WHEN TO SELL YOUR IDEA: THEORY AND EVIDENCE FROM THE MOVIE INDUSTRY*

(Job Market Paper)

Hong Luo

New York University, Stern School of Business

hluo@stern.nyu.edu

April, 2011

ABSTRACT. In order to commercialize their ideas, most entrepreneurs need to cooperate with another party. When there are significant and irreversible investments at stake, how far should an entrepreneur develop his idea before selling it? This question is important for entrepreneurs in a variety of contexts, such as research alliances, technology licensing, and VC financing.

I study this question in the context of the U.S. movie industry, in which the screenwriter decides to sell a storyline versus a complete script. I first build a formal model, in which the writer and the buyer meet to transact an idea. The model incorporates important features of a market for ideas: uncertainty, information asymmetry, expropriation risk, and the heterogeneous observable quality of the seller. I then test the model's predictions on a novel sample of idea sales from Hollywood.

The empirical results confirm the predictions: 1) conditional on sale, the likelihood of a complete script has a non-monotonic relationship with respect to the writer's observable quality; and 2) conditional on release, a movie purchased as a complete script, on average, performs better than a movie purchased as a storyline when the writer's observable quality is relatively high.

The paper highlights the access barriers to a desirable audience faced by a seller of relatively low observable quality, as well as the seller's selection behavior in choosing when to sell. These results have interesting implications for the strategies of both the entrepreneurs and the buyers, as well as for policies on protection for idea/intellectual property transactions.

*The Kauffman Dissertation Fellowship Program generously funded this study. I especially thank my advisors Luís Cabral, Adam Brandenburger, John Asker, and Rob Seamans for their constant guidance, inspiration, and encouragement. I would also like to thank Mariagiovanna Baccara, Heski Bar-Issac, Wilbur Chung, Allan Collard-Wexler, J.P. Eggers, Ignacio Esponda, Ricard Gil, Matt Grennan, Yinghua He, Deepak Hegde, Michael Katz, Nicola Lacetera, Ramana Nanda, Gabriel Natividad, Gonçalo Pacheco-de-Almeida, Peter Rousseau, Xiaoxia Shi, Scott Stern, Marie Thursby, Heidi Williams, Sid Winter, as well as workshop participants at the Roundtable on Engineering Entrepreneurship Research (REER) 2010 at Georgia Tech, the BPS Dissertation Consortium 2010 at AOM Montreal, International Industrial Organization Conference 2010 at UBC, CCC 2010 at U Michigan, Kauffman Entrepreneurship Workshop 2010 at Atlanta, campus visits during the winter of my job market, and Stern's Economics and Management Department. All errors are my own.

1. Introduction

In order to commercialize their ideas, most entrepreneurs need to sell them to another party.[1] Since Arrow (1962), a lot of attention has been given to the problem of how to capture rents of an idea under incomplete information and weak property rights (e.g., Anton and Yao (1994, 2005) among many others). But if we take a step back and consider the actual development process of these ideas, frequently, there are huge uncertainty and significant (and irreversible) investments at stake. This leads to the important question of "How far should an entrepreneur develop his idea before selling it?"

The stage of sale is important for entrepreneurs because it affects how much rents they are able to capture from their ideas; it determines the extent of sunk costs, as well as the severeness of information asymmetry and expropriation risk at the point of sale. Understanding these trade-offs could help us address questions across a range of industries, such as: "At what phase should a biotech firm form a research alliance with a pharmaceutical firm?" "How much of his own capital should an entrepreneur invest in his business idea before pursuing VC financing?" "Should a fashion designer approach the retailer with an idea of a clothing line or a complete line?"

I study this question in a specific context; namely, the market for original movie ideas in Hollywood. Besides being a significant and fascinating market in and of itself,[2] it provides a relatively clean and simple setup to study the question of sale stage. Empirically, the data are reasonably rich, and in particular, the industry's relatively short life cycle allows me to observe the sale's final outcome.

Here, screenwriters (writers, hereafter) sell original movie ideas to movie studios. The stage of sale is equivalent to the writer's choice of selling the storyline and (if sold) getting paid by the buyer to write the script (*pitch*), versus investing effort to write up the idea and sell a complete script (*spec*).[3] Thus, in this context, the relevant research question is: "When should a writer pitch versus spec an idea?"

I first build a formal model, in which a writer and a buyer meet to transact an idea. Interviews and understanding of the industry lead to a particular set of assumptions and the sequential form.

[1] I use "idea" to mean various types of intellectual properties, and use "sale" to mean various forms of transactions, including a research alliance, a licensing agreement, or an outright sale of an idea.

[2] The price of an idea ranges from a few thousand dollars to possibly a few millions; thousands of writers supply ideas to about half of all movies released in the U.S., with the other half being adaptations and sequels. If I am only counting box office sales in the U.S., this is an industry of annual revenues of over ten billion dollars.

[3] The writers' union requires that the buyer hire the writer who sells the pitch to write the first draft. In keeping with the industry convention, I use "spec" and "pitch" as both a verb and a noun in this paper.

I then test the model's predictions using a novel data set consisting of 1,638 original movie ideas sold in Hollywood between 1998 and 2003.

The key elements of the model are as follows. Given an idea, the writer decides whether to pitch, to spec, or to drop it. The writer has an observable quality (e.g., his previous industry experience) and also possesses private information regarding this particular idea's value. The buyer, in turn, decides whether or not to meet the writer, meaning either to listen to his pitch or to read his script. Additional information is realized (to both parties) after the meeting because of the buyer's knowledge and expertise, though the extent of additional information realized depends on whether there is a complete script.[4] Finally, disclosing the idea exposes the writer to the risk of expropriation; however, this risk decreases as the idea is further developed largely because copyright protection is more effective.[5]

The model puts an emphasis on the buyer's participation incentives because meeting a writer is costly. In addition to the physical costs and the opportunity cost of time for evaluating an idea, it is costly also because agreeing to meet exposes the buyer to the potential legal disputes that come with the first-hand knowledge of an idea (e.g., idea-theft or copyright infringement claims).[6] Adding the buyer's meeting decision confronts another challenge an idea seller faces: the access barrier to a desirable audience that determines whether he gets a chance to disclose his idea at all.

From the writer's point of view, there is a trade-off between "pitching" and "specing." Writing a full script is more expensive to the writer than talking about the idea. In this sense, one advantage of pitching is that if the idea turns out not to be that interesting, the writer can save himself the cost of writing the full script.[7] Another advantage of pitching is that the buyer shares the writing cost, instead of the writer bearing it alone. However, pitching implies a greater risk of expropriation, leaving the writer with a smaller share of the idea's surplus. The problem becomes more complex to the extent that the writer's choice signals private information that he possesses.

When would we expect the writer to pitch his idea rather than spec it? I show that writers of relatively low observable quality face barriers to selling pitches (i.e., earlier-stage ideas) because the buyer does not find it worthwhile meeting them. Hence, they either have to write up the complete script or to drop the idea entirely. When the writer's observable quality is good enough

[4] The buyer's knowledge might come from her extensive experience in commercialization, or might simply reflect her idiosyncratic taste, which is unknown to the writer beforehand.

[5] Interviews with writers and an internal lawyer from one of the biggest talent agencies confirm that stronger legal protection is an important motivation to spec.

[6] This is why movie production companies routinely return unsolicited scripts unopened to writers, and toy companies do the same with unsolicited toy designs.

[7] This "option value" of pitching is reflected in a quote from *The Artful Writer*: "It takes about 1-2 weeks to develop a solid pitch. If a studio/producer/company bites, you go into draft knowing you're not wasting your time."

such that the buyer would want to meet him regardless of how complete the idea is, the writer would want to spec if his private signal of the idea's value is above a certain threshold, and to pitch otherwise. Furthermore, the better the writer's observable quality, the more often he would want to spec.

Two predications come out of the model that are testable using the data: First, conditional on sale, the likelihood of a spec is high for writers of low and high observable qualities and low for writers in the middle; Second, when the writer's observable quality is relatively high, conditional on release, a spec is expected to perform better than a pitch.

To empirically test these predictions, I construct a novel sample of 1,638 original movie ideas sold in Hollywood between 1998 and 2003.[8] The sales data come from *Done Deal Pro*, from which I observe the sale stage. Matching the sales data to *IMDb* and *TheNumbers*, I also observe the writer's characteristics and the sale's final outcome.

The data confirm my theoretical predictions. Here, I use the number of the writer's movies released through the major studios in the previous five years to measure the writer's observable quality. First, the predicted likelihood of a spec sale is 0.59 for writers without any major writing credits. It drops to 0.42 and 0.37 for writers with one and two credits, and increases to 0.63 and 0.74 for writers with three and four (and more) credits. Both the decline and the increase are statistically significant at the 5% level. Second, for writers with two or more credits, when looking at movies that are released, other things being equal, a spec, on average, performs at least 60% better than a pitch at the U.S. box office, and takes at least 13 months less time to get to the theater.[9]

Qualitatively, these results are robust to different ways of counting the credits, and partly robust to alternative ways to measure the writer's observable quality. Evidence from other aspects of the data, such as prices paid for these ideas, suggests that these results are robust to the potential bias caused by the writer's characteristics unobservable to the researcher but observable to the buyer. The data also refute that the buyer's negative influence on the quality of a project after pitches are sold explains the worse performance of pitches (conditional on release), and that given the writer's observable quality, his choice is mainly driven by heterogenous writing costs.

It is generally challenging for entrepreneurs to get access to buyers — especially so with earlier-stage ideas — because in the absence of these barriers, the model suggests an increasing likelihood of a spec sale with respect the writer's observable quality. Allowing for institutional differences,

[8]The only work that I am aware of using similar data is a working paper by Ravid et al (2011) on the contractual design of intellectual properties.

[9]For the time to market, because specs are at a more advanced stage than pitches, I add four months to this variable for all specs.

the (differential) access barriers might help explain the decreasing relationship found in other studies, such as Jensen et al. (2003) and Katila and Mang (2003).[10]

These barriers are less severe when the buyer's meeting costs are lower, when the information is more symmetric, and when the buyer's expected payoffs are greater. A few managerial implications follow immediately for inexperienced sellers (or sellers of less impressive track records): developing the idea further might be better than banging on the buyer's door with a half-baked idea; engaging an intermediary is worthwhile when the benefits of selling the idea earlier is big; waiving the confidentiality rights could help lower the barriers since it essentially lowers the buyer's meeting cost.[11] Lastly, stronger IP protection might not have a uniform impact on idea sellers because it decreases the buyer's expected payoffs and as a result, the barriers increase.[12] In particular, it decreases the buyer's expected payoff and as a result, the barriers increase. While benefiting some sellers, the increased barriers would make some sellers worse off since they are now excluded from the market for earlier-stage ideas.

This paper is rooted in the literature of disclosure of private information under weak property rights (the dilemma raised in Arrow (1962)). See, for example, Bhattacharya and Ritter (1983), Anton and Yao (1994, 2002), Baccara and Razin (2004), Hellmann and Perotti (2010) among others. Instead of assuming a developed IP, this paper combines the problem of investment under uncertainty with the problem of selling ideas.[13] Here, the seller uses the idea's development stage to signal (to the buyer) his private information, which affects whether he gets a chance to disclose at all.

A more developed idea is better protected, hence, selling at a later stage gives the seller a bigger share of the idea's surplus. This setup has the flavor of the seller's endogenous choice affecting the buyer's expropriation incentives or valuation (i.e., a different sale environment). A few papers could be interpreted along these lines. Anton and Yao (2005) shows that when a portion of the IP can be protected, a sequential strategy of selling the protected portion before the unprotected

[10]Jensen et al. (2003) find that universities with higher academic rankings for their faculty have a higher proportion of disclosures licensed in the proof-of-concept stage (early stage). They argue that this is because the marginal impact of quality on the probability of success is highest at this stage. Katila and Mang (2003) find that biotech firms with more partnering experiences and higher R&D intensities are associated with earlier collaboration with pharmaceutical companies.

[11]Release-forms are commonly used in the movie industry.

[12]There has been a recent trend of strengthening contract law protections for idea sales across the states. Nine out of the 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 any written expression (such as a book) could be protected as well. A notable court case is the ruling by the Ninth Circuit court on *Gross vs Miramax* in September 2004.

[13]Pindyck (1991) reviews the literature on investment under uncertainty. Roberts and Weitzman (1981) is among the earliest papers that address the value of information gathering in R&D investment decisions.

portion is superior to selling both in a bundle because selling the protected portion first increases the winning bidder's willingness to pay for the unprotected portion. Gallini and Wright (1990) shows that an ex ante license contract creates a contractual reward structure which alters the future expropriation incentives of the licensee. Bhattacharya and Guriev (2004) analyze how the choice of selling with and without patents affects expropriation incentives which, in turn, affects seller payoffs and R&D incentives.

Previous works on idea sales are, largely, not concerned with the buyer's participation incentives. The only exception, as far as I am aware of, is Anton and Yao (2008). They point out that buyers might be reluctant to participate in bidding because of the exposure to expropriation lawsuits, and provide an explanation for the prevalent practice of the seller's waiving confidentiality rights. This paper shares the emphasis on the buyer's participation incentives. It shows that buyers are more reluctant to participate for earlier-stage ideas, and find supporting empirical evidence.

A few other papers study the stage of idea sales in empirical settings.[14] See Allain et al. (2009) for impacts of the downstream market structure on the stage of licensing in the pharmaceutical industry, Jensen et al. (2003) on faculty's choice of the disclosure stage of their inventions to their universities' technology transfer office and its impact on the terms of licensing, Gans et al. (2008) on the impact of resolving the uncertainty over the scope of IP rights on the timing of licensing across a few technology markets, and finally Katila and Mang (2003) for some empirical evidence on relationships between the stage of alliance formation in the biotechnology industry with characteristics of the biotech firms and related institutions.

In contrast to previous works' focus on technology markets, this paper provides some fresh evidence from a market for creative ideas. It also has a different focus: how does the choice of sale stage depend on the entrepreneur's observable experience and quality? Conceptually, the model explicitly models the role of the buyer's information feedback in the R&D investment decisions.[15] Finally, by linking the sales to their outcome, it is able to infer the underlying quality of ideas associated with different sale stages.

Section 2 presents the model and derives the writer's choice in equilibrium. Section 3 derives predictions from the model. Section 4 describes the data and variables. Section 5 presents and Section 6 discusses the empirical results. Section 7 concludes with a discussion of the managerial and policy implications.

[14]More broadly, the stage of idea sales could be generalized to the question of vertical integration in R&D. See, for example, Gans et al. (2002), Gans and Stern (2003), Teece (1986), Arora (1995), Arora et al. (2001).

[15]The value of information gathering has been studied in the literature on R&D investment (Roberts and Weitzman (1981)); the timing of product market entry (Dechenaux et al. (2003)); and the timing of IPO (Maug (2001)).

2. Model

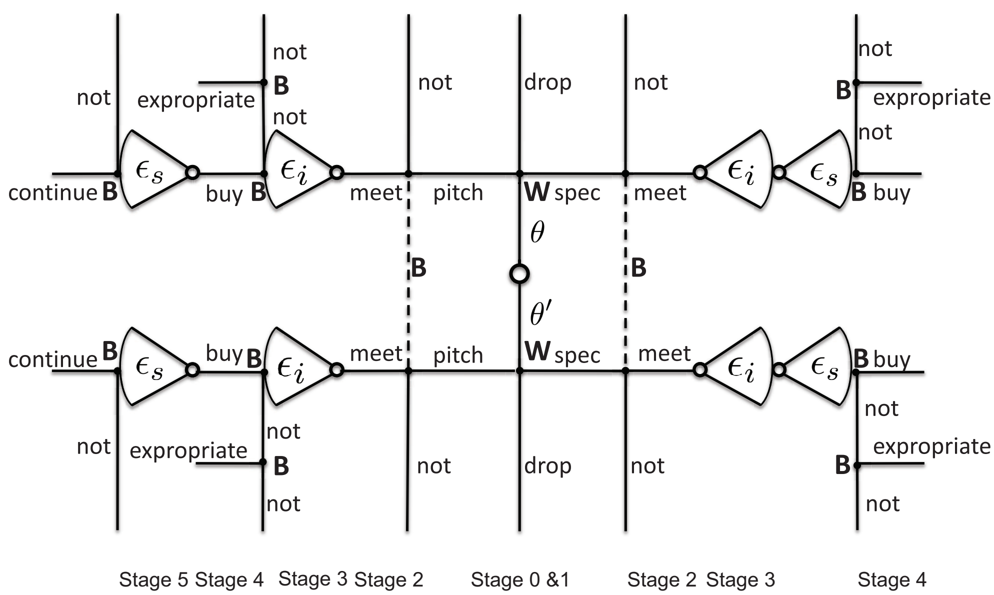
I examine the market exchange of ideas between a writer, W, and a buyer, B. Both players are risk-neutral and maximize expected net payoffs.

