<td>-0.025 (0.292)</td>
<td>0.214 (0.232)</td>
</tr>
<tr>
<td>WEXP</td>
<td>0.005 (0.111)</td>
<td>-0.050 (0.139)</td>
<td>-0.130 (0.170)</td>
<td>0.049 (0.125)</td>
</tr>
<tr>
<td>SPEC × WEXP</td>
<td>0.273** (0.129)</td>
<td>0.316** (0.148)</td>
<td>0.525*** (0.199)</td>
<td>0.304* (0.162)</td>
</tr>
<tr>
<td>log(SCREEN)</td>
<td>0.419*** (0.081)</td>
<td>0.328*** (0.085)</td>
<td>0.290*** (0.072)</td>
<td>0.325*** (0.087)</td>
</tr>
<tr>
<td>log(PRODBUDGET)</td>
<td>0.522*** (0.141)</td>
<td>-0.960*** (0.194)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STAR</td>
<td>0.459** (0.178)</td>
<td>0.273 (0.191)</td>
<td>0.332 (0.289)</td>
<td>0.387* (0.212)</td>
</tr>
<tr>
<td>DIRECTOR_EXP</td>
<td>-0.012 (0.017)</td>
<td>-0.015 (0.016)</td>
<td>-0.001 (0.024)</td>
<td>0.021 (0.017)</td>
</tr>
<tr>
<td>MAJOR_STUDIO</td>
<td>0.432** (0.167)</td>
<td>-0.000 (0.157)</td>
<td>-0.033 (0.241)</td>
<td>0.212 (0.171)</td>
</tr>
<tr>
<td>Constant</td>
<td>1.044 (1.197)</td>
<td>0.553 (1.235)</td>
<td>6.146*** (2.210)</td>
<td>5.119*** (0.792)</td>
</tr>
<tr>
<td>YEARRELEASE dummies</td>
<td>Y Y Y Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEEKRELEASE dummies</td>
<td>Y Y Y Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GENRE dummies</td>
<td>Y Y Y Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CREATIVE_TYPE dummies</td>
<td>Y Y Y Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adj. R²</td>
<td>0.736 (0.577)</td>
<td>0.577 (0.335)</td>
<td>0.335 (5.037)</td>
<td>0.5037 (0.04)</td>
</tr>
<tr>
<td>N</td>
<td>217 191 191 176</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Numbers in parentheses are robust standard errors. 26 movies have no production-budget information, and 41 movies have no international box office revenues. ***, **, and * are, respectively, significant levels of 1%, 5%, and 10%.

Recall from Section 2.2 that given \( w \), there are two differences between specs and pitches that have countervailing effects on their relative performance: The writer-selection effect makes specs
better, while the extra-screening effect makes pitches better. The empirical results that specs perform better than pitches for experienced writers is consistent with Hypothesis 3 in that the writer-selection effect dominates the extra-screening effect for relatively high $w$. The results that specs and pitches are not very different for inexperienced writers is also consistent with Hypothesis 3 because the extra-screening effect counterbalances the writer-selection effect for relatively low $w$.

Three caveats about this set of results. First, the relationship between performance and the sale stage that I try to explain here is about correlation rather than causality. In fact, to interpret these patterns in terms of what happens at the sourcing stage, I implicitly assume that there is little causal effect. In other words, the fact that a project is sourced at an earlier versus a later stage should not, alone, cause different treatments of these projects—e.g., having a different threshold criteria at later stages. In the next section, I discuss possible causal explanations, such as the impact of the buyer’s involvement in the creative process and agency problems.

Second, the number of observations once conditional on release is extremely small for the higher end of WEXP.\(^{28}\) The identification of the increasing difference (i.e., the coefficient of the interaction term being positive) relies mainly on the variation from writers with a lower number of credits and the parametric specification of equation (5), which is not an unreasonable assumption. In order not to stretch the data too much, I also run the regression on two subsamples: movies from writers with zero credits, and movies from writers with some. The results (available in the online appendix) are consistent with the basic results: Performance of specs and pitches is not different for inexperienced writers, and specs generally perform better than pitches for relatively experienced writers.

Third, potential mismeasurement of $w$ implies that the writer-selection effect might be underestimated. For each WEXP, there will be some writers whose true $w$ is low and are forced to spec, and others whose true $w$ is high and choose to spec. The latter is the writer-selection effect by definition. Specs observed in the data are pooling together these two types. The writer-selection effect is likely to be underestimated because the first type of specs are lower-quality. The underestimation would be stronger for lower values of WEXP.

As for other variables, across specifications, the number of screens, whether there are stars in the movie, and the production budget (when it is available) are important predictors of box office revenues. Dummies for release year, release week and creative types are all jointly significant.

\subsection*{D. Robustness checks}

In the basic analysis, I use the number of the writer’s major writing credits in the previous five years as the main measure of the writer’s experience, and use other characteristics of the writer as controls. As robustness checks, I construct alternative measures in two ways.

First, following the count measure used in the basic analysis, I weight a writing credit by the movie’s box office revenues. Specifically, I employ two different weighting methods: One is to weight movies with above-median performance by 1 and movies with below-median performance

\(^{28}\)For example, for writers with three credits, there are seven specs released and three pitches released; and for writers with four credits, seven specs and one pitch are released.
or without any box office information by 0; the other is to weight movies with performance from the first to fourth quantile by 1/4, 1/2, 3/4, and 1, respectively, and to weight movies without any box office information by 0. In columns (2) through (5) of Table 5 and 6, I use the two weighting methods for movies in the previous five years, and for movies in the writer’s entire working history.