An idea's ultimate value, V, is the sum of four parts:

$$V = w + \theta + \epsilon_i + \epsilon_s. \tag{1}$$

Before the writer reveals the idea's details, the buyer has an assessment of the idea's expected value, w. The buyer's assessment is largely based on information she observes about the writer (such as his past experience), therefore, I call it *the writer's observable quality*. This particular idea, however, might be better or worse than average. *The writer's private signal*, $\theta \sim G$, is a signal that the writer observes before deciding what to do. Ex-ante, the idea's value is still uncertain. *The uncertain quality of the abstract idea* is denoted by $\epsilon_i \sim F^i$, and *the uncertain quality of the script* is denoted by $\epsilon_s \sim F^s$. The distributions, G, F^i, and F^s (densities g, f^i, and f^s), are common knowledge. For simplicity, assume that the support of all random variables is $\mathbb{R}$, and they are independent of each other.

FIGURE 1. **Timing**



Notes: Given any writer's observable quality, w, the game starts in the middle, where nature determines the value of the writer's private signal, θ. θ is continuous in the model, though I show only two distinctive values in the graph for illustration purpose. W stands for the writer and B stands for the buyer.

Given w, the game proceeds as follows (illustrated in Figure 1).

At Stage 0, nature determines the value of θ, which is privately observed by the writer.

At Stage 1, the writer decides to spec, to pitch, or to drop the idea given w and θ. If the writer decides to spec (i.e., sell a complete script), he pays the writing cost, $c_s > 0$. If the writer decides to pitch (i.e., sell the storyline), he pays no cost upfront. If the writer drops the idea, the game ends.

At Stage 2, given that the idea is not dropped, the buyer decides whether to meet the writer. "Meeting" means reading the script or listening to the storyline. The buyer observes the writer's quality, w, and the writer's choice of specing versus pitching. Based on what she observes, the buyer forms posterior beliefs of the writer's private signal, θ. The buyer pays the meeting cost, $c_m > 0$, if she decides to meet, and the game ends otherwise.

At Stage 3, the writer discloses the idea during the meeting. The buyer now also observes the writer's private signal, θ. In addition, more information is realized and observed by both parties. For a spec, qualities of both the abstract idea and the script,ϵ_i and ϵ_i, are realized; and for a pitch, only the quality of the abstract idea, ϵ_i, is realized.

At Stage 4, the buyer and the writer bargain over the idea's surplus. I assume a generalized Nash bargaining solution. The writer's and the buyer's bargaining power are, respectively, α and $1 - \alpha$. If the negotiation breaks down, the buyer decides whether to expropriate the idea. For example, she may pass the content of the idea on to another writer and commission a script. If the idea's expected surplus is negative, the buyer's expected net payoff from expropriation is also negative. If the idea's expected surplus is positive, the buyer expects a gross payoff of δ proportion of the idea's surplus. $0 < \delta < 1$ reflects some frictions in expropriation, such as the opportunity cost of the time spent to find another writer. Given expropriation, the writer expects to be compensated with λ_s (λ_p) proportion of the buyer's gross payoff in the case of a spec (a pitch). λ's reflect the strength of the legal protection, such as the probability of the buyer getting caught. Most importantly,

Assumption 1. $\lambda_s > \lambda_p$.

An important reason is that copyright protection is more effective when there is a complete script. First of all, copyright does not protect oral conveyance of an idea. Second, to seek some protection by copyright, writers usually write down one- or two-page outline of the story before they pitch. However, the plots, dialogues, and characters in a complete script are much better defined than in a few-page outline. Thus, in general, the protection of a complete script is more effective.

At Stage 5, in the case of a pitch sale, the writer pays the writing cost, c_s, and the quality of the script, ϵ_s, is realized. The buyer then decides whether to continue with the project.

I make two additional assumptions. First, for a spec, the writer's upfront cost (the cost of writing) is relatively greater than the buyer's upfront cost (the cost of meeting); that is,

Assumption 2. $\frac{c_s}{\alpha(1-\delta)+\lambda_s\delta} > \frac{c_m}{1-(\alpha(1-\delta)+\lambda_s\delta)}$,

where $\alpha(1-\delta)+\lambda_s\delta$ and $1-(\alpha(1-\delta)+\lambda_s\delta)$ are, respectively, the writer's and the buyer's effective shares of the idea's surplus (derived later). Implicitly, the opposite is true for a pitch because the writer pays no upfront cost. This assumption simplifies the analysis greatly, and broadly speaking, it is also consistent with my understanding of the industry.

Second, the distributions of the random variables have the following properties.

Assumption 3. *The probability distribution of θ has 1) a monotone increasing hazard rate (i.e., $\frac{g(\theta)}{1-G(\theta)}$ increases with θ), and 2) a monotone decreasing reverse hazard rate (i.e., $\frac{g(\theta)}{G(\theta)}$ decreases with θ). So do the probability distributions of ϵ_i, ϵ_s, and $\epsilon_i+\epsilon_s$.*

This assumption is standard. Many common probability distributions have such properties, such as normal, logistic, and exponential. See Bagnoli and Bergstrom (2005) for a list of examples.

2.1. **Equilibrium.**

I solve for a perfect Bayesian equilibrium (PBE) for this sequential-move game of incomplete information. The writer's strategy at Stage 1 is to spec, to pitch, or to drop the idea, in anticipation of the buyer's meeting decision. The writer's strategy is a function of his observable quality, w, and his private signal, θ. The buyer's strategy at Stage 2 is whether or not to meet the writer, which is a function of w and the writer's choice at Stage 1. The buyer's beliefs about θ following the writer's choice are consistent with the writer's strategy under Bayes' rule wherever possible.

I focus on PBEs in which the writer's strategy is semiseparating, meaning that given w, some values of θ make the writer choose the same action, while others make him choose different actions.

The following proposition describes the buyer's and the writer's strategies in equilibrium.

Proposition 1. *There exists a unique semiseparating equilibrium, in which the buyer always meets the writer given a spec; and meets the writer if and only if $w \geq \bar{w}$ given a pitch. In the equilibrium,*

(i) *when $w \geq \bar{w}$, the writer chooses to spec if $\theta \geq r_0(w)$, and to pitch otherwise;*

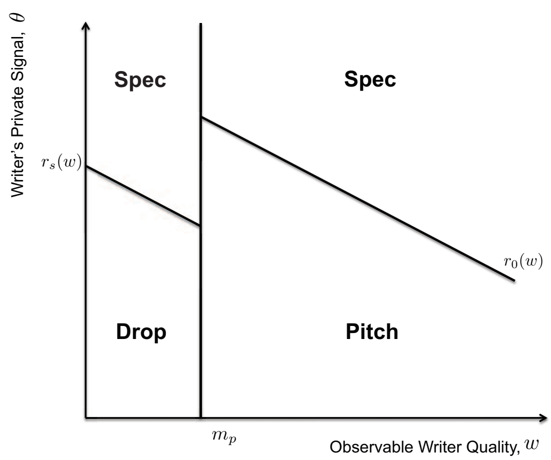
(ii) *when $w < \bar{w}$, the writer chooses to spec if $\theta \geq r_s(w)$, and to drop the idea otherwise.*

Furthermore, $r_0'(w) = r_s'(w) = -1$, and $r_0(w) > r_s(w)$.

Proof. Proof in the Appendix. □

Figure 2 illustrates the writer's choice in equilibrium. Three key features: first, a writer of relatively low observable quality (i.e., $w < \bar{w}$) faces barriers to selling a pitch, with his choices being restricted to specing and dropping the idea; second, when the writer's observable quality is good enough (i.e., $w \geq \bar{w}$), he selects better ideas (i.e., $\theta \geq r_0(w)$) to spec and worse ideas (i.e., $\theta < r_0(w)$) to pitch; and finally, as the writer's observable quality gets better, the threshold separating specs and pitches decreases (i.e., $r_0'(w) < 0$).

FIGURE 2. **Writer's Choice in Equilibrium**



Notes: This graph describes the writer's choice in the semi-separating equilibrium described in Proposition 1. $\bar{w}$ is the buyer's meeting threshold for pitches. When $w < \bar{w}$, the writer is indifferent between specing and dropping the idea when $\theta = r_s(w)$; and when $w \geq \bar{w}$, the writer is indifferent between specing and pitching the idea when $\theta = r_0(w)$.

To understand the equilibrium, I begin by solving the bargaining problem at Stage 4, the results of which enter into the writer's and the buyer's expected payoffs at Stages 1 and 2.

A. ***Bargaining at Stage 4***

In the case of a spec, all uncertainties are realized and observed by both players after meeting, so the idea's surplus is its ultimate value, V. The buyer's outside option is $\max\{(1-\lambda_s)\delta V, 0\}$ because if the negotiation breaks down, she expropriates if and only if $V \geq 0$. The writer's outside option is $\max\{\lambda_s \delta V, 0\}$, which means that he can recoup part of the idea's surplus through legal protection given that expropriation happens.

A spec is sold if and only if $V \geq 0$. Under the generalized Nash bargaining rule, the writer's and the buyer's payoffs given a sale are, respectively,

$$(\alpha(1-\delta) + \lambda_s\delta)V \quad \text{and} \quad (1 - \alpha(1-\delta) - \lambda_s\delta)V.$$

In the case of a pitch, because the quality of the script is yet realized, the idea's expected value, $v(w,\theta,\epsilon_i)$, is the probability that the idea's ultimate value is greater than zero, multiplied by the conditional expected value; that is,

$$v(w,\theta,\epsilon_i) = \mathbb{P}(w+\theta+\epsilon_i+\epsilon_s \geq 0)\mathbb{E}[w+\theta+\epsilon_i+\epsilon_s | w+\theta+\epsilon_i+\epsilon_s \geq 0]. \tag{2}$$

The expected surplus of a pitch is the idea's expected value minus the writing cost to be invested; that is, $v(w,\theta,\epsilon_i) - c_s$.

Similar to the case of a spec, the buyer's and the writer's outside options are, respectively, $\max\{(1-\lambda_p)\delta(v(w,\theta,\epsilon_i)-c_s),0\}$ and $\max\{\lambda_p\delta(v(w,\theta,\epsilon_i)-c_s),0\}$. A pitch is sold if and only if $v(w,\theta,\epsilon_i) \geq c_s$. Given a sale, the writer's and the buyer's payoffs are, respectively,

$$(\alpha(1-\delta)+\lambda_p\delta)(v(w,\theta,\epsilon_i)-c_s) \quad \text{and} \quad (1-(\alpha(1-\delta)+\lambda_p\delta))(v(w,\theta,\epsilon_i)-c_s).$$

Let $\alpha_s = \alpha(1-\delta)+\lambda_s\delta$ and $\alpha_p = \alpha(1-\delta)+\lambda_p\delta$. They are the writer's effective shares of the idea's surplus given a spec and a pitch. The complements, $1-\alpha_s$ and $1-\alpha_p$, are the buyer's shares. Intuitively, the writer's shares increase with his bargaining power, α, and the strength of the legal protection, λ_s (or λ_p). Because α and λ's affect the players' payoffs only through these effective shares, I use α_s and α_p directly hereafter.

Note that $\alpha_s > \alpha_p$ because the strength of protection for a spec is stronger than that for a pitch (i.e., $\lambda_s > \lambda_p$).

B. *Writer's problem at Stage 1*

At Stage 1, given w and θ, the writer chooses to spec, to pitch, or to drop the idea, in anticipation of the buyer's meeting decision at Stage 2. If he drops the idea, he gets zero payoff.

Conditional on being met, the writer's expected payoff from spec, $S^{\mathrm{W}}(w,\theta)$, and his expected payoff from pitch, $P^{\mathrm{W}}(w,\theta)$, are

$$\begin{aligned} S^{\mathrm{W}}(w,\theta) &= \alpha_s\mathbb{P}(V \geq 0)\mathbb{E}[V|V \geq 0] - c_s \\ P^{\mathrm{W}}(w,\theta) &= \alpha_p\mathbb{P}(v(w,\theta,\epsilon_i) \geq c_s)\mathbb{E}[v(w,\theta,\epsilon_i) - c_s | v(w,\theta,\epsilon_i) \geq c_s]. \end{aligned}$$

In words, $S^{\mathrm{W}}(w,\theta)$ is the probability that a spec is sold, $\mathbb{P}(V \geq 0)$, multiplied by the writer's payoff given a sale, $\alpha_s\mathbb{E}[V|V \geq 0]$, and minus the writing cost, c_s. $P^{\mathrm{W}}(w,\theta)$ is the probability that the pitch is sold, $\mathbb{P}(v(w,\theta,\epsilon_i) \geq c_s)$, multiplied by the writer's payoff given a sale, $\alpha_p\mathbb{E}[v(w,\theta,\epsilon_i) - c_s | v(w,\theta,\epsilon_i) \geq c_s]$.

Proposition 1 says that when the writer expects to be met for both a spec and a pitch (i.e., when $w \geq \bar{w}$), he would choose better ideas to spec and worse ideas to pitch. To see this, write the writer's trade-offs between a spec and a pitch, $S^{\mathrm{W}}(w,\theta) - P^{\mathrm{W}}(w,\theta)$, in the following way (notice that $\mathbb{P}(V \geq 0)\mathbb{E}[V|V \geq 0] = \mathbb{E}[v(w,\theta,\epsilon_i)]$):

$$\begin{aligned} \Delta^{\mathrm{W}}(w,\theta) &= \alpha_p \mathbb{P}(v(w,\theta,\epsilon_i) < c_s)\mathbb{E}[v(w,\theta,\epsilon_i) - c_s | v(w,\theta,\epsilon_i) < c_s] \\ &\quad -(1-\alpha_p)c_s + (\alpha_s - \alpha_p)\mathbb{E}[v(w,\theta,\epsilon_i)]. \end{aligned} \tag{3}$$

On the one hand, pitching is desirable for two reasons. First, pitching has an *option value* in the sense that if the idea turns out not to be that interesting, the writer can save himself the cost of writing the full script. The writer's share of this option value is reflected by $\alpha_p \mathbb{P}(v(w,\theta,\epsilon_i) < c_s)\mathbb{E}[v(w,\theta,\epsilon_i) - c_s | v(w,\theta,\epsilon_i) < c_s] < 0$. Second, the buyer is sharing the writing cost with the writer rather than the latter bearing it alone. This *cost advantage* of pitching is reflected by $-(1-\alpha_p)c_s < 0$. On the other hand, specing is desirable because it implies a smaller risk of expropriation, leaving the writer with a greater share of the idea's surplus. This *bargaining advantage* of specing is reflected by $(\alpha_s - \alpha_p)\mathbb{E}[v(w,\theta,\epsilon_i)] > 0$.

Given w, the writer follows a threshold strategy because specing becomes more attractive as the writer's private signal, θ, increases (i.e., $\frac{\partial \Delta^{\mathrm{W}}(w,\theta)}{\partial \theta} > 0$). This is because a better value of θ means a better expected value of the idea, resulting in 1) a bigger incentive to capture a greater share of the surplus, and 2) less need for interim feedback from the buyer.

Furthermore, the threshold decreases as the writer's observable quality increases (i.e., $r_0'(w) < 0$). This is because specing is more attractive for better w, which is true for similar reasons described in the previous paragraph.

When the writer anticipates being met only for a spec (i.e., when $w < \bar{w}$), he chooses to spec if the idea is good enough to justify the writing cost (i.e., $\theta > r_s(w)$).

C. *Buyer's problem at Stage 2*

At Stage 2, the buyer decides whether or not to meet the writer. To obtain a PBE, the buyer's posterior beliefs of the writer's private signal, $H(\theta|\Omega)$ (density h), need to be consistent with Bayes' theorem wherever possible. The buyer's information set Ω includes w the the writer's choice; that is, $\Omega = \{(w,S),(w,P)\}$.

The buyer's expected payoffs from meeting the writer for a spec and a pitch are, respectively,

$$\begin{aligned} S^{\mathrm{B}}(w) &= \int_{-\infty}^{\infty}\{(1-\alpha_s)\mathbb{E}[v(w,\theta,\epsilon_i)] - c_m\}dH(\theta|w,S) \\ P^{\mathrm{B}}(w) &= \int_{-\infty}^{\infty}\{(1-\alpha_p)\mathbb{P}(v(w,\theta,\epsilon_i) \geq c_s)\mathbb{E}[v(w,\theta,\epsilon_i) - c_s | v(w,\theta,\epsilon_i) \geq c_s] - c_m\}dH(\theta|w,P). \end{aligned}$$

Given w and θ, the buyer's expected payoff from meeting a spec is her effective share of the surplus, $(1-\alpha_s)\mathbb{E}[v(w,\theta,\epsilon_i)]$, minus the meeting cost, c_m. $S^{\mathrm{B}}(w)$ integrates the buyer's payoff over θ according to $H(\theta|w,S)$. Similar explanations apply to $P^{\mathrm{B}}(w)$.