Second, I use the Principle Component Analysis (PCA) to reduce multiple variables measuring various aspects of the writer’s experience to a one-dimensional measure. PCA is a statistical technique used to reduce the dimensionality of data. Here, PCA is performed on the following eight variables: WEXP, WRITER_TENURE, WRITER_TV, WRITER_ACTOR, WRITER_DIRECTOR, WRITER_PRODUCER, WRITER_USBO, WRITER_RATING. All but two variables are already explained. WRITER_USBO is the average U.S. box office revenues, in $ millions, for all of the writer’s past movies for which he has a writing credit and that the box office revenue is not missing. For writers without any writing credits and for writers whose movies have no revenue information, I replace the variable with zero. WRITER_RATING is defined similarly using IMDB ratings for the writer’s movies. I then use the eigenvector with the largest eigenvalue to transform the original eight variables into a one-dimensional measure of the writer’s experience, WEXP_pca. Column (6) of Table 5 and 6 report results using WEXP_pca.

Overall, the results using these two alternative approaches are consistent with the basic results: 1) there exhibits a non-monotonic relationship between the likelihood of a spec sale and the writer’s experience, and 2) conditional on release, the performances of specs and pitches exhibit increasing difference.

4.1 Discussion

A. Measurement error in writer experience

It is not likely that WEXP captures all that the buyer observes about the writer, w, before her meeting decision. It is important to understand how the potential measurement error might help us understand the data and possibly bias the results. For example, it helps to explain why writers with zero credits still pitch quite often (42%). Though appearing inexperienced in the data, many of them might be working on projects currently in development, or they might have demonstrated good craftsmanship in previous failed projects, so are still good enough to get their pitches heard.

The part of w that is not captured by WEXP is in the error term. It is not unreasonable to assume that the omitted variable is independent of WEXP, in which case the interpretation of the results is fine. But it would be problematic if the correlation were sufficiently negative for some data ranges. For example, suppose that writers with three or four credits appear to be of highly experienced in the data, while, in fact, the distribution of their true w is inferior to that of writers with one or two credits. Then, we cannot conclude that once writers overcome the barriers to

---

29 The quantiles are −2.21m, $14.6m, and $46.8m, which are based on all movies released by the major and minor studios between 1992 and 2005.
30 Please see “Principal Components Analysis” by Jason Hsu for more details. It is available at http://www.jasonhsu.org/uploads/1/0/0/7/10075125/principal_components_analysis2.pdf.
pitching, they spec more often as they get better.

That a spec’s expected box office performance significantly increases with WEXP already provides some evidence that it is not a bad measure. In the data, we also observe two other outcome variables that are intuitively correlated with \( w \): whether the movie is eventually released and the purchasing price. These variables could also provide evidence addressing this concern.

Price is available for 56% of the sales. Typically, for both specs and pitches, the reported price consists of two parts: “front-end payment,” which is roughly the price for two drafts and a polish; and “credit bonus,” which is paid only if the movie is produced.\(^{31}\) The overall quality of price data is not great, which is why they are not used in the main analysis. Many report only the sum of the two parts and often in rough ranges (e.g., high six figures). For the purpose of analysis, I divide the sum by a typical 1:1.5 proportion between the two parts, and impute a number considered reasonable by practitioners for rough ranges.\(^{32}\) Despite the inaccuracy, the price data provide useful information on the evaluation of an idea.

Table 4 reports the OLS estimates for a price equation, where \( \text{PRICE} \) is defined as the front-end payment; and a linear probability model for \( \text{RELEASE} \). All regressions use the same set of controls as in Column (1) of Table 2, whose coefficients are omitted in the interest of space.

### Table 4: OLS Estimates for Price and Release

<table>
<thead>
<tr>
<th>DV = \log(\text{PRICE})</th>
<th>DV = \text{RELEASE}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WEXP</strong></td>
<td><strong>WEXP</strong></td>
</tr>
<tr>
<td>\text{Specs}</td>
<td>\text{Pitches}</td>
</tr>
<tr>
<td>\text{(1)}</td>
<td>\text{(2)}</td>
</tr>
<tr>
<td>( 0.322^{***} )</td>
<td>( 0.131^{**} )</td>
</tr>
<tr>
<td>( (0.056) )</td>
<td>( (0.062) )</td>
</tr>
<tr>
<td><strong>SPEC</strong></td>
<td><strong>SPEC</strong></td>
</tr>
<tr>
<td>\text{Specs} \times \text{WEXP}</td>
<td>\text{Pitches}</td>
</tr>
<tr>
<td>\text{(4)}</td>
<td>\text{(5)}</td>
</tr>
<tr>
<td>( -0.118^{**} )</td>
<td>( -0.066^{**} )</td>
</tr>
<tr>
<td>( (0.052) )</td>
<td>( (0.017) )</td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td><strong>Constant</strong></td>
</tr>
<tr>
<td>\text{Specs}</td>
<td>\text{Pitches}</td>
</tr>
<tr>
<td>\text{(7)}</td>
<td>\text{(8)}</td>
</tr>
<tr>
<td>( 5.278^{***} )</td>
<td>( 5.350^{***} )</td>
</tr>
<tr>
<td>( (0.176) )</td>
<td>( (0.213) )</td>
</tr>
<tr>
<td><strong>Controls</strong></td>
<td><strong>Controls</strong></td>
</tr>
<tr>
<td>( Y )</td>
<td>( Y )</td>
</tr>
<tr>
<td>\text{Adj } R^2</td>
<td>\text{Adj } R^2</td>
</tr>
<tr>
<td>\text{Specs}</td>
<td>\text{Pitches}</td>
</tr>
<tr>
<td>\text{(10)}</td>
<td>\text{(11)}</td>
</tr>
<tr>
<td>( 0.285 )</td>
<td>( 0.204 )</td>
</tr>
<tr>
<td>( 0.131 )</td>
<td>( 0.082 )</td>
</tr>
<tr>
<td>\text{N}</td>
<td>\text{N}</td>
</tr>
<tr>
<td>\text{Specs}</td>
<td>\text{Pitches}</td>
</tr>
<tr>
<td>\text{(13)}</td>
<td>\text{(14)}</td>
</tr>
<tr>
<td>( 547 )</td>
<td>( 463 )</td>
</tr>
<tr>
<td>( 994 )</td>
<td>( 840 )</td>
</tr>
</tbody>
</table>

*Note: \( \text{PRICE} \) is roughly the payment for two drafts and a polish. \( \text{RELEASE} \) is a dummy variable indicating whether the movie is released by 2009. The regressions on \( \text{RELEASE} \) are simple linear probability models. All columns use the same controls as in Table 2. Numbers in parentheses are robust standard errors. \( ^{***}, ^{**}, \text{and } ^{*} \) are, respectively, significant levels of 1%, 5%, and 10%.

Broadly speaking, we observe the following patterns for both outcome variables. First, for both specs and pitches, the purchasing price increases significantly with WEXP, and so does the likelihood of release. This rejects the concern of a large negative correlation between WEXP and

\(^{31}\)If the original writer(s) share the writing credit with other writer teams, they also share the bonus.

\(^{32}\)See the online appendix for a detailed description of how writers are typically compensated, what price information is available in the data, and the coding mechanism.
the omitted variable.

Second, a pattern similar to that for the movie’s performance emerges. Take the purchasing price, for example. The price of specs starts lower than pitches for writers with zero credits, but it increases faster than that of pitches and is eventually higher. Similarly, we also observe an increasing difference in the likelihood of release between specs and pitches. Note that neither of the two outcome measures is suitable to test Hypothesis 3. The likelihood of release is problematic because specs have a lower degree of uncertainty by definition. Price is problematic because it depends not only on an idea’s value, but also on the writer’s share, which is supposed to be bigger for a spec and possibly varies with WEXP, as well. Nonetheless, it is comforting to find that both measures at least do not reject Hypothesis 3.

B. Alternative explanations

Could the choice of the sale stage be driven mainly by heterogeneous writing costs? Recall that in the model, given \( w \), the choice of the sale stage is driven by the variation in \( \theta \), which is the deviation of this particular idea’s value from \( w \). Consider a simple variation of the model, where given \( w \), ideas vary in how costly it is to write it up, \( c_s \); and where the writing cost does not affect the idea’s value, \( V = w + \epsilon_i + \epsilon_s \). If the writing cost is the main driver, the writer would choose ideas with higher costs to pitch and ideas with lower costs to spec. This is because the higher the cost, the more valuable is the interim feedback and having the buyer share the cost. Then, conditional on release, the expected performance of a pitch is higher than that of a spec for all \( w \). This is because there is only the extra-screening effect that makes pitches better, but no writer-selection effect to counteract it. However, this prediction is not consistent with the data. We can reasonably conclude that heterogeneous writing costs are not the main driver of the choice of the sale stage.

Buyer’s meeting costs versus writer’s pitching costs. In my model, the non-monotonicity in the choice of the sale stage with respect to \( w \) relies critically on the interaction between costly buyer participation and the writer’s incentives. It is possible, however, to construct a simpler single-agent decision problem that would yield similar predictions, in which neither buyer participation nor different IP appropriability levels is necessary. What is needed, instead, is that pitching incurs non-trivial extra costs, and that the pitching costs decrease with \( w \). This would be essentially an information acquisition story, in that interim feedback is valuable but costly to obtain. Then, the writer would choose to sell his best ideas as specs and drop his worst ones. Pitching is preferable only for an intermediate range of ideas, where the value is sufficiently high to justify not giving up, but not high enough that more information about its demand is still worth the extra cost. Because the pitching cost decreases in \( w \), for sufficiently low \( w \), the extra cost of pitching is sufficiently large that going through this intermediate step is not at all preferable.

Empirically, it is hard to differentiate these two stories. From what I understand about the industry, however, I believe that, in explaining the data, costly buyer participation has more bite than the writer’s pitching costs. You rarely hear writers complain about extra preparation they would have to do if they could get a pitch meeting, while many express frustration at not being able to secure a meeting. Because we are comparing between specs and pitches, this extra cost of
pitching is what the writer needs not incur for a spec. This rules out most transaction costs (e.g., search costs) that are likely to be substantial. The only cost that I can think of is the extra effort to come up with a good presentation without a complete script, which is probably not significant enough to generate the substantial percentage of specs (55%). Furthermore, in order to obtain the non-monotonicity, the pitching cost needs to decrease with $w$. It is hard to say whether this is a reasonable assumption because even though experienced writers might take less time to develop a good presentation, their opportunity cost of time is also higher.