Given the writer's strategy described in Proposition 1, facing a spec, the buyer's posterior $h(\theta|w,S)$ is $\frac{g(\theta)}{1-G(r_0(w))}$ for $\theta \in [r_0(w),\infty)$, and 0 elsewhere if $w \geq \bar{w}$; and it is $\frac{g(\theta)}{1-G(r_s(w))}$ for $\theta \in [r_s(w),\infty)$, and 0 elsewhere if $w < \bar{w}$. Under these beliefs, the buyer always wants to meet the writer. The reason is that the writer bears a greater proportion of the upfront costs (Assumption 2), hence, as long as the writer finds it worthwhile to write the script, the buyer should find it worthwhile reading.

Facing a pitch, the buyer's posterior $h(\theta|w,P)$ is $\frac{g(\theta)}{G(r_0(w))}$ for $\theta \in (-\infty, r_0(w))$, and 0 elsewhere for all w. This belief is consistent with Bayes' rule when $w \geq \bar{w}$. When $w < \bar{w}$, because pitching is a probability zero event in equilibrium, I impose the buyer's off-the-equilibrium posterior to be so. Under these beliefs, the buyer's expected payoff from meeting the writer is monotone increasing in w, and the buyer also follows a threshold strategy: meet if and only if $w \geq \bar{w}$. The buyer does not always find it optimal to meet a pitch because, opposite to a spec, it is the buyer who bears a greater proportion of the upfront costs (i.e., it costs the writer nothing to pitch, while is costly for the buyer to meet the writer).

The essential result is that the buyer is stricter about meeting a pitch than meeting a spec. To get intuitions, consider the buyer's trade-offs. A spec is desirable for two reasons: 1) the idea's expected value is better because the writer selects better ideas to spec; and 2) the buyer does not need to share the writing cost. However, a spec puts the buyer in a worse bargaining position because $1-\alpha_s < 1-\alpha_p$. Then, for the threshold writer, $\bar{w}$, the buyer would still be happy to meet the writer for a spec as long as she does not need to give up too much more share of the surplus (i.e., α_s is not too much higher than α_p). Assumption 2 implies that $\alpha_s < \frac{c_s}{c_m+c_s}$, and is a (stronger than necessary) condition to guarantee that this is the case.

2.2. **Discussion.**

The writer's private signal, θ. In the model, the buyer observes the private signal, θ, after the idea is disclosed during the meeting. This might be a more reasonable assumption for a spec than for a pitch because information is more likely to be symmetric after disclosure when there is a complete script. With a pitch, the writer might have an incentive to brag about the idea; or simply find it difficult to articulate the idea without a script.[16]

[16]From *Variety*: "It is the preferred thing to buy scripts," said Amy Pascal, president of Columbia Pictures. "It's so much closer to a movie. Plus, it's on paper. In a script, everyone has the same information." "What happens with the pitch,"

In the Appendix, I discuss a variation of the model, in which the buyer does not observe the writer's private signal, θ, until the script is finished. The equilibrium results are qualitatively similar to Proposition 1. The differences are (1) the buyer's meeting threshold is stricter (i.e., $\bar{w}$ is higher); and (2) for writers who expect to be met for both a spec and a pitch, they spec more (i.e., $r_0(w)$ is lower). The writer chooses to spec more because now he also needs to differentiate his idea from worse ones — a classic adverse selection problem. The reasons for a higher barrier to pitching are two-fold: the average quality of pitches offered for sale decreases as a result of the adverse-selection problem; and there is efficiency loss because the buyer does not observe the true θ when deciding whether to buy a pitch.

Interpretation of λ**.** In the model, choosing to spec gives the writer a better bargaining position because of a stronger legal protection level. In reality, however, there might be multiple reasons that contribute to a better bargaining position when selling an idea at a later stage, and they are likely to coexist. For example, Anton and Yao (1994) argue that the threat of selling the idea to a rival buyer, thus undermining the current buyer's monopoly position, allows the idea seller to capture a sizable share of the surplus even when there are no effective property rights. Applying the intuition to this context, λ_s and λ_p might represent *the degrees of threat* to the current buyer's monopoly position. $\lambda_s > \lambda_p$ means that an idea at a more advanced stage imposes a more immediate threat to the current buyer because a rival buyer is able to bring a spec to the next stage, and eventually to the market, much quicker than a pitch.

Bargaining. The generalized Nash bargaining solution is a great simplification of the actual contract (see the Appendix of Luo (2011) for a detailed description of how writers are compensated). Typically, the writer (and his agent if applicable) first contacts the buyer who is likely to value the idea the most. If the negotiated price is reasonable and competitive, he will sell the idea. An auction is sometimes used, but it is a fairly restricted practice. For example, only four percent of sales in the data have been identified going through some sort of auctioning mechanism.[17]

The qualitative results hold both in the intermediate cases, where the writer and the buyer both have some bargaining power (i.e., $0 < \alpha < 1$), and in the extreme cases, where either one has all of the bargaining power (i.e., $\alpha = 0$ or 1). In related studies, Ueda (2004) gives the entrepreneur all the bargaining power; Anton and Yao (1994 and 2002) give the buyer all the bargaining power; Aghion and Tirole (1994) and Gans et al. (2002) assume a Nash bargaining solution.[18]

Pascal explained, "is that somebody comes in and tells you an idea. To you, that idea means one thing. To them, it means something else. It's a very imprecise science."

[17] 1.2% of the sales indicate "bidding war," and 2.8% percent indicate "preemptive bid."

[18] Gans et al. (2000) develop non-cooperative foundations for this bargaining assumption in the context of a licensing game where the timing of licensing is endogenous.

An idea's ultimate value, V**.** The model assumes that $V = w + \theta + \epsilon_i + \epsilon_s$, and the random variables are independent of each other. Together with Assumption 3, this set-up simplifies the analysis greatly and yields reasonable properties of quantities of interests. In general, however, any set up that yields properties of an idea's expected value, $v(w, \theta, \epsilon_i)$, that are similar to Lemma A1 and A2 in the Appendix would have qualitatively similar equilibrium results. For example, the symmetry in each factor's contribution to V and the independence among the random variables are not critical.

A couple of less obvious points. First, though V takes an additive form, there is actually complementarity among different factors in the players' payoffs. For example, Lemma A1 shows that the cross partial of $v(w, \theta, \epsilon_i)$ with respect to w and θ is positive. This is because the downside risk is eliminated by the option of terminating the project.

Second, the variances of θ, ϵ_i, and ϵ_s might depend on the writer's observable quality, w. For example, suppose that the uncertain quality of the script, ϵ_s, is normally distributed, but its variance decreases with w. It might be argued that the better the writer's past experience the smaller the variance of the script quality. $\sigma'_{\epsilon_s}(w) > -1$ is a sufficient condition to sustain the property that $v(x, \theta, \epsilon_s)$ is an increasing function of w. The intuition is that even though $v(x, \theta, \epsilon_s)$ increases with $\sigma_{\epsilon_s}(w)$ because the upside benefit is greater and the downside risk is eliminated, as long as $\sigma_{\epsilon_s}(w)$ does not decrease with w too fast, $v(x, \theta, \epsilon_s)$ still increases with w.

The writer's and the buyer's costs. Assumptions on the writer's and the buyer's costs and their relative magnitudes can also be relaxed. First, the buyer's meeting cost for a spec might be different from that for a pitch. In fact, under Assumption 2 that the writer's writing cost is relatively greater than the buyer's meeting cost for a spec, the model's results hold qualitatively as long as the buyer's meeting cost for a pitch is positive.[19] Second, Assumption 2 is also stronger than necessary. It can be shown that a more relaxed condition is sufficient for the result that the buyer has a meeting threshold for both a spec and a pitch, but the former is less strict than the latter.

3. Empirical Implications

In this section, I derive empirical implications from the model. These implications are then tested using data from Hollywood.

[19] That said, the buyer's meeting cost for evaluating an earlier-stage idea might not be trivial: they might require higher-skilled staff to evaluate because the information embodied is less concrete than a later-stage idea, and higher-skilled staff's opportunity cost of time is higher.

Ideally, I would like to observe a random sample of ideas *offered for sale* by the writer, both sold and rejected, so that I could directly test the writer's choice. However, such data are hard to find.[20] Therefore, I take particular care when deriving the following predictions, so that they are conditional on sale.

Likelihood of a spec conditional on sale. The model implies a non-monotonic relationship between the likelihood that the sale is a spec and the writer's observable quality; that is:

Prediction 1. *Conditional on sale, the likelihood of a spec is high for writers of low and high observable qualities and low for writers of intermediate observable quality.*

Proof. Proof in the Appendix. □

It is easier to explain the results using Figure 2, which describes the writer's choice in equilibrium. Writers of low observable quality ($w < \bar{w}$) encounter barriers to selling a pitch, so the likelihood of a spec sale for these writers is one. Once a writer's observable quality is good enough ($w \geq \bar{w}$) to get his pitches heard, the likelihood of a spec sale drops. As a writer's observable quality increases, he chooses to spec more often. As a result, the likelihood that the sale is a spec increases again.[21]

Movie performance conditional on release. Suppose that after pitches are transformed into scripts and considered by the buyer as qualified for the next stage, pitches and specs are subject to similar go/no-go criteria as they go through the subsequent development and production processes. Then, the difference in the average quality of these two types of projects come mainly from the early stages.

On the one hand, the model shows that the writer chooses better ideas to spec. *The writer-selection effect*, when considered alone, makes projects purchased as a spec better, on average. On the other hand, projects purchased as a pitch are selected for an extra round by the buyer at the idea stage. *The extra-evaluation effect*, when considered alone, makes projects purchased as a pitch better, on average.[22]

In the data, I observe one such subsequent point where the average qualities of specs and pitches are comparable: when the movie is released. Combining the above two effects, the model predicts:

[20]The only study I am aware of that has both ideas sold and rejected is Kerr et al. (2010), where they have both business ideas financed and rejected from two angel investment groups.

[21]The likelihood of a spec conditional on sale approaches to one as w approaches infinity.

[22]Arguments such as "the producer and the studio contribute positively to the process of transforming the pitch into a script" have effects similar to the extra-evaluation effect, making pitches, on average, better.

Prediction 2. *Conditional on release, when observable writer quality is sufficiently high, the expected performance of movies purchased as a pitch is worse than that of movies purchased as a spec.*

Proof. Proof in the Appendix. □

The extra-evaluation effect is minimal when the writer's observable quality is high enough. This is because their ideas are most likely good enough to be sold if they choose to pitch; thus, the extra round of evaluation does not make much of a difference. The writer-selection effect dominates for these writers. Thus, spec performs better, on average. However, it is possible that pitch performs, on average, better when the writer has low observable quality because the extra-evaluation effect might dominate the writer-selection effect.

4. Data and Variables

The sales data come from an internet database called *Done Deal Pro*, which tracks original movie idea sales and script commissions for adaptations in Hollywood on a daily basis. The database is recommended by various industry organizations, including the Writers Guild of America, as a valuable resource for screenwriters and other industry professionals to stay up to date on what ideas are being developed. The database covers a decent portion of projects purchased by major studios and big production companies. For example, by manually checking all movies released by the major studios in 2008, I am able to track down over 80% of them in the sales database.

The online Appendix has a detailed description of the sales database, the matching procedure across different data sources, as well as the construction of the final sample and the variables. There are around 7,900 entries in the *Done Deal Pro* database between 1998 and early 2008. First, excluding cases where a writer is hired to adapt other people's ideas, such as a book, drops about 55% of the sales. Second, I use ideas sold by December 2003 to leave enough time to observe the final outcome of a project.[23] This leaves about 2,100 sales. Lastly, not all sales clearly indicate whether they are specs or pitches. Complemented by information from two additional sources, *Hollywood Literary Sales* and *Who's Buying What*, I am able to complete this information for 81% of the sales. My final sample for analysis contains 1,638 sales.

To test the model's predictions, I need more information. First, I obtain writer characteristics by matching writers in the sales sample to people on *IMDb*. *IMDb* has comprehensive information on a person's past experiences in the industry. Second, I obtain an idea's final outcome by matching the sales to (released) movies on *IMDb* and *TheNumbers*.

[23] There are 70 months between January 2004 and November 2009. The data show that of ideas sold between January 1998 and December 2003, if the movies are released, 92 percent of them are released within 70 months (the average is 38 months).

I use the data as a cross-sectional sample because about 80% of the writers have only one original idea sale over the six years. The unit of analysis is a sale. Whenever there are multiple players in a player type, I take the maximum of different players' characteristics as that of the player type. For example, for a two-writer team, I define the measure of writer quality as the better of the two. The same rule applies to the buyer and the intermediary.

4.1. **Variables.** Table 1 summarizes the variables, which are defined below.

A. Dependent variables

SPEC equals 1 if the sale is a spec, and 0 if a pitch. 55% of the sample are specs.

RELEASED equals 1 if the movie is theatrically released and generates positive box office revenues in the U.S.[24] The rate of release is 0.12 on average. A spec is significantly more likely to be released than a pitch (0.15 versus 0.09), which is not surprising given its more advanced stage.

For movies that are released, I construct the following two measures of movie performance:

US_BO is the average U.S. box office is $44.4 millions. There is no significant difference in the average U.S. box office between specs and pitches ($43.88 versus $43.53 million).

TIME_TO_MARKET measures the number of months a project takes from sale to market. It is defined as $(\text{release date in the U.S.} - \text{sale date})/30$ if it is a pitch, and $(\text{release date in the U.S.} - \text{sale date})/30 + 4$ if it is a spec. The reason for adding four months is to compensate for the fact that specs are at a more advanced stage of development.[25] I define the date on which the sale is entered into the database as the sale date. This is a reasonable approximation because 1) the database is updated daily; and 2) relevant parties have incentives to announce the sale in a timely measure.[26] The mean of TIME_TO_MARKET is 41.3 months. A one-sided test suggests that a pitch, on average, takes longer to reach the market than a spec (44.38 versus 39.7 months), even after adjusting for the four months' leading time.

B. Independent variables

WRITEREXP is a measure of the writer's observable quality. I use the number of the writer's movies that have been released by the major studios in the previous five years.[27] First, restricting the count to the previous five years avoids simply measuring the length of the writer's industry

[24]There are three cases where the movies are released in film festivals, and the box office revenues are recorded as $0. I count them as not released.

[25]The reason for four months is that a typical writing contract for a pitch requires the writer to finish the first draft in three months, and it allows the buyer about a month to decide whether to continue.

[26]The studio might want to preempt a project, and the writer (and his agent) wants the positive advertising effect.

[27]A writing credit includes *written by*, *screenplay by*, and *story by*. I count movies where the idea is the writer's original, as well as movies where the writer is hired to do the adaptation.

experience and also captures the current status. Second, having a released movie is more convincing than simply selling a spec/pitch or working on a project that has failed. 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. 79% of the sales are from writers with no major writing credit in the previous five years. For the other 21%, the average number of writing credits is 1.52. Later in the empirical section, I discuss alternative ways to count the number, and alternative types of measures for the writer's observable quality.

The following defines a few sets of control variables.

Other writer characteristics: NUM_WRITER is the number of writers on the team. 67.3% of the sales are from single writers, 31% are from two-writer teams, and 1.7% are from teams of three or more. WRITER_ANYMOVIE equals 1 if the writer has ever written any feature film that was then released. WRITER_TV equals 1 if the writer has written for any TV program aired by the major TV networks. WRITER_DIRECTOR, WRITER_ACTOR, and WRITER_PRODUCER indicate whether the writer has ever obtained a directing, acting (top five listed actors in a movie, by importance), or producing credit for movies released by the major studios.

Buyer characteristics: There can be several buyers for each sale. They can be major studios or independent production companies. I create ten dummies to indicate that there is one of the ten major studios among the buyers,[28] and the leftout group would be an average non-studio buyer. 60% of the ideas are purchased by the major studios, and they are more likely to purchase pitches than an average non-studio buyer (53% versus 34%). Producers are also involved in most cases. PRODUCER_EXP measures a producer's experience, which is the number of the producer's movies that have been released by the major studios in the previous five years. I impute the variable as zero for sales where there is no producer listed.

Intermediary characteristics: BIG_AGENT equals 1 if the writer is affiliated with one of the five biggest agencies in Hollywood.[29] I focus on the effects of these biggest agencies because having an agent or other minor type of intermediary is more or less the norm in Hollywood. However, in terms of reputation capital and market power, the difference between being represented by one of these "big five" and not is greater than the difference between being represented by any

[28]Six traditional major studios were Walt Disney, Warner Bros., Paramount, Universal, Twentieth Century-Fox, and Sony Pictures Entertainment (Columbia Pictures). Four smaller ones were DreamWorks SKG, New Line Cinema, MGM, and Miramax Films. There are three cases where there are two major studios involved. I count the first major studio listed as the buyer.

[29]During the sample period that I study, the biggest five agencies were Creative Artists Agency, United Talent Agency, William Morris, International Creative Management, and Endeavor.

intermediary and not.[30] 40% of the sales are from writers affiliated with a big agency. On average, sales from writers with a big agency are significantly more likely to be a pitch (54%) than are sales from writers without a big agency (38%). MANAGER equals 1 if the writer also has a manager. 27% of the sales are from writers that have a manager.

Idea characteristics at the time of sale: Nine dummy variables are used to control for the main GENRE of the idea.[31] The largest genre is comedy (45%), which is followed by drama (21%) and action (11.5%). This is important because ideas of some genres (e.g., comedy) are more execution-dependent than others, so it is more important for the writer to illustrate that he can write well. Genres also matter because they differ in profitability and production budget, which might affect the size of the expected surplus. ATTACH_DIRECTOR and ATTACH_ACTOR indicate whether a director or an actor is already attached to the project at the time of sale. WRITER_DIRECT, WRITER_ACT, and WRITER_PRODUCE indicate whether the writer is going to direct, act in, or produce this particular movie, as well. The year of sale, YEAR_SALE, controls for the general market demand and supply in a particular year. It also affects whether the movie is released within the considered time frame.

Movie characteristics at the time of release: This information is available only for movies that are released. Of the 197 released movies, 149 have information on PROD_BUDGET (in $millon). The mean production budget levels are not significantly different between specs and pitches ($40.3 versus $37.3 millions). NUM_SCREEN is the number of screens (in 100s) in the first weekend of the film's release, and is used to control for the movie's marketing strategy. Pitches have significantly greater number of screens in the first weekend of the film's release than specs (2345.75 versus 1947.95). STAR indicates whether the movie is on the list of top 1,000 "Highest Combined Star Gross" defined by *TheNumbers*.[32] 36 movies are on the list. FRANCHISE indicates whether the movie has a franchised name to bank on. GENRE and MPAA_RATING[33] of the movie affect the potential market size. Dummies of YEAR_RELEASE are also included. Seasonal effects are

[30]From the perspective of size, the total number of unique sales by the five biggest agencies between January 1998 and February 2008 range from 610 to 998, while the highest number from the other 267 agencies is 181. The number of sales is based on the complete sample of 7980 sales that include both adaptations and original ideas dropped due to missing information on the sale stage.

[31]They are action, comedy, drama, horror, thriller, family, adventure, fantasy, and sci-fi. The baseline is other genres.

[32]"Highest Combined Star Gross" is calculated by adding together the total box office for all the credited actors and actresses in the movie, where the total box office for an actor/actress is the sum of the U.S. box office for all the movies he/she has been in.

[33]MPAA rating refers to the film-rating system by the Motion Picture Association of America. It is used in the U.S. and its territories to rate 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 younger audience. For the 198 movies that are released, there is one G and one not-rated. I group them into PG.

controlled for using 51 WEEK_RELEASE dummies, as in Einav (2007). Distributor dummies are also included.[34]

TABLE 1. **Variables and Descriptive Statistics**

Variable	Obs.	Mean	S.D.	Min.	Max.
Sale stage					
SPEC	1638	0.55	0.5	0	1
Outcome					
RELEASED	1638	0.12	0.33	0	1
US_BO ($millions)	197	44.33	48.25	0	242.71
TIME_TO_MARKET (months)	197	41.21	20.21	12.47	124.97
Writer characteristics					
WRITEREXP	1638	0.32	0.72	0	4
NUM_WRITER	1638	1.35	0.51	1	3
WRITER_ANYMOVIE	1638	0.59	0.49	0	1
WRITER_TV	1638	0.04	0.2	0	1
WRITER_DIRECTOR	1638	0.07	0.26	0	1
WRITER_ACTOR	1638	0.15	0.36	0	1
WRITER_PRODUCER	1638	0.13	0.34	0	1
Buyer characteristics					
PRODUCER_EXP	1429	3.67	4.03	0	26
WARNER BROS.	1638	0.09	0.28	0	1
PARAMOUNT	1638	0.05	0.21	0	1
FOX	1638	0.08	0.26	0	1
UNIVERSAL	1638	0.06	0.23	0	1
SONY	1638	0.07	0.26	0	1
DREAMWORKS	1638	0.03	0.17	0	1
NEWLINE	1638	0.08	0.27	0	1
MIRAMAX	1638	0.02	0.15	0	1
MGM	1638	0.04	0.19	0	1
DISNEY	1638	0.09	0.29	0	1
Intermediary characteristics					
BIG_AGENT	1638	0.4	0.49	0	1
MANAGER	1638	0.27	0.45	0	1
Idea characteristics at the time of sale					
WRITER_DIRECT	1638	0.08	0.27	0	1
WRITER_ACT	1638	0.02	0.13	0	1
WRITER_PRODUCE	1638	0.05	0.23	0	1
ATTACH_ACTOR	1638	0.14	0.34	0	1
ATTACH_DIRECTOR	1638	0.21	0.41	0	1
YEAR_SALE	1638	2000.77	1.57	1998	2003
ACTION	1638	0.12	0.32	0	1
ADVENTURE	1638	0.01	0.12	0	1
DRAMA	1638	0.21	0.41	0	1
COMEDY	1638	0.45	0.50	0	1
FAMILY	1638	0.02	0.13	0	1
FANTASY	1638	0.02	0.14	0	1
HORROR	1638	0.02	0.15	0	1
SCI-FI	1638	0.02	0.14	0	1

Continued on Next Page...

[34]The distributor data are from *TheNumbers*. The distributors are in charge of the movie's release strategy. When an idea is purchased by a movie studio at the time of sale, the purchasing studio is likely to be both the producer and the distributor of the movie, while if an idea is purchased by an independent production company, a distributor will be involved later on. There are three cases, in which the movie is purchased by a movie studio but distributed by another. This is mostly likely because the movie is put into turnaround and picked up by the second studio.

Table 1 – Continued

Variable	Obs.	Mean	S.D.	Min.	Max.
THRILLER	1638	0.12	0.32	0	1
OTHER GENRES	1638	0.01	0.10	0	1
Movie characteristics at the time of release					
PROD_BUDGET($millions)	149	38.27	26.52	1.5	150
NUM_SCREEN	197	2077.18	1107.19	1	3965
STAR	197	0.2	0.4	0	1
FRANCHISE	197	0.06	0.23	0	1
YEAR_RELEASE	197	2003.76	2.26	1999	2009
WEAK_RELEASE	197	26.25	15.04	2	52
MPAA_R	197	0.37	0.48	0	1
MPAA_13	197	0.52	0.50	0	1
MPAA_PG	197	0.11	0.31	0	1

Notes: The unit of analysis is a sale. Whenever there are multiple players in a player type, I take the maximum of different players' characteristics as that of the player type.

5. EMPIRICAL ANALYSIS

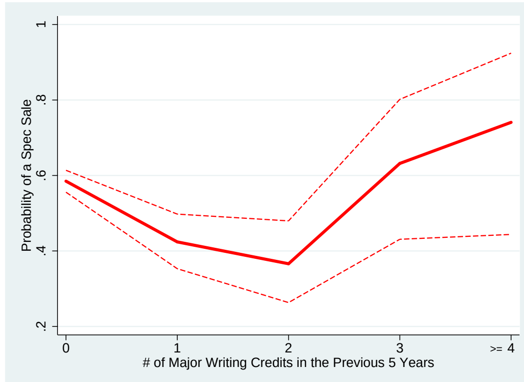
5.1. **Likelihood of a spec conditional on sale.** The model predicts that the likelihood of a spec sale has a non-monotonic relationship with respect to the writer's observable quality. To test this prediction, consider the following Probit model:

$$\mathbb{P}(\text{SPEC}_i = 1) \quad = \quad \mathbb{P}(\beta_0 + \textstyle\sum_{w=1}^{4} \beta_w \mathbb{1}_{[\text{WRITEREXP}_i = w]} + \beta_X X_i + u_i \geq 0), \tag{4}$$

where a group of dummies indicating the number of the writer's major writing credits in the previous five years are used (the group left out are writers without any major writing credits). X_i are control variables, including other writer characteristics, the characteristics of the agency, the buyer, and the idea at the time of sale. $u_i \sim N(0, 1)$ is the unobservable factors that affect whether an idea is sold as a spec verus a pitch.

The Probit estimates are reported in Column (1) of Table 2. Figure 3 plots the predicted probability of a spec sale at different values of WRITEREXP, keeping the control variables at their sample means. The predicted probability is 0.59 for writers without any major writing credits. It drops to 0.42 and 0.37 for writers with one and two credits, and increases to 0.63 and 0.74 for writers with three and four and more credits. Both the decrease and the increase are significantly different from zero at the 5% level.[35]

[35] For example, comparing the predicted likelihood of a spec sale for writers with one and zero major writing credits, $t = \frac{0.42-0.59}{0.0411} = -3.92$. The standard errors are calculated using the delta method. The difference between writers with one and two credits is not significantly different from zero, and neither is the difference between writers with three and four credits.



Notes: The solid line is the predicted probability of a spec sale as a function of the number of the writer's major writing credits in the previous five years. The control variables are kept at their sample means. The dashed lines are the 95% confidence interval. The predicted probabilities are calculated using regression results of equation (4). The regression results are reported in Column (1) of Table 2.

FIGURE 3. **Predicted Probability of a Spec Sale**

Interpreting the results using the theory model, the non-monotonic relationship is a result of the interaction between the writer's and the buyer's incentives. Sales from writers without any major writing credits are more likely to be a spec because the buyer is not willing to listen to their ideas unless they are written up. Writers in the middle want to pitch and are also good enough to get their pitches heard. Sales from top writers are again more likely (than writers in the middle) to be a spec because it is to their interest to spec more often in the first place; by selling at a later stage, the writer is able to capture a greater share of the idea's surplus.

Interestingly, when WRITEREXP = 0, the likelihood of a pitch sale is still quite big (41%). This seems to contradict the discontinuity result from the model (i.e., writers of observable quality lower than the buyer's meeting threshold for pitches cannot sell a pitch at all). An explanation, if not the best, is that even after controlling for many factors, the buyer still has more information about the writer than what is available in the data. Some writers who appear low-quality in the data might be working on projects that are still in the pipeline, or their movies are dropped for reasons unrelated to their capability. This could explain why a fairly large proportion of writers without any major writing experience are still good enough to get their pitches heard. See Section 6 for a detailed discussion on what is observable and what is unobservable from the data.

In terms of the control variables, all else being equal, the sale is less likely to be a spec if the writer has written any feature film that is released, is producing this movie at the same time, is affiliated with a big agency, or if the idea is bought by a less-experienced producer. The dummies for YEAR_SALE are jointly significant. So are the dummies for GENRE and MAJOR_STUDIO.

TABLE 2. **Probit Estimates for the Probability of a Spec Sale**

	(1)	(2)	(3)	(4)
WRITEREXP = 1	-0.092**		-0.093*	-0.072
	(0.044)		(0.055)	(0.046)
WRITEREXP = 2	-0.151**		-0.216***	-0.139**
	(0.062)		(0.078)	(0.065)
WRITEREXP = 3	0.114		-0.028	0.042
	(0.097)		(0.124)	(0.120)
WRITEREXP = 4	0.217*		0.249*	0.273***
	(0.124)		(0.131)	(0.105)
WRITEREXP		-0.171***		
		(0.050)		
$WRITEREXP^2$		0.060***		
		(0.017)		
NUMWRITER	-0.039	-0.041		-0.016
	(0.026)	(0.026)		(0.028)
WRITER_SALE_LASTYR				-0.104***
				(0.024)
WRITER_ANYMOVIE	-0.133***	-0.132***	-0.149***	-0.126***
	(0.030)	(0.030)	(0.036)	(0.031)
WRITER_DIRECTOR	0.070	0.070	0.080	0.074
	(0.056)	(0.056)	(0.071)	(0.058)
WRITER_ACTOR	-0.057	-0.055	0.002	-0.051
	(0.041)	(0.041)	(0.055)	(0.044)
WRITER_PRODUCER	0.042	0.047	0.121**	0.032
	(0.045)	(0.044)	(0.053)	(0.047)
WRITING_TV	-0.090	-0.094	-0.128	-0.117*
	(0.068)	(0.068)	(0.091)	(0.071)
BIG_AGENT	-0.112***	-0.113***	-0.138***	-0.091***
	(0.028)	(0.028)	(0.034)	(0.029)
WRITER_DIRECT	0.011	0.011	0.006	-0.003
	(0.061)	(0.061)	(0.073)	(0.064)
WRITER_ACT	-0.001	-0.010	-0.136	-0.033
	(0.107)	(0.107)	(0.156)	(0.110)
WRITER_PRODUCE	-0.127**	-0.123**	0.006	-0.111*
	(0.061)	(0.060)	(0.078)	(0.063)
ATTACH_ACTOR	-0.020	-0.018	-0.029	-0.025
	(0.041)	(0.040)	(0.048)	(0.042)
ATTACH_DIRECTOR	0.063	0.061	0.065	0.082**
	(0.038)	(0.038)	(0.044)	(0.040)
PRODUCER_EXP	0.015**	0.015**	0.017*	0.014*
	(0.008)	(0.008)	(0.009)	(0.008)
$PRODUCER\_EXP^2$	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
YEAR_SALE F.E.	Y	Y	Y	Y
MAJOR_STUDIO F.E.	Y	Y	Y	Y
GENRE F.E.	Y	Y	Y	Y
N	1638	1638	1100	1472

Note: The unit of analysis is a sale. The dependent variable is SPEC, which is a dummy indicating whether the sale is a spec versus a pitch, for all regressions. The marginal effects of the Probit estimations are reported. (1) and (3) correspond to equation (4), with the former using the complete sample, and the latter using sales from single writers. (2) uses linear and quadratic terms WRITEREXP. (4) uses ideas sold after 1999 so that the number of the writer's original idea sales and writing assignments in the previous year is controlled for. ***, **, and * are respectively significant levels of 1%, 5%, and 10%.

For robustness checks, results in the other columns of Table 2 show that the non-monotonic relationship is statistically significant when a) only sales from single writers are used (Column (3)); b) a linear and a quadratic terms of WRITEREXP are used (Column (2)); and c) when controlling for the writer's number of original idea sales and writing assignments (such as adaptations) in the previous year (Column (4)).[36]

Table 5 shows that the non-monotonic relationship is also robust to different ways of counting the number of credits. The measures are defined the same as WRITEREXP, except that (1) counts the previous six years, (2) counts the previous five years plus the future three years,[37] (3) counts the writer's entire history in the industry, and (4) counts the previous five years and only if the movie also has a U.S. box office greater than $2.5 million, which is the median U.S. box office for movies released between 1993 and 2003.[38]

5.2. **Movie performance conditional on release.** The model predicts that when the writer's observable quality is high enough, conditional on release, the expected performance of a spec is better than that of a pitch.

Before going to the regressions, it is important to understand what I try to accomplish here. First, the purpose of the regression is to detect the writer's selection behavior (i.e., given any observable quality, the writer selects better ideas to spec and worse ideas to pitch when he has the choice). I look for evidence from the movie's performance because it is correlated with an idea's quality; and it needs to be conditional on release because the conditional performance is comparable between specs and pitches. To elaborate on the latter, I cannot compare the probability of release because specs are more developed than pitches, hence, have a smaller degree of uncertainty. Then, results that specs are more likely to be released tell me little about the writer's selection behavior.

Second, equation (5) is about correlation rather than causality because SPEC_i is endogenous. The exercise here is to see whether the data, after controlling for other factors that might affect movie performance, present a pattern that is consistent with what the model predicts. To compare the data and the model (either to confirm it or to reject it), an implicit assumption of "no causal effects" needs to hold: once the first draft of the script is written up for a pitch and considered qualified to proceed to the next step, projects sold as specs and projects sold as pitches are subject to the same selection criteria as they go through subsequent development and production process. This is a reasonable assumption given that the studio's decision-making is largely a function of the project's net expected value.

[36]Regression in Column (4) deleted ideas sold in the first year from the sample (i.e., 1998) so that the writer's sales in the previous year can be computed.