**Competing reasons for writers to spec.** There might be competing reasons why a writer wants to spec. For example, it might be easier to generate a bidding war with a spec. The data, however, show that only 4% of the sales are through bidding, and the percentage is not significantly different between specs and pitches. This is consistent with the claim from a number of practitioners that the vast majority of ideas are sold through bilateral negotiations. The results do not change much after controlling for bidding or excluding these cases. Another factor that favors specs is that the likelihood of obtaining a sole credit—thus not having to share the credit bonus—is higher if the writer is guaranteed the authorship for two drafts rather than one. This is likely to be an important consideration, but the likelihood of a movie actually being produced is so slim (especially for inexperienced writers) that people in the industry say that the purchasing price for the various drafts is what the writer really goes after.

The above two considerations both fall into the category of a more favorable appropriability condition with a spec, which is essentially what is modeled in this paper. In the model, I attribute the appropriability advantage of a spec to a specific reason: better protection level for both legal and non-legal reasons. In addition to the benefit of making the arguments concrete, I make the choice also because IP protection is one of the most important considerations coming out of my conversations with writers and agents and is of broad interest in other contexts and to the literature. But it is important to recognize that this is not the only reason.

**Alternative explanations for performance differences between specs and pitches.** I observe that there is an increasing difference in performance between a spec and a pitch and that specs perform significantly better than pitches for experienced writers. This is consistent with the model, in which the writer selects better ideas to spec. There might be alternative reasons why pitches perform worse. For example, the buyer’s early involvement during the scripting process might have a negative impact; or pitches might suffer from more-severe agency problems. For the impact of the buyer’s early involvement, it is possible to argue both ways. For the agency problem, a typical contract allows the buyer to terminate the project or to change writers after each stage. This staging arrangement is similar to what is commonly observed in contracts in the venture capital industry and is designed partly to alleviate agency problems. Despite the staging arrangement and reputation concerns, it is possible that pitches still suffer more from moral hazard because the writer is guaranteed payment for the first draft as long as he delivers a craftsmanlike work.

---

33 Conventionally, with a spec, the writer is typically hired for a revision, and with a pitch, he would be hired for the first draft. After that, the producer might or might not change the writing team.
These causal explanations might suggest that pitches perform worse. However, in order to predict an increasing difference between specs and pitches, the buyer’s influence (or the agency problem) not only must be sufficiently negative, but, more importantly, also must be increasingly negative to a sufficient extent as $w$ increases. The latter condition seems rather implausible.

5 Conclusion

This paper studies the seller’s choice of the sale stage in a market for ideas. In the scriptwriting context, I develop a theoretical framework featuring costly buyer participation and differential IP appropriability levels for ideas developed to different stages. I find support from the data for the following results. First, for earlier-stage ideas, inexperienced sellers face difficulty in even getting their ideas heard, let alone making a sale. Often, they are forced to develop their ideas further independently or to drop them, even though it might be more efficient to try selling them to the downstream buyer at an earlier stage. For relatively experienced sellers who have no access problem, as their experience grows, they prefer to sell at a later stage more often because it allows them to appropriate a greater share of the surplus. Second, the data show evidence that are consistent with the seller’s selection behavior, whereby they select better ideas to spec and worse ideas to pitch. Finally, the access difficulty for selling earlier-stage ideas is caused by partly information asymmetry, which raises a role for the intermediary. The data show that inexperienced sellers affiliated with a more reputable intermediary are significantly more likely to sell at an earlier stage. Such affiliation makes no significant difference in the sale stage for relatively experienced sellers.

An implication from the model is that there might be undesirable consequences from strengthening the IP protection.\(^{34}\) In addition to incentivizing the seller to innovate, stronger IP rights might reduce the buyer’s incentive to participate. The buyer might be more reluctant to participate because her share of the surplus would be reduced. The buyer might also be more reluctant because the risk of potential legal disputes rises since it is now easier for sellers to sue and differentiating idea-stealing from independent creation is intrinsically difficult. Therefore, in settings in which transferring ideas is important for commercialization and in which small inventors and entrepreneurs are critical for the supply of creative ideas, depressing the buyer’s incentive to participate would likely shut many sellers out of the idea market.

Even though the paper focuses on a specific industry, the basic trade-offs faced by writers and studios are also important considerations in many other contexts, such as VC financing, technology licensing, and the formation of research alliances. Lastly, it is important to note that the paper has its limitations. As discussed in the previous section, the data and institutional knowledge could

\(^{34}\) There has been a recent trend of strengthening contract law protections for idea sales (e.g., Miller (2006)). Nine out of eleven U.S. courts of appeals have established that idea-theft claims based on contract law are no longer preempted by federal copyright law, which implies that as long as a contractual relationship is established, abstract ideas underlying a written expression could also be protected. A notable court case is the ruling by the Ninth Circuit court (the federal appellate court governing states including California) on Grosso vs Miramax in September 2004. One of the consequences is that it is easier for writers to sue at the local courts for idea theft on the basis of contract law. Practitioners in law have observed a surge in the number of lawsuits filed afterwards.
help us to rule out a number of alternative explanations, but not all. Future research that could differentiate and/or quantify the relative importance of different forces that drive the sale stage either within a specific industry, or across industries, would be valuable.

References


26


I am not sure for the moment
A. Proof

Lemmas 1 and 2 are useful for the proofs. They describe properties of the expected value of a pitch, $v(w, \theta, \epsilon_i)$ (defined in equation (1)). They are obtained through straightforward differentiation, integrating by parts, and taking limits of functions.