[37]The reason for including the number of writing credits in the future three years is that a movie, on average, takes between three and four years from idea/script to the theater if it is released at all. At the time of sale, the writer might be already writing scripts for projects which are still in progress. I cannot observe these projects in the data, however, the buyer might observe them when she decides whether to meet the writer.

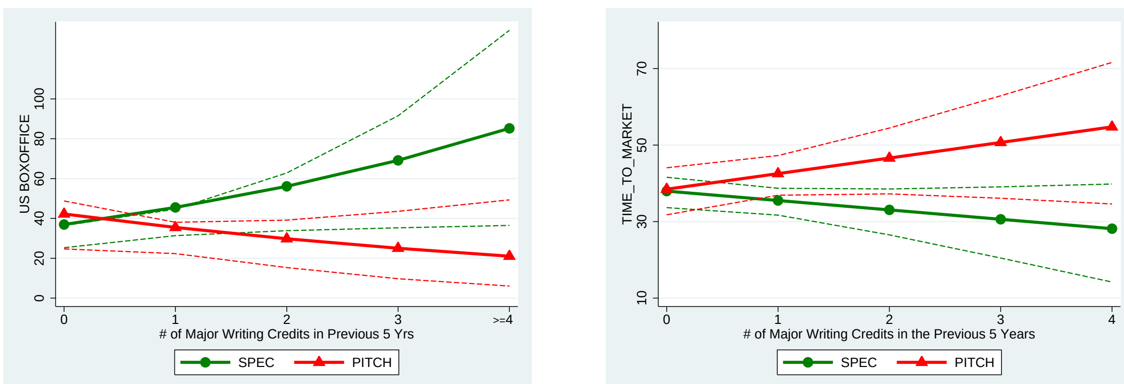
[38]These measures are all censored at the high end, 4, 6, 8, and 4 respectively, to achieve a reasonable number of observations. They use the same set of controls as in Table 2, but their coefficients are omitted to save space.

Finally, due to the small number of observations for movies that are released, I impose a parametric structure in equation (5) rather than using dummies for each cell.

I use two alternative measures for movie performance: US_BO measured in $million, and TIME_TO_MARKET, which is the number of months from sale to release (taking into account of the average four-month lead time of a spec). Consider the following OLS regression for movies that are released:

$$\begin{aligned} \text{Performance}_i &= \beta_0 + \beta_1 \text{SPEC}_i + \beta_2 \text{WRITEREXP}_i \\ &\quad + \beta_3 \text{SPEC}_i \times \text{WRITEREXP}_i + \beta_M M_i + u_i. \end{aligned} \tag{5}$$

M_i are control variables and might differ depending on the measure used. For example, for US_BO, dummies indicating the year and the week of release are used; and for TIME_TO_MARKET, dummies indicating the year of sale are used.



(A) **U.S. Box Office** (B) **Time to Market**

Notes: In (a), the solid lines are the predicted US_BO (transformed from log(US_BO)) for a spec and a pitch; in (b), the solid lines are the predicted TIME_TO_MARKET for a spec and a pitch. The control variables are kept at their sample means. The dashed lines are the 95% confidence intervals. The predictions are based on the regression results of equation (5), where the dependent variables are, respectively, log(US_BO) and TIME_TO_MARKET. See Columns (3) in Table 3 and Column (2) in Table 4 for the regression results.

FIGURE 4. **Predicted Movie Performance Conditional on Release**

Table 3 are results of equation (5) when the dependent variable is log(US_BO). The results and the significant levels are slightly different depending on the specifications, though qualitatively they are similar. Figure 4a plots the predicted U.S. box-office revenues (transformed from log(US_BO)), with the other variables kept at their sample means. The plot is based on results in Column (3), in which the most complete sets of controls are used. When the writer has zero or one major writing credit, the difference is not significantly different from zero. However, specs

TABLE 3. **OLS Estimates for U.S. Box Office**

	(1)	(2)	(3)
SPEC	-0.004	-0.180	-0.134
	(0.155)	(0.173)	(0.214)
WRITEREXP	-0.001	-0.160	-0.174
	(0.120)	(0.146)	(0.157)
SPEC × WRITEREXP	0.228	0.413**	0.383**
	(0.140)	(0.170)	(0.184)
PROD_BUDGET	0.008**	0.006	0.007
	(0.003)	(0.004)	(0.004)
NUM_SCREEN	0.056***	0.059***	0.054***
	(0.009)	(0.013)	(0.016)
STAR	0.329**	0.161	-0.040
	(0.154)	(0.185)	(0.214)
FRANCHISE	1.098***	1.258***	1.164***
	(0.288)	(0.396)	(0.431)
GENRE F.E.	Y	Y	Y
MPAA_RATING F.E.	Y	Y	Y
YEAR_RELEASE F.E.	Y	Y	Y
WEEK_RELEASE F.E.	N	Y	Y
DISTRIBUTOR F.E.	N	N	Y
R-squared	0.573	0.634	0.649
N	149	149	149

Note: The unit of analysis is a released movie. The dependent variable is log(US_BO), where US_BO is the U.S. box office in $millions. The regressions use all observations that are released by November 2009 and have production budget information. ***, **, and * are, respectively, significant levels of 1%, 5%, and 10%.

perform significantly better for writers with two or more credits. For example, when WRITEREXP = 2, spec performs 63.2% better than pitch (significant at the 5% level).[39]

The results using TIME_TO_MARKET (Table 4) are qualitatively similar. Figure 4b looks the opposite to Figure 4a because this measure is negatively correlated with an idea's quality.[40] Take the results in Column (2), for example. When the writer has zero or one major writing credit, the difference is not significantly different from zero; while for writers with two or more credits, a spec is expected to take significantly less time to get to the market. For example, when WRITEREXP = 2, a spec is expected to take 13.5 months less to be released (significant at the 5% level).[41]

These results are consistent with Prediction 2 that when the writer's observable quality is sufficiently high, a spec is expected to perform better than a pitch. Using the theory to interpret the result, this is because *the writer's selection effect* (i.e., choosing better ideas to spec and worse ideas to pitch) dominates *the extra evaluation effect* (i.e., pitches are selected by the buyer for one more

[39]When using log(US_BO), $\text{sd}(\hat{\beta}_1 + 2\hat{\beta}_3) = 0.30$. Thus, $t = \frac{-0.134+2*0.383}{0.30} = \frac{0.632}{0.30} = 2.07$.

[40]The positive correlation between box-office revenue and an idea's quality is relatively straightforward. The negative correlation between TIME_TO_MARKET and an idea's quality is because other things being equal, worse ideas may require more rounds of rewrites and take longer to get commitments from other talents, etc.

[41]When using TIME_TO_MARKET, $\text{sd}(\hat{\beta}_1 + 2\hat{\beta}_3) = 5.894$. Thus, $t = \frac{-0.494-6.527*2}{5.894} = \frac{-13.548}{5.894} = -2.30$.

TABLE 4. **OLS Estimates for Time-to-Market**

	TIME_TO_MARKET (1)	TIME_TO_MARKET (2)	log(TIME_TO_MARKET) (3)
SPEC	-0.115	-0.494	-0.047
	(3.979)	(4.293)	(0.097)
WRITEREXP	3.159	4.072	0.096
	(3.002)	(2.993)	(0.068)
SPEC × WRITEREXP	-5.994*	-6.527*	-0.137*
	(3.442)	(3.532)	(0.080)
PROD_BUDGET	0.455**	0.595**	0.017***
	(0.215)	(0.227)	(0.005)
$PROD\_BUDGET^2$	-0.002	-0.003*	-0.000**
	(0.002)	(0.002)	(0.000)
NUM_SCREEN	-0.014	0.009	-0.000
	(0.227)	(0.249)	(0.006)
STAR	-10.303***	-12.772***	-0.300***
	(3.897)	(3.917)	(0.089)
FRANCHISE	-12.659*	-13.986*	-0.362**
	(6.921)	(7.616)	(0.172)
WRITER_DIRECT	-5.904	-4.044	-0.133
	(5.406)	(5.590)	(0.127)
WRITER_ACT	-0.336	-5.278	-0.202
	(9.438)	(10.003)	(0.226)
WRITER_PRODUCE	-6.031	-3.987	-0.181
	(6.066)	(6.111)	(0.138)
ATTACH_ACTOR	-10.630**	-10.675**	-0.337***
	(4.471)	(4.572)	(0.104)
ATTACH_DIRECTOR	0.824	0.467	0.018
	(4.145)	(4.348)	(0.098)
PRODUCER_EXP	-1.871**	-1.552*	-0.037*
	(0.901)	(0.915)	(0.021)
$PRODUCER\_EXP^2$	0.104*	0.102*	0.002*
	(0.057)	(0.059)	(0.001)
Constant	38.802***	29.417**	3.309***
	(11.114)	(12.110)	(0.274)
GENRE F.E.	Y	Y	Y
MPAA_RATING F.E.	Y	Y	Y
YEAR_SALE F.E.	Y	Y	Y
DISTRIBUTOR F.E.	N	Y	Y
R-squared	0.134	0.210	0.325
N	149	149	149

Note: The unit of analysis is a released movie. The dependent variable is TIME_TO_MARKET for (1) and (2), and log(TIME_TO_MARKET) for (3). TIME_TO_MARKET measures the number of months between the sale date and the release date in the U.S., with four months added to all specs to account for the average leading time of a spec. The regressions use all observations that are released by November 2009 and have production budget information. ***, **, and * are, respectively, significant levels of 1%, 5%, and 10%.

round at the idea stage) when the writer's observable quality is relatively high. Under the assumption of "no causal effects" discussed at the beginning of this section, these results provide evidence for the writer's selection behavior.

A few caveats about the empirical results. First, the strength of the evidence presented here is subject to the small number of observations. In particular, for writers with three credits, there are only two pitches released; and for writers with four credits no pitches are released. Therefore, the predicted performance at the high end of Figure 4a and 4b are extrapolations, which is subject to the validity of the parametric form of equation (5). In order not to stretch the data too much, instead of estimating equation (5), I split the released movies into two groups: those from writers without any credits, and those from writers with some. Table 6 shows that the results are also consistent with the model's predictions. In particular, for writers without any credits, specs, on average, generate less U.S. box office and take similar length of time to get to the market than pitches. For writers with some credits, specs, on average, generate more U.S. box office and take less time to get to the market.

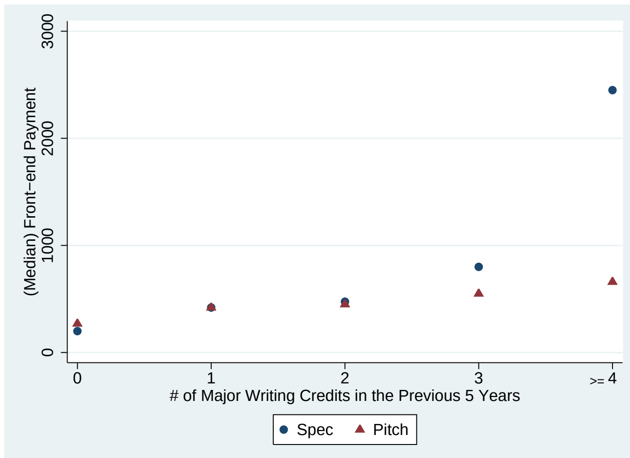
Second, I am likely to underestimate the writer-selection effect. For any value of WRITEREXP, there will be some writers whose true observable quality is below the buyer's meeting threshold for pitches (hence are forced to spec), and others who are good enough to have a choice (and choose better ideas to spec). The latter is the writer-selection effect that I am interested in. Because specs observed in the data are pooling together these two types of ideas, the writer-selection effect is likely to be underestimated because the former type of specs are of lower quality. The underestimation is stronger for lower values of WRITEREXP.

6. DISCUSSION

Writer characteristics observable to the buyer versus to the researcher. Think of the writer's true observable quality, w, in the theory model consisting of two parts: w^o and w^u. w^o is observable to both the buyer and the researcher, such as the number of the writer's major writing credits (i.e., WRITEREXP); and w^u is observable to the buyer but not to the researcher, such as projects that the writer has been working on but yet released. The predictions from the theory model are on w, while in the data, I observe w^o only.

The key question here is "Under what relationship between w^u and w^o does the non-monotonicity result observed in the data reflect a reality different from what the model predicts?" For example, what if the reality is that the likelihood of a spec (conditional on sale) is monotone decreasing in the writer's true observable quality, w? This would be true if the distribution of w^u for writers with three or four (and more) credits (i.e., when w^o is high) is sufficiently dominated by the distribution of w^u for writers with one or two credits. In other words, w^u and w^o are sufficiently negatively correlated for these two groups of writers. Then, the reason why I observe a higher likelihood of a spec when w^o is higher is because the writers are, in fact, of low w.

Evidence from the price data refutes that w^u and w^o are negatively correlated. If they were, the price paid for ideas from writers with three or four (and more) credits should, on average, be lower than that for ideas from writers with one or two credits. Figure 5 tabulates the median of the front-end payment (roughly the price for two drafts and a polish) for specs and pitches.[42] It shows that for specs at least, the median price paid for ideas from writers with three or four (and more) credits is statistically greater than that for ideas from writers with one or two credits.[43] This is also true for mean prices. For pitches, the difference in medians is not statistically significant perhaps due to the small number of observations. A one-sided test for the difference in means is statistically significant at the 5% level.[44]



Notes: This figure tabulates the median front-end payment for specs (dots) and pitches (triangles). For both, the front-end payment is roughly the price for the first two drafts and a polish. 54% of specs and 60% of pitches have price information. See the Appendix for what pricing information is available in the data and the coding mechanism.

FIGURE 5. **Price**

Another possible reality is that the likelihood of a spec (conditional on sale) is monotone increasing in the writer's true observable quality, w. By similar argument, this would be true if w^u and w^o were sufficiently negatively correlated for writers without any credits and for writers with one or two credits. The price data also refutes this possibility. For both specs and pitches, the median (and mean) price paid for ideas sold by writers with one or two credits is statistically greater than that for ideas from writers without any credits.

Alternative measures for the writer's observable quality. The main measure that I use for the writer's observable quality is the number of the writer's movies that have been released by the major studios in the previous five years. There are alternative ways to construct the measure. Define WRITER_BO as the average U.S. box office (in $millions) of the writer's movies that

[42]About 57% of the sales have pricing information. See the online Appendix for detailed description of how writers are compensated, what price information is available, and the coding mechanism.

[43]For specs, Pearson $\chi^2(1)$ test for the difference in medians of these two groups is 14.8 (p-value is 0.0000). t-statistics for the difference in means is 3.90 (p = 0.0001 for a one-sided test).

[44]For pitches, Pearson $\chi^2(1)$ test for the difference in medians of two groups is 2.39 (p-value is 0.122). t-statistics for the difference in means is 1.75 (p = 0.04 for a one-sided test).

had been released by the major studios before the sale; and WRITER_RATING as the average *IMDb* user ratings of these movies. For the 431 ideas that are sold by writers with some major credits (so that these two measures are available), WRITER_BO is missing for 97 sales, and WRITER_RATING is available for almost all sales. The mean of WRITER_BO is $43.67 million, and that of WRITER_RATING is 5.71 (the possible range is from 0 to 10).

Table 10 are Probit estimates for the probability of a spec sale using these alternative measures. Note that they use only the sample of writers that have some major writing credits. The results show that for these writers, the likelihood of a spec sale does not change with the average U.S. box office revenues of the writer's previous movies, while increases significantly (at the 10% level) with the average *IMDb* user ratings for the writer's previous movies.[45]

For writers with some major experience, these results are partly consistent with the results obtained by using the number of credits. Nonetheless, I argue that the number of credits might be the best among these three types of measures. Conceptually for two reasons: First, the pure count makes a consistent measure for all writers because there is no rating nor box office information available for those who haven't written any released movies; Second, the pure count might be less noisy because it is hard to control for factors that are critical to the ratings and the box office performance, such as the production budget, other talents, and genre.

Two pieces of evidence from the data also corroborate with this judgement. First, Table 9 provides correlations between these measures and the sale's outcome.[46] Among these three measures, WRITEREXP has the highest correlations with both RELEASED and log(PRICE). In particular, the other two measures have zero correlations with the probability of release. Second, the clear increasing relationship between price and the number of credits (Figure 5) confirms that it is a reasonably good measure,[47] while price patterns using the other two measures are less clean (Figure 6).[48]

An alternative explanation for the worse performance for pitches. An alternative explanation for the worse performance for pitches is that the buyer's involvement during the process of transforming an idea into a script has a negative impact on the project's value. Though it might be argued that the buyer's involvement could contribute positively, it is more convincing to look for evidence in the data.

[45] The regressions use the same set of controls as in Table 2, but their coefficients are omitted to save space.

[46] For comparison, these correlations are computed using the 333 observations that all three measures are available, and 185 of them have price information.

[47] The significant and increasing relationship between price and WRITEREXP remains true after controlling for other factors.

[48] The lack of clean (increasing) patterns in the price data relative to the other two measures might be partly explained by the small number of observations in each cell.

What is unstated in Prediction 2 is that the difference in the movie's (conditional on release) performance between specs and pitches also increases with w. Intuitively, this is true because when the writer selects better ideas (i.e., $\theta \geq r_0(w)$) to spec and worse ideas (i.e., $\theta < r_0(w)$) to pitch, the average quality of pitches has an upper bound; while the average quality of specs does not. Results using both the U.S. box office and the time to market are cosistent with this prediction.

Now consider the argument that the worse performance of pitches in not caused by the writer's selection behavior, but by the negative influence from the buyer's involvement. More formally, this is to say that the posterior distribution of the writer's private signal, θ, is $G(\theta)$ for both specs and pitches (i.e., the writer randomly selects ideas to spec and to pitch), but the distribution of the quality of the script, ϵ_s, is worse for pitches than for specs. For example, ϵ_s could have a lower mean for pitches. This would only create the pattern in the data if the buyer's influence is 1) sufficiently negative to overcome the extra-evaluation effect that makes pitches, on average, better, and 2) increasingly negative to a sufficient extent as the writer's observable quality gets better to obtain the increasing difference. These conditions seem to be rather implausible.

Heterogenous writing costs. In the model, given w, the writer's choice is driven by the variation in θ, which represents the deviation of this particular idea's value from an average idea from the writer. Is it possible that the writer's choice is driven mainly by some other factors?

One possible variation might be that the writing cost is different for different ideas. A simple variation of the model shows that the writer tends to choose ideas with higher cost to pitch and ideas with lower costs to spec.[49] The intuition is that when the writing cost is big, pitching is particularly valuable because of its option value and cost advantage. The data refutes this model because one of its predictions is that conditional on release, the average performance of pitches is higher than that of specs regardless of the writer's observable quality. The reason is that there is only extra-evaluation effect that makes pitches better, but no writer-selection effect that counteracts it. This is opposite to what the data suggest.

7. Concluding Remarks

This paper studies the question: "At what stage of development should an entrepreneur sell his idea?" In the context of Hollywood, I model how the seller and the buyer interact in idea sales. Consistent with the model's predictions, I find evidence that relatively unknown sellers face barriers to selling early-stage ideas; and that when he has the choice, the seller tends to choose better ideas to sell at a later stage.

[49]See this variation of the model in the Appendix of Luo (2011).

The market for original movie ideas is, in and of itself, an important and fascinating market — the price of an idea ranges from a few thousand dollars to possibly a few million; thousands of aspiring and experienced writers supply ideas to an industry that grosses over ten billion dollars annually,[50] and, most importantly, the market is active despite all the typical frictions in idea sales. However, the basic trade-offs that writers and studios face are, by no means, specific to this market alone. They are fundamental considerations in many other markets for ideas, where uncertainty is high; where information asymmetry and expropriation risk might hinder efficient transfers of ideas; where established firms might contribute critical information about an idea's value; and where idea sellers are heterogeneous in their observable experience and qualities. Interesting examples include book publishing, software programming, and biotech industries, as well as the financing of entrepreneurial start-ups.

The stage of sales is an important question both for entrepreneurs who want to commercialize their ideas, and for downstream firms or investors that acquire these ideas. Understanding the trade-offs the seller and the buyer face is also a first-step towards addressing policy questions to correct inefficient stage of idea transfers or exclusion of good ideas. A few managerial and policy implications follow from the main results of the paper.

The result on the access barriers faced by a relatively unknown seller (or a seller with mediocre track record) when trying to sell an early-stage idea implies a few actions they can take. First, instead of wasting time and effort trying to obtain access to a buyer with only a half-baked idea, the seller might be better off developing the idea further. Second, engaging an intermediary is worthwhile if the benefits of selling the idea earlier outweigh the commission. It should be noted, however, that both these paths are subject to the seller's liquidity/time constraint and the access barriers to a reputable intermediary.

Comparative statics results on the access barriers also have a few policy implications. First, policies that limit the buyer's risk of potential legal disputes (e.g., claims of idea theft or IP infringement) might be important to lower these barriers. An example is a clear and enforceable release-form policy. Second, strengthening legal protection levels might benefit some sellers, while hurt others. The intuition is that the buyer's expected payoff from meeting the seller decreases when the protection level increases, and as a result, the barriers increase. An interesting case is the recent trend of strengthening contract law protection for idea sales across the states.[51] While benefiting

[50] I am counting only ticket sales in the U.S. The data are from *TheNumbers*.

[51] Recent rulings by nine out of the 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 means that abstract ideas underlying any written expression (such as a book) are protected as well, as long as a contractual relationship is established. The most prominent court case is the ruling by the Ninth Circuit court on Grosso v.s. Miramax in September 2004.

sellers that have either very good or very low observable qualities, the increased barriers to selling early-stage ideas would make intermediate writers worse off because they are now excluded from the early market.[52] Lastly, institutions that help reduce information asymmetry are critical for a market where small entrepreneurs are critical in supplying ideas. These institutions can be market intermediaries, or government and industry associations.[53]

This paper finds evidence that the seller has an incentive to develop better ideas further on his own. This has important implications for the buyer: First, it is important to take into account of the differences in the underlying qualities, in addition to the uncertainty levels, when evaluating ideas offered for sale at different stages; Second, moving first to solicit ideas from the seller might not make a difference because the seller still has the incentive to pitch mediocre ideas early, and to keep the best ideas to themselves.

The model also suggests that for sellers at some intermediate levels, the buyer prefers that the seller offers a more complete idea. This implies that the buyer might be better off if she can commit not to listen to early-stage ideas. An interesting anecdote about the scriptwriting world is from an article in *Variety* that quotes a sign outside producers David Zucker's and Jerry Zucker's office at Columbia Pictures: "Thank you for not pitching us your idea." Of course, the credibility of such a commitment might be limited especially when there is competition from rival buyers.

References

Aghion, P. and J. Tirole (1994). The management of innovation. *The Quarterly Journal of Economics 109*(4), 1185–1209.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics 84*, 488–500.

Allain, M., E. Henry, and M. Kyle (2009). The timing of licensing: Theory and empirics.

Anton, J. J. and D. A. Yao (1994). Expropriation and inventions: Appropriable rents in the absence of property rights. *The American Economic Review 84*(1), 190–209.

Anton, J. J. and D. A. Yao (1995). Start-ups, spin-offs, and internal projects. *Journal of Law, Economics, & Organization 11*, 362–378.

Anton, J. J. and D. A. Yao (2002). The sale of ideas: Strategic disclosure, property rights, and contracting. *Review of Economic Studies 69*(3), 513–531.

[52] See Proposition 5.3 in Luo (2011) for comparative statics results.

[53] For example, CORDIS provides information on technologies coming from E.U.-funded R&D programmes, as well as other online services that facilitate licensing, R&D alliances, and manufacturing and marketing agreements.

Anton, J. J. and D. A. Yao (2005). Markets for partially-contractible knowledge: Bootstrapping vs. bundling. *Journal of the European Economic Association 3:2-3*, 745–754.

Anton, J. J. and D. A. Yao (2008). Attracting skeptical buyers: Negotiating for intellectual property rights. *International Economic Review 49*, 319–348.

Arora, A. (1995). Licensing tacit knowledge: Intellectual property rights and the market for know-how. *Economics of Innovation and New Technology 4*, 41 – 60.

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

Arrow, K. (1962). Economic welfare and the allocation of resources for invention. In R. Nelson (Ed.), *The Rate and Direction of Inventive Activity: Economic and Social Factors*. Princeton University Press.

Baccara, M. and R. Razin (2004). From thought to practice: Appropriation and endogenous market structure with imperfect intellectual property rights. CEPR Discussion Papers 4419, C.E.P.R. Discussion Papers.

Bagnoli, M. and T. Bergstrom (2005). Log-concave probability and its applications. *Economic Theory 26*(2), 445–469.

Bhattacharya, S. and S. Guriev (2004). Knowledge disclosure, patents, and optimal organization of research and development. Working paper, Princeton University.

Bhattacharya, S. and J. R. Ritter (1983). Innovation and communication: Signalling with partial disclosure. *Review of Economic Studies 50*, 331–346.

Dechenaux, E., B. D. Goldfarb, M. C. Thursby, and S. A. Shane (2003). Appropriability and the timing of innovation: Evidence from mit inventions. NBER Working Paper Series.

Einav, L. (2007). Seasonality in the u.s. motion picture industry. *Rand Journal of Economics 38*(1), 127–145.

Gallini, N. T. and B. D. Wright (1990). Technology transfer under asymmetric information. *RAND Journal of Economic 21*, 147–160.

Gans, J. S., D. H. Hsu, and S. Stern (2000). When does start-up innovation spur the gale of creative destruction? NBER Working Papers 7851, National Bureau of Economic Research, Inc.

Gans, J. S., D. H. Hsu, and S. Stern (2002). When does start-up innovation spur the gale of creative destruction? *RAND Journal of Economics 33*(4), 571–586.

Gans, J. S., D. H. Hsu, and S. Stern (2008). The impact of uncertain intellectual property rights on the market for ideas: Evidence from patent grant delays. *Management Science 54*, 982–997. NBER Working Paper.

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

Gans, J. S. and S. Stern (2010, ne). Is there a market for ideas? *Industrial and Corporate Change 19*(3), 805–837.

Gompers, P. A. (1995). Optimal investment, monitoring, and the staging of venture capital. *The Journal of Finance 50*(5), 1461–1489.

Gompers, P. A. and J. Lerner (1998). The determinants of corporate venture capital successes: Organizational structure, incentives, and complementarities. NBER Working Papers 6725, National Bureau of Economic Research, Inc.

Greene, W. (2010). Testing hypotheses about interaction terms in nonlinear models. *Economic Letters 107*, 291–296.

Hellmann, T. and E. Perotti (2010). The circulation of ideas in firms and markets.

Hsu, D. H. (2004). What do entrepreneurs pay for venture capital affiliation? *Journal of Finance 59*, 1805–1844.

Jensen, R. A., J. G. Thursby, and M. C. Thursby (2003). The disclosure and licensing of university inventions. *International Journal of Industrial Organization 21*(9), 1271–1300.

Katila, R. and P. Y. Mang (2003). Exploiting technological opportunities: the timing of collaboration. *Research Policy 32*, 317–332.

Kerr, W. R., J. Lerner, and A. Schoar (2010). The consequences of entrepreneurial finance: A regression discontinuity analysis. National Bureau of Economic Research Working Paper No. 15831 and Harvard Business School Working Paper No. 10-086.

Lerner, J. and R. P. Merges (1998). The control of strategic alliances: An empirical analysis of biotechnology collaborations. *Journal of Industrial Organization 46*(2), 125–156.

Luo, H. (2011). Markets for ideas: Theory and evidence from the movie industry. Ph.D. dissertation, New York University.

Maug, E. (2001). Ownership structure and the life-cycle of the firm: A theory of the decision to go public. *European Finance Review 5*, 167–200.

Pindyck, R. S. (1991). Irreversibility, uncertainty, and investment. *Journal of Economic Literature XXIX*, 1110–1148.

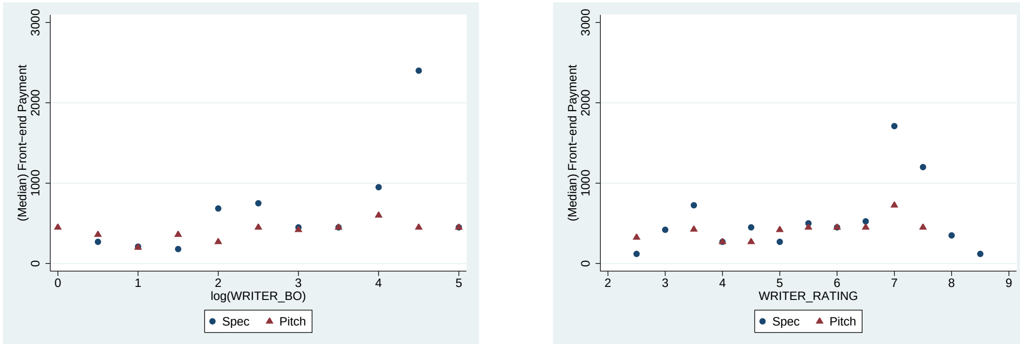
Roberts, K. and M. L. Weitzman (1981). Funding criteria for research, development, and exploratoin projects. *Econometrica 49*(5), 1261–1288.

Teece, D. J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy 15*(6), 285–305.

Ueda, M. (2004). Banks versus venture capital: Project evaluation, screening, and expropriation. *Journal of Finance 59*, 601–621. CEPR Discussion Papers 3411, C.E.P.R. Discussion Papers.

Wooldridge, J. M. (2002). *Econometrics Analysis of Cross Section and Panel Data*. The MIT Press.

FIGURE 6. **Price by Sale Stage (alternative measures of writer's observable quality)**



(A) log(WRITER_BO) (B) WRITER_RATING

Notes: Both log(WRITER_BO) and WRITER_RATING are grouped by increments of 0.5 log point. In addition, log(WRITER_BO) is censored from below at zero to avoid single observation cells.

TABLE 5. **Probit Estimates for the Probability of a Spec Sale (alternative counts of the writer's writing credits)**

	(1)	(2)	(3)	(4)
WRITEREXP_6YRS	-0.151**			
	(-3.20)			
WRITEREXP_6YRS^2	0.0475**			
	(3.14)			
WRITEREXP_5YRS_3YRS		-0.0975***		
		(-3.68)		
WRITEREXP_5YRS_3YRS^2		0.0197***		
		(3.64)		
WRITEREXP_ALL			-0.0745**	
			(-2.77)	
WRITEREXP_ALL^2			0.0125**	
			(3.29)	
WRITEREXP_5YRS_BO				-0.084*
				(-3.07)
WRITEREXP_5YRS_BO^2				0.024*
				(3.17)
Controls	Y	Y	Y	Y
YEAR_SALE F.E.	Y	Y	Y	Y
MAJOR_STUDIO F.E.	Y	Y	Y	Y
GENRE F.E.	Y	Y	Y	Y
N	1638	1638	1638	1638

Note: The unit of analysis is a sale. The dependent variable is SPEC, which is a dummy indicating whether the sale is a spec versus a pitch, for all regressions. The marginal effects of the Probit estimations are reported. All regressions use the number of the writer's major writing credits as the measure for the writer's observable quality. These measures are defined in the same way as is WRITEREXP, except that (1) counts the previous 6 years, (2) counts the previous five years plus the future three years, (3) counts the writer's entire history in the industry, and (4) counts the previous five years but only if the movie also has a U.S. box office greater than \$2.5 million. They use the same set of controls as in Table 2, but their coefficients are omitted to save space. ***, **, and * are respectively significant levels of 1%, 5%, and 10%.

TABLE 6. **OLS Estimates for Movie Performance (split samples)**

	log(US_BO) WRITEREXP = 0 (1)	log(US_BO) WRITEREXP > 0 (2)	TIME_TO_MARKET WRITEREXP = 0 (3)	TIME_TO_MARKET WRITEREXP > 0 (4)
SPEC	-0.414*	0.613***	-1.572	-8.886*
	(0.232)	(0.141)	(4.703)	(5.003)
PROD_BUDGET	0.002	0.007**	0.802***	0.564
	(0.008)	(0.003)	(0.269)	(0.342)
NUM_SCREEN	0.054**	0.052***	0.351	-0.610*
	(0.024)	(0.009)	(0.322)	(0.314)
STAR	-0.701	0.269	-22.663***	-8.219
	(0.438)	(0.160)	(5.183)	(5.757)
FRANCHISE	1.293**	0.574	-11.382	-1.362
	(0.562)	(0.342)	(10.135)	(8.916)
WRITEREXP		0.225***		-0.877
		(0.070)		(2.676)
PROD_BUDGET^2			-0.003	-0.003
			(0.002)	(0.003)
WRITER_DIRECT			4.046	-16.549**
			(8.400)	(6.945)
WRITER_ACT			-6.788	4.756
			(21.279)	(10.753)
WRITER_PRODUCE			-8.492	0.522
			(11.225)	(6.283)
ATTACH_ACTOR			-8.904	-11.685**
			(6.920)	(5.655)
ATTACH_DIRECTOR			-2.387	5.644
			(5.452)	(6.235)
PRODUCER_EXP			-1.652	-1.987
			(1.088)	(1.830)
PRODUCER_EXP^2			0.107	0.188
			(0.068)	(0.123)
Constant	8.077***	0.577	5.786	49.478***
	(1.518)	(0.624)	(15.024)	(9.423)
GENRE F.E.	Y	Y	Y	N
MPAA_RATING F.E.	Y	Y	Y	N
DISTRIBUTOR F.E.	Y	N	Y	Y
YEAR_RELEASE F.E.	Y	Y	N	N
WEEK_RELEASE F.E.	Y	N	N	N
YEAR_SALE F.E.	N	N	Y	N
R-squared	0.733	0.790	0.294	0.285
N	99	50	99	50

Note: The unit of analysis is a released movie. The dependent variable for (1) and (2) is log(US_BO), where US_BO is measured in $millions. The dependent variable for (3) and (4) is the number of months between the sale date and the release date in the U.S. (taking into account of the average of four months leading time of specs). The observations here are movies that are released and have production budget information. (1) and (3) use movies sold from writers without any major writing credits in the previous five years (i.e., WRITEREXP = 0); while (2) and (4) use movies sold from writers with some major writing credits (i.e., WRITEREXP > 0). ***, **, and * are, respectively, significant levels of 1%, 5%, and 10%.