**Lemma 1.**
\[
\frac{\partial v(w, \theta, \epsilon_i)}{\partial w} = \frac{\partial^2 v(w, \theta, \epsilon_i)}{\partial \epsilon_i} > 0; \quad \frac{\partial^2 v(w, \theta, \epsilon_i)}{\partial w^2} = \frac{\partial^2 v(w, \theta, \epsilon_i)}{\partial \epsilon_i^2} = \frac{\partial^2 v(w, \theta, \epsilon_i)}{\partial w \partial \epsilon_i} > 0.
\]

**Lemma 2.**
\[
\lim_{\theta \to \infty} v(w, \theta, \epsilon_i) = \infty \quad \text{and} \quad \lim_{\theta \to -\infty} v(w, \theta, \epsilon_i) = 0. \quad \text{By symmetry, } v(w, \theta, \epsilon_i) \text{ have the same limits with respect to } w \text{ and } \epsilon_i.
\]

**Proof. Proof of Proposition 1**

A. Writer’s strategy at stage 1

When $w \geq \bar{w}$, anticipating being met for both a spec and a pitch, the writer compares his expected payoffs between these two choices. Their difference is defined in equation (2). First,

\[
\frac{\partial \Delta^W(w, \theta)}{\partial \theta} = \lambda_p \int_{-\infty}^{\epsilon^*_i(w, \theta)} \frac{\partial v(w, \theta, \epsilon_i)}{\partial \theta} dF^i(\epsilon_i) + (\lambda_s - \lambda_p)E_{\epsilon_i} \left[ \frac{\partial v(w, \theta, \epsilon_i)}{\partial \theta} \right] > 0,
\]

where $\epsilon^*_i(w, \theta)$ is the solution of $v(w, \theta, \epsilon_i) = c_s$. The inequality holds because both terms are positive by Lemma 1. Second,

\[
\lim_{\theta \to \infty} \Delta^W(w, \theta) = \lambda_p \lim_{\theta \to \infty} F^i(\epsilon^*_i(w, \theta)) \lim_{\theta \to \infty} \int_{-\infty}^{\epsilon^*_i(w, \theta)} (v(w, \theta, \epsilon_i) - c_s) \frac{f_i(\epsilon_i)}{F^i(\epsilon_i(w, \theta))} d\epsilon_i - (1 - \lambda_p)c_s
\]

\[
= 0 - (1 - \lambda_p)c_s + \infty
\]

\[
= \infty.
\]

We can switch limit and expectation for the last term because $v(w, \theta, \epsilon_i)$ is nonnegative and monotone increasing; hence, the monotone convergence theorem applies. $\lim_{\theta \to \infty} F^i(\epsilon^*_i(w, \theta)) = 0$ because

\[
\frac{\partial \epsilon^*_i(w, \theta)}{\partial \theta} = -1.
\]

Lastly,

\[
\lim_{\theta \to -\infty} \Delta^W(w, \theta) = \lambda_p \lim_{\theta \to -\infty} F^i(\epsilon^*_i(w, \theta)) \lim_{\theta \to -\infty} \int_{-\infty}^{\epsilon^*_i(w, \theta)} (v(w, \theta, \epsilon_i) - c_s) \frac{f_i(\epsilon_i)}{F^i(\epsilon_i(w, \theta))} d\epsilon_i - (1 - \lambda_p)c_s
\]

\[
+ (\lambda_s - \lambda_p)E_{\epsilon_i} \left[ \lim_{\theta \to -\infty} v(w, \theta, \epsilon_i) \right]
\]

\[
= -\lambda_p c_s - (1 - \lambda_p)c_s + 0
\]

\[
= -c_s.
\]

We can switch limit and expectation because $v(w, \theta, \epsilon_i)$ is nonnegative and decreasing; hence, the dominated convergence argument applies.

Therefore, given $w$, $\exists$ a unique $r_0(w) \in R$ such that $\Delta^W(w, \theta) \geq 0$ if and only if $\theta \geq r_0(w)$.
When \( w < \bar{w} \), the writer anticipates being met only if he chooses to spec. It can be shown that \( \frac{\partial S_{w, \theta}^{W}}{\partial \theta} > 0 \), \( \lim_{\theta \to \infty} S_{w, \theta}^{W} = \infty \), and \( \lim_{\theta \to -\infty} S_{w, \theta}^{W} = -c_s \). Thus, given \( w \), there exists a unique \( r_s(w) \in R \) such that the writer specs if \( \theta \geq r_s(w) \) and drops the idea otherwise.

By the implicit function theorem and by symmetry, \( r'_0(w) = -\frac{\partial \Delta_{w, r_0(w)}}{\partial \theta} \frac{\partial w}{\partial \theta} = -1 \). Similarly, \( r'_1(w) = -1 \). \( r_0(w) > r_s(w) \) because \( S_{w, r_0(w)}^{W} = P_{w, r_0(w)}^{W} > 0 \).

### B. Buyer’s strategy at stage 2

By Bayes’ rule and the writer’s strategy, the buyer’s posterior of \( \theta \) seeing a spec is

\[
 h(\theta|w, S) = \begin{cases} 
 \frac{g(\theta)}{1 - G(r_0(w))} & \text{for } \theta \in [r_0(w), \infty), \text{ and } 0 \text{ otherwise} \\
 \frac{g(\theta)}{1 - G(r_s(w))} & \text{for } \theta \in [r_s(w), \infty), \text{ and } 0 \text{ otherwise} 
\end{cases}
\]

if \( w \geq \bar{w} \).