TABLE 7. **OLS Estimates for Price**

	(1)	(2)	(3)
WRITEREXP	0.309***	0.196***	0.182***
	(0.049)	(0.044)	(0.041)
SPEC			-0.096**
			(0.045)
SPEC × WRITEREXP			0.177***
			(0.053)
NUMWRITER	0.002	0.062	0.020
	(0.058)	(0.054)	(0.039)
WRITER_ANYMOVIE	0.017	0.028	0.031
	(0.062)	(0.066)	(0.044)
WRITER_TV	-0.179	-0.093	-0.157
	(0.188)	(0.137)	(0.109)
WRITER_DIRECTOR	0.208	-0.024	0.034
	(0.180)	(0.123)	(0.100)
WRITER_ACTOR	0.015	0.117	0.067
	(0.104)	(0.080)	(0.062)
WRITER_PRODUCER	0.128	-0.065	0.030
	(0.117)	(0.089)	(0.069)
BIG_AGENT	0.115*	0.246***	0.180***
	(0.061)	(0.057)	(0.041)
WRITER_DIRECT	0.179	0.120	0.168
	(0.175)	(0.146)	(0.109)
WRITER_ACT	-0.692	0.375	0.116
	(0.471)	(0.238)	(0.204)
WRITER_PRODUCE	0.511***	0.277**	0.370***
	(0.166)	(0.120)	(0.096)
ATTACH_ACTOR	0.296***	0.214**	0.249***
	(0.093)	(0.095)	(0.065)
ATTACH_DIRECTOR	0.223**	0.206**	0.238***
	(0.092)	(0.083)	(0.060)
PRODUCEREXP	0.016	-0.012	0.008
	(0.015)	(0.015)	(0.010)
$PRODUCEREXP^2$	-0.000	0.001	0.000
	(0.001)	(0.001)	(0.001)
Constant	5.115***	5.354***	5.279***
	(0.138)	(0.144)	(0.103)
YEAR_SALE F.E.	Y	Y	Y
MAJOR_STUDIO F.E.	Y	Y	Y
GENRE F.E.	Y	Y	Y
(Adj.) R^2	0.31	0.21	0.28
N	488	444	932

Note: The dependent variable is log(PRICE) for all regressions. PRICE is roughly the price for the first two drafts and a polish. 932 observations have price information. (1) is for specs, (2) is for pitches, and (3) uses all observations that have price information. ***, **, and * are respectively significant levels of 1%, 5%, and 10%.

TABLE 8. **Descriptive Statistics of Alternative Measures of Writer's Observable Quality**

Variable	Obs.	Mean	S.D.	Min.	Max.
WRITER_BO	333	43.67	42.19	0.01	245.85
WRITER_RATING	429	5.71	0.97	2.8	8.6

Note: WRITER_BO is the average U.S. box office (in $millions) of the writer's movies that had been released by major studios before the sale; WRITER_RATING is the average *IMDb* rating of those movies. WRITER_BO is available for 333 sales, and WRITER_RATING is available for 429 sales.

TABLE 9. **Correlations between Alternative Measures of Writer's Observable Quality and Idea's Quality**

	WRITEREXP	log(WRITER_BO)	WRITER_RATING	RELEASED	log(PRICE)
WRITEREXP	1.000				
(Obs.)	(333)				
log(WRITER_BO)	0.24***	1.000			
(Obs.)	(333)	(333)			
WRITER_RATING	0.060	0.080	1.000		
(Obs.)	(333)	(333)	(333)		
RELEASED	0.16***	0.010	0.040	1.000	
(Obs.)	(333)	(333)	(333)	(333)	
log(PRICE)	0.38***	0.28***	0.32***	0.36***	1.000
(Obs.)	(185)	(185)	(185)	(185)	(185)

Note: 333 sales have all three measures, and 185 of them have price information. WRITEREXP is the number of the writer's major writing credits in the previous five years; WRITER_BO is the average U.S. box office (in $millions) of the writer's movies that had been released by major studios before the sale; WRITER_RATING is the average *IMDb* rating of those movies; RELEASED is a dummy variable indicating whether the movie is released by November 2009; and PRICE is roughly the price for the first two drafts and a polish. ***, **, and * are respectively significant levels of 1%, 5%, and 10%.

TABLE 10. **Probit Estimates for the Probability of a Spec Sale (alternative measures of the writer's observable quality)**

	(1)	(2)	(3)	(4)	(5)	(6)
log(WRITER_BO)	0.025	0.018				
	(0.025)	(0.051)				
log(WRITER_BO)^2		0.002				
		(0.010)				
WRITEREXP			-0.276**			-0.216***
			(0.112)			(0.079)
WRITEREXP^2			0.083***			0.069***
			(0.029)			(0.023)
WRITER_RATING				0.048*	0.186	
				(0.029)	(0.229)	
WRITER_RATING^2					-0.012	
					(0.020)	
Controls	Y	Y	Y	Y	Y	Y
YEAR_SALE F.E.	Y	Y	Y	Y	Y	Y
MAJOR_STUDIO F.E.	Y	Y	Y	Y	Y	Y
GENRE F.E.	Y	Y	Y	Y	Y	Y
N	333	333	333	429	429	429

Note: The dependent variable is SPEC for all regressions, and the marginal effects of the Probit estimations are reported. They use the same set of controls as in Table 2 although, their coefficients are omitted to save space. For comparison, (3) and (6) are results using WRITEREXP for the same observations when WRITER_BO and WRITER_RATING are available. ***, **, and * are respectively significant levels of 1%, 5%, and 10%.

APPENDIX

A. Proof

Lemma A1 and A2 are useful to prove the results of the paper. The expected value of an idea given (w, θ, ϵ_i) (defined in (2)) is

$$\begin{aligned} v(w,\theta,\epsilon_i) &= \mathbb{P}[w+\theta+\epsilon_i+\epsilon_s \geq 0]\mathbb{E}[w+\theta+\epsilon_i+\epsilon_s | w+\theta+\epsilon_i+\epsilon_s \geq 0] \\ &= (1-F^s(-w-\theta-\epsilon_i)) \int_{-w-\theta-\epsilon_i}^{\infty} (w+\theta+\epsilon_i+\epsilon_s) \frac{f^s(\epsilon_s)}{(1-F^s(-w-\theta-\epsilon_i))} d\epsilon_s. \end{aligned} \tag{A1}$$

Lemma A1. $\frac{\partial v(w,\theta,\epsilon_i)}{\partial w} = \frac{\partial v(w,\theta,\epsilon_i)}{\partial \theta} = \frac{\partial v(w,\theta,\epsilon_i)}{\partial \epsilon_i} > 0$; $\frac{\partial^2 v(w,\theta,\epsilon_i)}{\partial w^2} = \frac{\partial^2 v(w,\theta,\epsilon_i)}{\partial \theta^2} = \frac{\partial^2 v(w,\theta,\epsilon_i)}{\partial \epsilon_i^2} = \frac{\partial^2 v(w,\theta,\epsilon_i)}{\partial w \partial \epsilon_i} = \frac{\partial^2 v(w,\theta,\epsilon_i)}{\partial w \partial \theta} = \frac{\partial^2 v(w,\theta,\epsilon_i)}{\partial \theta \partial \epsilon_i} > 0.$

Proof. $\frac{\partial v(w,\theta,\epsilon_i)}{\partial w} = 1 - F^s(-w-\theta-\epsilon_i) > 0$. $\frac{\partial^2 v(w,\theta,\epsilon_i)}{\partial w^2} = f^s(-w-\theta-\epsilon_i) > 0$. The other results hold by symmetry. □

Lemma A2. $\lim_{w\to\infty} v(w,\theta,\epsilon_i) = \infty$ *and* $\lim_{w\to-\infty} v(w,\theta,\epsilon_i) = 0$. *By symmetry,* $v(w,\theta,\epsilon_i)$ *have the same limits with respect to* θ *and* ϵ_i*.*

Proof. Given θ and ϵ_i, $\lim_{w\to\infty} v(w,\theta,\epsilon_i) = 1 \times (\infty + \theta + \epsilon_i) + \mathbb{E}[\epsilon_s] = \infty$.

Integrating (A1) by parts, one gets

$$v(w,\theta,\epsilon_i) = (1-F^s(-w-\theta-\epsilon_i)) \int_{-w-\theta-\epsilon_i}^{\infty} \frac{1-F^s(\epsilon_s)}{(1-F^s(-w-\theta-\epsilon_i))} d\epsilon_s. \tag{A2}$$

$\lim_{w\to-\infty} v(w,\theta,\epsilon_i) = \lim_{w\to-\infty} (1-F^s(-w-\theta-\epsilon_i)) \lim_{w\to-\infty} \int_{-w-\theta-\epsilon_i}^{\infty} \frac{1-F^s(\epsilon_s)}{(1-F^s(-w-\theta-\epsilon_i))} d\epsilon_s = 0 \times c = 0$. As $w \to -\infty$, $\int_{-w-\theta-\epsilon_i}^{\infty} \frac{1-F^s(\epsilon_s)}{(1-F^s(-w-\theta-\epsilon_i))} d\epsilon_s$ is a decreasing series (explained below), and is bounded from below by 0. This implies that it must converge to some nonnegative finite number c.

The derivative of $\int_{-w-\theta-\epsilon_i}^{\infty} \frac{1-F^s(\epsilon_s)}{(1-F^s(-w-\theta-\epsilon_i))} d\epsilon_s$ w.r.t. w is

$$\frac{\int_{-w-\theta-\epsilon_i}^{\infty} ((1-F^s(-w-\theta-\epsilon_i)) f^s(\epsilon_s) - (1-F^s(\epsilon_s)) f^s(-w-\theta-\epsilon_i)) d\epsilon_s}{(1-F^s(-w-\theta-\epsilon_i))^2} > 0.$$

The inequality holds because under Assumption 3, $(1-F^s(-w-\theta-\epsilon_i)) f^s(\epsilon_s) > (1-F^s(\epsilon_s)) f^s(-w-\theta-\epsilon_i)$, for all $\epsilon_s > -w-\theta-\epsilon_i$.

The other results hold by symmetry. □

***Proof of Proposition 1**.* I first prove the existence, the uniqueness follows.

□ *The writer's strategy.*

When $w \geq m_p$, the writer anticipates being met for both spec and pitch. Dropping the idea is dominated by pitch, because the latter always yields a positive payoff. The difference between spec and pitch, $\Delta^{W}(w,\theta)$, is (3). As will be explained below, $\frac{\partial\Delta^{W}(w,\theta)}{\partial\theta} > 0$, $\lim_{\theta\to\infty}\Delta^{W}(w,\theta) = \infty$, and $\lim_{\theta\to-\infty}\Delta^{W}(w,\theta) = -c_s$. Then, $\exists\, r_0(w) \in R$ such that $\Delta^{W}(w,\theta) \geq 0$ if and only if $\theta \geq r_0(w)$.

Let $\epsilon_i^*(w,\theta)$ be the solution of $v(w,\theta,\epsilon_i) = c_s$. Because $\frac{\partial v(w,\theta,\epsilon_i)}{\partial\theta} > 0$ and $\frac{\partial v(w,\theta,\epsilon_i)}{\partial\epsilon_i}$,

$$\frac{\partial\Delta^{W}(w,\theta)}{\partial\theta} = \alpha_p \int_{-\infty}^{\epsilon_i^*(w,\theta)} \frac{\partial v(w,\theta,\epsilon_i)}{\partial\theta} dF^i(\epsilon_i) + (\alpha_s - \alpha_p)\int_{-\infty}^{\infty} \frac{\partial v(w,\theta,\epsilon_i)}{\partial\theta} dF^i(\epsilon_i) > 0.$$

As $\theta \to \infty$, $v(w,\theta,\epsilon_i)$ is nonnegative and monotone increasing. By monotone convergence theorem,

$$\begin{aligned}\lim_{\theta\to\infty}\Delta^{W}(w,\theta) &= \alpha_p \lim_{\theta\to\infty}\textstyle\int_{-\infty}^{\epsilon_i^*(w,\theta)}(v(w,\theta,\epsilon_i) - c_s)dF^i(\epsilon_i) - (1-\alpha_p)c_s \\ &\quad + (\alpha_s - \alpha_p)\mathbb{E}[\lim_{\theta\to\infty} v(w,\theta,\epsilon_i)] \\ &= 0 - (1-\alpha_p)c_s + \infty = \infty.\end{aligned}$$

$\lim_{\theta\to\infty}\int_{-\infty}^{\epsilon_i^*(w,\theta)}(v(w,\theta,\epsilon_i) - c_s)dF^i(\epsilon_i) = \lim_{\theta\to\infty} F^i(\epsilon_i^*(w,\theta)) \lim_{\theta\to\infty}\int_{-\infty}^{\epsilon_i^*(w,\theta)}(v(w,\theta,\epsilon_i) - c_s)\frac{f^i(\epsilon_i)}{F^i(\epsilon_i^*(w,\theta))}d\epsilon_i = 0 \times c = 0$. $\lim_{\theta\to\infty} F^i(\epsilon_i^*(w,\theta)) = 0$ because $\epsilon_i^*(w,\theta) \to \infty$ ($\frac{\epsilon_i^*(w,\theta)}{\partial\theta} = -1$). As $\theta \to \infty$, $\int_{-\infty}^{\epsilon_i^*(w,\theta)}(v(w,\theta,\epsilon_i) - c_s)\frac{f^i(\epsilon_i)}{F^i(\epsilon_i^*(w,\theta))}d\epsilon_i$ is an increasing series, and is bounded from above by zero. Then, it must converge to a finite non-positive number, c.

As $\theta \to -\infty$, $v(w,\theta,\epsilon_i)$ is nonnegative and monotone decreasing. By dominated convergence argument, and $\lim_{\theta\to-\infty} v(w,\theta,\epsilon_i) = 0$.

$$\begin{aligned}\lim_{\theta\to-\infty}\Delta^{W}(w,\theta) &= \alpha_p \textstyle\int_{-\infty}^{\infty}(\lim_{\theta\to-\infty} v(w,\theta,\epsilon_i) - c_s)dF^i(\epsilon_i) - (1-\alpha_p)c_s \\ &\quad + (\alpha_s - \alpha_p)\mathbb{E}[\lim_{\theta\to-\infty} v(w,\theta,\epsilon_i)] \\ &= -c_s.\end{aligned}$$

When $w < m_p$, the writer anticipates being met only if he chooses to spec. Similarly, it can be shown that $\frac{\partial S^{W}(w,\theta)}{\partial\theta} > 0$, $\lim_{\theta\to\infty} S^{W}(w,\theta) = \infty$, and $\lim_{\theta\to-\infty} S^{W}(w,\theta) = -c_s$. Thus, $\exists\, r_s(w) \in R$ such that $S^{W}(w,\theta) \geq 0$ if and only if $\theta \geq r_s(w)$.

By the implicit function theorem, $r_0'(w) = -\frac{\frac{\partial\Delta^{W}(w,\theta)}{\partial w}}{\frac{\partial\Delta^{W}(w,\theta)}{\partial\theta}}|_{w,r_0(w)} = -1$. $\frac{\partial\Delta^{W}(w,\theta)}{\partial w} = \frac{\partial\Delta^{W}(w,\theta)}{\partial\theta}$, $\forall\,(w,\theta)$, by symmetry. $r_s'(w) = -1$ is proved similarly.

Furthermore, $r_0(w) > r_s(w)$ because $S^{W}(w, r_0(w)) > 0$.