The buyer’s posterior of \( \theta \) seeing a pitch is

\[
 h(\theta|w, P) = \frac{g(\theta)}{G(r_0(w))} \quad \text{for } \theta \in (-\infty, r_0(w)), \text{ and } 0 \text{ otherwise.}
\]

\( h(\theta|w, P) \) is consistent with Bayes’ rule when \( w \geq \bar{w} \). When \( w < \bar{w} \), I impose such a belief because choosing to pitch is a probability zero event in the equilibrium.

The buyer always meets the writer for a spec because \( S_{W, r_0(w)}^{B} > 0 \) for all \( w \). When \( w < \bar{w} \),

\[
 S_{W}^{B} = (1 - \lambda_s) \int_{r_s(w)}^{\infty} \mathbb{E}_{\xi_i}[v(w, \theta, \epsilon_i)] \frac{g(\theta)}{1 - G(r_s(w))} d\theta - c_m
\]

The second equality holds because when \( \theta = r_s(w) \), the writer is indifferent between speccing and dropping the idea; that is, \( S_{w, r_s(w)}^{W} = \lambda_s \mathbb{E}_{\xi_i}[v(w, r_s(w), \epsilon_i)] - c_s = 0 \). The last inequality holds under Assumption 2. For \( w \geq \bar{w} \), \( S_{W}^{B} > 0 \) because \( r_0(w) > r_s(w) \) (i.e., the buyer’s posterior belief given a spec is even more favorable).

Given a pitch, the buyer’s expected payoff from meeting is

\[
P_{W}^{B} = \int_{-\infty}^{r_0(w)} P_{w, \theta}^{B} g(\theta) G(r_0(w)) d\theta,
\]

where \( P_{W}^{B}(w, \theta) = (1 - \lambda_p) \mathbb{P}[v(w, \theta, \epsilon_i) \geq c_\epsilon] \mathbb{E}_{\xi_i}[v(w, \theta, \epsilon_i) - c_s|v(w, \theta, \epsilon_i) \geq c_s] - c_m \).

First, \( P_{W}^{B}(w) \) increases with \( w \).

\[
\frac{dP_{W}^{B}(w)}{dw} = \frac{\partial P_{W}^{B}(w)}{\partial w} + \frac{\partial P_{W}^{B}(w)}{\partial r_0(w)} r'_0(w)
\]

The second equality is obtained through operating integration by part first on \( P_{W}^{B}(w) \) before deriving \( \frac{\partial P_{W}^{B}(w)}{\partial r_0(w)} \); and that \( r'_0(w) = -1 \) and \( \frac{\partial P_{W}^{B}(w, \theta)}{\partial \theta} = \frac{\partial P_{W}^{B}(w, \theta)}{\partial w} \). The inequality holds because under Assumption 3, \( \frac{g(\theta)}{G(\theta)} > \frac{g(r_0(w))}{G(r_0(w))} \) for all \( \theta < r_0(w) \).
Second, it is straightforward to show that \( \lim_{w \to -\infty} P_B(w) = -c_m \).

Lastly, \( \lim_{w \to \infty} P_B(w) = P_B(w, r_0(w)) - \lim_{w \to \infty} \int_{-\infty}^{r_0(w)} \frac{\partial P_B^r(w, \theta)}{\partial w} G(\theta) d\theta = P_B(w, r_0(w)) > 0 \). The first equality is obtained through integration by part. The first term after the first equality is a constant because \( r_0'(w) = -1 \) and \( \frac{\partial P_B^r(w, r_0(w))}{\partial w} = \frac{\partial P_B^r(w, r_0(w))}{\partial w} \). The second term is a decreasing sequence and approaches zero. It can be shown that \( P_B(w, r_0(w)) > \frac{1}{\lambda_s} c_s - c_m \), which is positive under Assumption 2.

Thus, \( \exists \) a unique \( \tilde{w} \in R \) such that the buyer meets the writer for a pitch if and only if \( w \geq \tilde{w} \).

**D. Uniqueness**

This equilibrium is unique in the sense of the players’ equilibrium strategies, which is a result of the monotonicity of both players’ payoffs given the other’s strategy and Bayes’ rule. However, any posterior that makes the buyer find it not worthwhile to meet the writer if he chooses to pitch when \( w < \tilde{w} \) (a probability-zero event) sustains the equilibrium.

**Proof. Proof Sketch of Prediction 1**

By Bayes’ rule, the likelihood of a spec conditional on \( w \) and sale is

\[
\mathbb{P}({\text{spec|sale}}, w) = \frac{\mathbb{P}({\text{spec \cap sale|w}})}{\mathbb{P}({\text{spec \cap sale|w}}) + \mathbb{P}({\text{pitch \cap sale|w}})}.
\]

When \( w < \tilde{w} \), \( \mathbb{P}({\text{pitch \cap sale|w}}) = 0 \) implies that \( \mathbb{P}({\text{spec|sale}}, w) = 1 \).

When \( w \geq \tilde{w} \), the following shows that \( \mathbb{P}({\text{spec|sale}}, w) \) is monotone increasing in \( w \) when \( w \) is sufficiently high. Furthermore, \( \lim_{w \to \infty} \mathbb{P}({\text{spec|sale}}, w) = 1 \). Thus, \( \mathbb{P}({\text{spec|sale}}, w) \) is at its minimum for intermediate values of \( w \).

Note that \( \mathbb{P}({\text{spec|sale}}, w) \) is monotone increasing in \( w \) if and only if

\[
\frac{\partial \mathbb{P}({\text{spec \cap sale|w}})/\partial w}{\mathbb{P}({\text{spec \cap sale|w}})} \geq \frac{\partial \mathbb{P}({\text{pitch \cap sale|w}})/\partial w}{\mathbb{P}({\text{pitch \cap sale|w}})}.
\]  \( \text{A1} \)

The LHS of (A1) is always positive because \( \mathbb{P}({\text{spec \cap sale|w}}) = \mathbb{P}({\text{spec|w}}) \mathbb{P}({\text{sale|spec, w}}) \) and both terms increase with \( w \).

It can be shown that the RHS of (A1) is negative when \( w \) is sufficiently high, hence, (A1) holds. To see this, note that \( \mathbb{P}({\text{pitch \cap sale|w}}) = \int_{-\infty}^{r_0(w)} [1 - F^i(e^*_i(w, \theta))] g(\theta) d\theta \). Taking derivatives and collecting terms, we have \( \partial \mathbb{P}({\text{pitch \cap sale|w}})/\partial w < 0 \) if and only if

\[
\frac{g(r_0(w))}{G(r_0(w))} \geq \mathbb{E}_\theta \left[ \left( \frac{f^i(e^*_i(w, \theta))}{1 - F^i(e^*_i(w, \theta))} \right) \left( 1 - \frac{F^i(e^*_i(w, \theta))}{1 - F^i(e^*_i(w, r_0(w)))} \right) | \theta < r_0(w) \right].
\]  \( \text{A2} \)

Under Assumption 3, \( \lim_{w \to \infty} \frac{g(r_0(w))}{G(r_0(w))} = \infty \), and \( \lim_{w \to \infty} \frac{f^i(e^*_i(w, \theta))}{1 - F^i(e^*_i(w, \theta))} = 0 \). In addition, \( \frac{1 - F^i(e^*_i(w, \theta))}{1 - F^i(e^*_i(w, r_0(w)))} < 1 \) for all \( \theta \in (-\infty, r_0(w)) \). Hence, the right-hand side of (A2) approaches zero as \( w \to \infty \). Therefore, \( \partial \mathbb{P}({\text{pitch \cap sale|w}})/\partial w < 0 \) when \( w \) is sufficiently high. \( \Box \)
Proof. Proof Sketch of Prediction 3

Given \( w \) and conditional on release, the expected value of a spec and a pitch are, respectively,

\[
\begin{align*}
\mathbb{E}[V|\text{released}, \text{spec}, w] &= \mathbb{E}_{\theta, \epsilon_i, \epsilon_s}[V|\theta \geq r_0(w), V \geq 0] \\
\mathbb{E}[V|\text{released}, \text{pitch}, w] &= \mathbb{E}_{\theta, \epsilon_i, \epsilon_s}[V|\theta < r_0(w), \epsilon_i \geq \epsilon_i^*(w, \theta), V \geq 0].
\end{align*}
\]

Recall that \( \epsilon_i^*(w, \theta) \) is defined as the solution of \( v(w, \theta, \epsilon_i) = c_s \).

What I mean by the writer-selection effect is that when ignoring the term \( \epsilon_i \geq \epsilon_i^*(w, \theta) \) in \( \mathbb{E}[V|\text{released}, \text{pitch}, w] \), the difference between the two conditional values is caused by the writer’s selection behavior only (that is, we are comparing \( \mathbb{E}_{\theta, \epsilon_i, \epsilon_s}[V|\theta \geq r_0(w), V \geq 0] \) and \( \mathbb{E}_{\theta, \epsilon_i, \epsilon_s}[V|\theta < r_0(w), V \geq 0] \)). Specs are expected to perform better because the conditional joint distribution of \( \theta, \epsilon_i, \epsilon_s \) is more favorable in the sense of first-order stochastic dominance.

The extra-evaluation effect means that when ignoring the selection in \( \theta \) (that is, comparing \( \mathbb{E}_{\theta, \epsilon_i, \epsilon_s}[V|V \geq 0] \) with \( \mathbb{E}_{\theta, \epsilon_i, \epsilon_s}[V|\epsilon_i \geq \epsilon_i^*(w, \theta), V \geq 0] \)), the joint conditional distribution of \( \theta, \epsilon_i, \epsilon_s \) for a pitch is more favorable in the sense of first-order stochastic dominance. But the difference diminishes as \( w \) increases because \( \lim_{w \to \infty} \epsilon_i^*(w, \theta) = -\infty \).

It can be shown that \( \mathbb{E}[V|\text{released}, \text{spec}, w] \) and \( \mathbb{E}[V|\text{released}, \text{pitch}, w] \) both increase with \( w \). However, specs tend to infinity, while pitches approach a finite number. To see this, rewrite \( \mathbb{E}[V|\text{released}, \text{spec}, w] \) as follow,

\[
\mathbb{E}[V|\text{released}, \text{spec}, w] = \int_{r_0(w)}^{\infty} \left\{ \mathbb{E}_{\epsilon_i, \epsilon_s}[V|\epsilon_i + \epsilon_s \geq -w - \theta] \right\} \frac{\mathbb{P}(\epsilon_i + \epsilon_s \geq -w - \theta)g(\theta)d\theta}{\int_{r_0(w)}^{\infty} \mathbb{P}(\epsilon_i + \epsilon_s \geq -w - \theta)g(\theta)d\theta},
\]

which is the expectation of \( \mathbb{E}_{\epsilon_i, \epsilon_s}[V|\epsilon_i + \epsilon_s \geq -w - \theta] \) over \( \theta \in (r_0(w), \infty) \). The integrand tends to \( \infty \) as \( w \to \infty \). Therefore, \( \mathbb{E}[V|\text{released, spec, w}] \) is unbounded. The conditional value of pitches is bounded because for all \( w \),

\[
\mathbb{E}_{\theta, \epsilon_i, \epsilon_s}[V|\theta < r_0(w), \epsilon \geq \epsilon_i^*(w, \theta), V \geq 0] < \mathbb{E}_{\theta, \epsilon_i, \epsilon_s}[V|\theta = r_0(w), \epsilon \geq \epsilon_i^*(w, r_0(w)), V \geq 0],
\]

which is a constant because \( r_0'(w) = -1 \).

Therefore, when \( w \) is sufficiently high, the conditional value of a spec is better; when \( w \) is sufficiently low, it is possible that the conditional value of a pitch is better due to the writer selection effect that is discussed above. \( \square \)
### B. Tables and Figures

Table 5: Robustness: Probit Estimates for a Spec Sale (DV = SPEC)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WEXP</td>
<td>-0.235***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.112)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEXP^2</td>
<td>0.066***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEXP_m</td>
<td>-0.144***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(WEXP_m)^2</td>
<td>0.041**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEXP_q</td>
<td>-0.151*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.080)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(WEXP_q)^2</td>
<td>0.049</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEXP_{all}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.116***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.027)</td>
<td></td>
</tr>
<tr>
<td>(WEXP_{all})^2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.014***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.004)</td>
<td></td>
</tr>
<tr>
<td>WEXP_{all}^2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.120***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.031)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>(WEXP_{all}^2)^2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.016***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td>WEXP_{pca}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.083***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.011)</td>
<td></td>
</tr>
<tr>
<td>WEXP_{pca}^2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.016***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.003)</td>
<td></td>
</tr>
</tbody>
</table>

**Controls** | **Y** | **Y** | **Y** | **Y** | **Y** | **Y** |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pseudo R^2</td>
<td>0.115</td>
<td>0.1069</td>
<td>0.1053</td>
<td>0.1110</td>
<td>0.1097</td>
<td>0.1177</td>
</tr>
<tr>
<td>N</td>
<td>1,834</td>
<td>1,834</td>
<td>1,834</td>
<td>1,834</td>
<td>1,834</td>
<td>1,834</td>
</tr>
</tbody>
</table>

*Note:* Numbers in parentheses are robust standard errors. Marginal effects are reported. (1) reproduces the second column of Table 2; (2) and (3) count movies in the previous five years, but weight them by median and by quantiles, respectively; (4) and (5) count movies in the writer’s whole writing history and weight them by median and by quantiles, respectively. (6) uses WEXP_{pca}, the measure produced by the Principle Component Analysis. All columns use the same set of controls as in Table 2. ***, **, and * are, respectively, significant levels of 1%, 5%, and 10%.
Table 6: Robustness: OLS Estimates for Movie Performance after Release (DV = log(US_BO))

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPEC</td>
<td>0.066</td>
<td>0.063</td>
<td>0.140</td>
<td>0.013</td>
<td>0.035</td>
<td>0.132</td>
</tr>
<tr>
<td></td>
<td>(0.178)</td>
<td>(0.176)</td>
<td>(0.176)</td>
<td>(0.187)</td>
<td>(0.184)</td>
<td>(0.176)</td>
</tr>
<tr>
<td>WEXP</td>
<td>0.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.111)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPEC × WEXP</td>
<td>0.273**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.129)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEXP_m</td>
<td></td>
<td>0.011</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.127)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPEC × WEXP_m</td>
<td></td>
<td>0.353**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.147)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEXP_q</td>
<td></td>
<td>0.120</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.214)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPEC × WEXP_q</td>
<td></td>
<td>0.325</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.241)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEXP_all_m</td>
<td></td>
<td></td>
<td>-0.046</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.056)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPEC × WEXP_all_m</td>
<td></td>
<td></td>
<td>0.228***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.074)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEXP_all_q</td>
<td></td>
<td></td>
<td></td>
<td>-0.041</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.067)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPEC × WEXP_all_q</td>
<td></td>
<td></td>
<td></td>
<td>0.251***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.084)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEXP_pca</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.020</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.052)</td>
<td></td>
</tr>
<tr>
<td>SPEC × WEXP_pca</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.157**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.064)</td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Y</td>
</tr>
<tr>
<td>Adj. R²</td>
<td>0.736</td>
<td>0.739</td>
<td>0.733</td>
<td>0.742</td>
<td>0.742</td>
<td>0.840</td>
</tr>
<tr>
<td>N</td>
<td>217</td>
<td>217</td>
<td>217</td>
<td>217</td>
<td>217</td>
<td>217</td>
</tr>
</tbody>
</table>

Note: Numbers in parentheses are robust standard errors. (1) reproduces the first column of Table 3; (2) and (3) count movies in the previous five years, but weight them by median and by quantiles, respectively; (4) and (5) count movies in the writer’s whole writing history and weight them by median and by quantiles, respectively. (6) uses WEXP_pca, the measure produced by the Principle Component Analysis. All columns use the same set of controls as Table 3. ***, **, and * are, respectively, significant levels of 1%, 5%, and 10%.