☐ *Buyer's strategy and beliefs.*

First, the buyer's meeting threshold for spec is $-\infty$, because $S^{B}(w) > 0$, $\forall\, w$. To see this, when $w \geq m_p$, given the writer's strategy, the buyer's belief $H(\theta|w,S)$ is $\frac{g(\theta)}{1-G(r_0(w))}$ for $\theta \in [r_0(w),\infty)$, and 0 elsewhere. Then,

$$\begin{aligned} S^{\mathrm{B}}(w) &= \int_{r_0(w)}^{\infty}\{(1-\alpha_s)\mathbb{E}[v(w,\theta,\epsilon_i)]-c_m\}\frac{g(\theta)}{1-G(r_0(w))}d\theta \\ &> \int_{r_s(w)}^{\infty}\{(1-\alpha_s)\mathbb{E}[v(w,\theta,\epsilon_i)]-c_m\}\frac{g(\theta)}{1-G(r_s(w))}d\theta \\ &> (1-\alpha_s)\mathbb{E}[v(w,r_s(w),\epsilon_i)]-c_m \\ &= \left(\frac{1-\alpha_s}{\alpha_s}\right)c_s-c_m \\ &> 0. \end{aligned} \tag{A3}$$

The first inequality holds because the signal $\theta \in [r_0(w),\infty)$ is more favorable than $\theta \in [r_s(w),\infty)$ in the sense of first-order stochastic dominance. The second equality holds, because $S^{\mathrm{W}}(w,r_s(w)) = 0 \Rightarrow \mathbb{E}[v(w,r_s(w),\epsilon_i)] = \frac{c_s}{\alpha_s}$. The last inequality holds under Assumption 2.

When $w < m_p$, the buyer's belief $H(\theta|w,S)$ is $\frac{g(\theta)}{1-G(r_s(w))}$ for $\theta \in [r_s(w),\infty)$, and 0 elsewhere. Line 2 onwards in (A3) show that the buyer's expected payoff from meeting the writer for spec is always positive as well. Thus, the buyer's best response for a writer who has chosen to spec is "always meet".

Second, if the writer has chosen to pitch, the buyer's posterior is $H(\theta|w,P)$ is $\frac{g(\theta)}{G(r_0(w))}$ for $\theta \in (-\infty,r_0(w))$, and 0 elsewhere. This belief is consistent with Bayes' rule when $w \geq m_p$. When $w < m_p$, because the writer's choosing to pitch is a probability zero event in the equilibrium, I impose the buyer's off-the-equilibrium posterior to be so.

As will be explained later, under $H(\theta|w,P)$, the buyer's expected payoff from meeting a pitch, $P^{\mathrm{B}}(w)$, is monotone increasing in w. Plus that $\lim_{w\to-\infty} P^{\mathrm{B}}(w) = -c_m$ and $\lim_{w\to\infty} P^{\mathrm{B}}(w) > 0$. Therefore, there exists $m_p \in R$ such that $P^{\mathrm{B}}(w) \geq 0$ if and only if $w \geq m_p$.

To see that $P^{\mathrm{B}}(w)$ is monotone increasing in w, let $P^{\mathrm{B}}(w,\theta) = (1-\alpha_p)\mathbb{P}(v(w,\theta,\epsilon_i) \geq c_s)\mathbb{E}[v(w,\theta,\epsilon_i) - c_s|v(w,\theta,\epsilon_i) \geq c_s] - c_m$, then, the buyer's expected payoff is

$$\begin{aligned} P^{\mathrm{B}}(w) &= \int_{-\infty}^{r_0(w)} P^{\mathrm{B}}(w,\theta)\frac{g(\theta)}{G(r_0(w))}d\theta \\ &= P^{\mathrm{B}}(w,r_0(w)) - \int_{-\infty}^{r_0(w)} \frac{\partial P^{\mathrm{B}}(w,\theta)}{\partial\theta}\frac{G(\theta)}{G(r_0(w))}d\theta. \end{aligned} \tag{A4}$$

Because $r_0'(w) = -1$, then

$$\begin{aligned} \frac{dP^{\mathrm{B}}(w)}{dw} &= \frac{\partial P^{\mathrm{B}}(w)}{\partial w} - \frac{\partial P^{\mathrm{B}}(w)}{\partial r_0(w)} \\ &= \frac{1}{G^2(r_0(w))}\int_{-\infty}^{r_0(w)} \frac{\partial P^{\mathrm{B}}(w,\theta)}{\partial\theta}\{g(\theta)G(r_0(w)) - g(r_0(w))G(\theta)\}d\theta > 0. \end{aligned}$$

The inequality holds because by Assumption 3, $g(\theta)G(r_0(w)) > g(r_0(w))G(\theta), \forall\, \theta < r_0(w)$.

$\lim_{w\to-\infty} P^{\mathrm{B}}(w) = \int_{-\infty}^{\infty} \lim_{w\to-\infty} P^{\mathrm{B}}(w,\theta)dG(\theta) = -c_m$ because $\lim_{w\to-\infty} P^{\mathrm{B}}(w,\theta) = -c_m$ given w and θ (the proof is similar to that for $\Delta^{\mathrm{W}}(w,\theta)$).

To see that $\lim_{w\to\infty} P^{\mathrm{B}}(w) > 0$, note that $\lim_{w\to\infty} P^{\mathrm{B}}(w) = \lim_{w\to\infty} P^{\mathrm{B}}(w, r_0(w)) - \lim_{w\to\infty} \int_{-\infty}^{r_0(w)} \frac{\partial P^{\mathrm{B}}(w,\theta)}{\partial\theta} \frac{G(\theta)}{G(r_0(w))} d\theta = P^{\mathrm{B}}(w, r_0(w))$. The second term tends to 0 because both terms in the integral are positive and bounded from above by 1, and $\lim_{w\to\infty} r_0(w) = -\infty$. $P^{\mathrm{B}}(w, r_0(w))$ is a constant because $r_0'(w) = -1$. To see that it is positive, write $P^{\mathrm{B}}(w, r_0(w)) = (1-\alpha_p)\bar{P} - c_m$, where $\bar{P} = \mathbb{P}(v(w, r_0(w), \epsilon_i) \geq c_s)\mathbb{E}[v(w, r_0(w), \epsilon_i) - c_s | v(w, r_0(w), \epsilon_i) \geq c_s]$. The following shows that $(1-\alpha_p)\bar{P} > \frac{1-\alpha_s}{\alpha_s} c_s$, which is greater than c_m.

$\mathbb{E}[v(w, r_0(w), \epsilon_i) - c_s] = \mathbb{P}(v(w, r_0(w), \epsilon_i) < c_s)\mathbb{E}[v(w, r_0(w), \epsilon_i) - c_s | v(w, r_0(w), \epsilon_i) < c_s] + \bar{P}$. Then, $S^{\mathrm{W}}(w, r_0(w)) = P^{\mathrm{W}}(w, r_0(w))$ implies that $\mathbb{P}(v(w, r_0(w), \epsilon_i) < c_s)\mathbb{E}[v(w, r_0(w), \epsilon_i) - c_s | v(w, r_0(w), \epsilon_i) < c_s] = \frac{1-\alpha_s}{\alpha_s} c_s - \frac{\alpha_s - \alpha_p}{\alpha_s}\bar{P}$. The LHS of the equation is between $-c_s$ and 0, which implies $\frac{(1-\alpha_p)(1-\alpha_s)}{\alpha_s - \alpha_p} c_s \leq (1-\alpha_p)\bar{P} \leq \frac{(1-\alpha_p)c_s}{\alpha_s - \alpha_p}$. $(1-\alpha_p)\bar{P} > \frac{1-\alpha_s}{\alpha_s} c_s$, because the lower bound $\frac{(1-\alpha_p)(1-\alpha_s)}{\alpha_s - \alpha_p} c_s > \frac{1-\alpha_s}{\alpha_s} c_s$.

□ *Uniqueness of the equilibrium.*

The equilibrium is the unique semi-separating equilibrium. The buyer's posterior beliefs matter only in terms of her meeting decisions. Anticipating the buyer's meeting decisions, the writer's problem, both when $w \geq m_p$ and when $w < m_p$, are not functions of the buyer's beliefs. The writer's problem has a unique solution. Conversely, given the writer's strategy, (**??**) is the only belief system that is consistent with Bayes' rule. Under these beliefs, the buyer's expected payoffs (for both spec and pitch) from meeting the writer are also monotone increasing in the writer's observable quality, w. Thus, the buyer's meeting threshold for pitch is unique as well.

Note that Bayes' rule does not determine the posterior distribution over the writer's choice to which the equilibrium assigns zero probability, i.e., if a writer of quality $w < m_p$ has chosen to pitch. Any posterior, including what is specified above, such that the buyer finds it not worthwhile to meet the writer for pitch sustains the equilibrium. To see that the buyer meets both spec and pitch for some $w < m_p$ can not be an equilibrium, suppose this were the case under some posterior belief, then, given the buyer's strategy, the writer's best response is to spec if $w \geq r_0(w)$, and pitch otherwise. This can not be an equilibrium, because, given the writer's strategy and Bayes' rule, the buyer's expected payoff given that the writer has chosen to pitch is negative, thus, her best response is to reject him. □

Proof of Prediction 1*.* By Bayes' rule, conditional on sale, the likelihood of spec is

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

$\mathbb{P}(\text{spec} \cap \text{sale}|w) = \mathbb{P}(\text{spec}|w)\mathbb{P}(\text{sale}|\text{spec}, w)$ is the joint probability that the writer has chosen to spec in the first place, and the spec is sold. $\mathbb{P}(\text{pitch} \cap \text{sale}|w)$ has similar interpretation.

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

When $w \geq m_p$, $\mathbb{P}(\text{spec}|\text{sale}, w)$ increases with w if and only if

$$\frac{\partial \mathbb{P}(\text{spec} \cap \text{sale}|w)/\partial w}{\mathbb{P}(\text{spec} \cap \text{sale}|w)} \geq \frac{\partial \mathbb{P}(\text{pitch} \cap \text{sale}|w)/\partial w}{\mathbb{P}(\text{pitch} \cap \text{sale}|w)}. \tag{A5}$$

First, $\frac{\partial \mathbb{P}(\text{spec} \cap \text{sale}|w)}{\partial w} > 0$ because $\mathbb{P}(\text{spec} \cap \text{sale}|w) = \mathbb{P}(\text{spec}|w)\mathbb{P}(\text{sale}|\text{spec}, w)$, and both terms increase with w. $\frac{\partial \mathbb{P}(\text{spec}|w)}{\partial w} = -G(r_0(w))r_0'(w) > 0$. $\frac{\partial \mathbb{P}(\text{sale}|\text{spec},w)}{\partial w} = \frac{\partial \int_{r_0(w)}^{\infty} \mathbb{P}(\epsilon_i + \epsilon_s \geq -w-\theta) \frac{g(\theta)}{1-G(r_0(w))} d\theta}{\partial w} = \int_{r_0(w)}^{\infty} \frac{\partial \mathbb{P}(\epsilon_i+\epsilon_s \geq -w-\theta)}{\partial \theta} \frac{g(\theta)(1-G(r_0(w)))-g(r_0(w))(1-G(\theta))}{(1-G(r_0(w)))^2} d\theta > 0$. The inequality holds because $g(\theta)$ has monotone increasing hazard rate.

Second, $\frac{\partial \mathbb{P}(\text{pitch} \cap \text{sale}|w)}{\partial w} < 0$, when w is sufficiently high. Then, condition (A5) holds when w is sufficiently high.

$$\begin{aligned} \mathbb{P}(\text{pitch} \cap \text{sale}|w) &= \mathbb{P}(\text{pitch}|w)\mathbb{P}(\text{sale}|\text{pitch}, w) \\ &= G(r_0(w)) \int_{-\infty}^{r_0(w)} [1 - F(\epsilon_i^*(w, \theta))] \frac{g(\theta)}{G(r_0(w))} d\theta, \end{aligned}$$

where $\epsilon_i^*(w, \theta)$ is the solution of $v(w, \epsilon_i, \theta) = c_s$. Because $\frac{\partial \epsilon_i^*(w,\theta)}{\partial w} = -1$ and $r_0'(w) = -1$, $\frac{\partial \mathbb{P}(\text{pitch} \cap \text{sale}|w)}{\partial w} < 0$ if and only if

$$\frac{g(r_0(w))}{G(r_0(w))} > \frac{\int_{-\infty}^{r_0(w)} f^i(\epsilon_i^*(w,\theta)) \frac{g(\theta)}{G(r_0(w))} d\theta}{\int_{-\infty}^{r_0(w)} [1 - F^i(\epsilon_i^*(w,\theta))] \frac{g(\theta)}{G(r_0(w))} d\theta} + \frac{-\frac{\partial \int_{-\infty}^{r_0(w)} [1-F^i(\epsilon_i^*(w,\theta))] \frac{g(\theta)}{G(r_0(w))} d\theta}{\partial r_0(w)}}{\int_{-\infty}^{r_0(w)} [1 - F^i(\epsilon_i^*(w,\theta))] \frac{g(\theta)}{G(r_0(w))} d\theta}. \tag{A6}$$

$\frac{\int_{-\infty}^{r_0(w)} f^i(\epsilon_i^*(w,\theta)) \frac{g(\theta)}{G(r_0(w))} d\theta}{\int_{-\infty}^{r_0(w)} [1 - F^i(\epsilon_i^*(w,\theta))] \frac{g(\theta)}{G(r_0(w))} d\theta} = \frac{f^i(\epsilon_i^*(w, r_0(w))) + \int_{-\infty}^{r_0(w)} f^{i'}(\epsilon_i^*(w,\theta)) \frac{G(\theta)}{G(r_0(w))} d\theta}{1 - F^i(\epsilon_i^*(w, r_0(w))) + \int_{-\infty}^{r_0(w)} f^i(\epsilon_i^*(w,\theta)) \frac{G(\theta)}{G(r_0(w))} d\theta}$ (integrate by parts). As as $w \to \infty$, $r_0(w) \to -\infty$, this term tends to a constant $\frac{f^i(\epsilon_i^*(w, r_0(w)))}{1 - F^i(\epsilon_i^*(w, r_0(w)))}$. On the other hand, $\frac{g(r_0(w))}{G(r_0(w))} \to \infty$. In addition, $-\frac{\partial \int_{-\infty}^{r_0(w)} [1-F^i(\epsilon_i^*(w,\theta))] \frac{g(\theta)}{G(r_0(w))} d\theta}{\partial r_0(w)} = -\int_{-\infty}^{r_0(w)} f^i(\epsilon_i^*(w,\theta)) \frac{G(\theta)g(r_o(w))}{G^2(r_0(w))} d\theta$ is negative. Thus, (A6) holds, i.e., $\frac{\partial \mathbb{P}(\text{pitch} \cap \text{sale}|w)}{\partial w} < 0$, when w is sufficiently high.

Finally, $\lim_{w \to \infty} \mathbb{P}(\text{pitch} \cap \text{sale}|w) = 0$ implies $\lim_{w \to \infty} \mathbb{P}(\text{spec}|\text{sale}, w) = 1$. Thus, $\mathbb{P}(\text{spec}|\text{sale}, w)$ is at its minimum for intermediate value of w. □

Proof of Prediction 2. Define “released" as the terminal value (if realized) is greater than zero, i.e., $V \geq 0$. Given writer quality w, the conditional expected value of movies purchased as spec is

$$\mathbb{E}[V|\text{released}, \text{spec}, w] = \mathbb{E}[V|\theta \geq r_0(w), V \geq 0], \tag{A7}$$

where the expected value is conditional on 1) the writer has chosen to spec, i.e., $\theta \geq r_0(w)$, and 2) the finished script is released, i.e., $V \geq 0$. The conditional expected value of movies purchased as pitch is

$$\mathbb{E}[V|\text{released}, \text{pitch}, w] = \mathbb{E}[V|\theta < r_0(w), v(w,\theta,\epsilon_i) \geq c_s, V \geq 0] \tag{A8}$$

where the expected value is conditional on 1) the writer has chosen to pitch, i.e., $\theta < r_0(w)$, 2) the pitch is sold, i.e., $v(w,\theta,\epsilon_i) \geq c_s$, and 3) finished script is released, i.e., $V \geq 0$.

It is clear from (A7) and (A8) that the writer selection effect is reflected by $\theta \geq r_0(w)$ for spec, and $\theta < r_0(w)$ for pitch. The extra evaluation effect is reflected by the extra condition $v(w,\theta,\epsilon_i) \geq c_s$ for pitch. In the following, I first consider the writer selection effect only, and then add the extra evaluation effect.

□ **Writer selection effect.** When considering the writer selection effect only, the conditional expected value of spec is greater than that of pitch, irrespective of the writer quality, i.e., $\mathbb{E}[V|\theta \geq r_0(w), V \geq 0] > \mathbb{E}[V|\theta < r_0(w), V \geq 0]$, for all w.

To see this, let $\delta(w,\theta) = \mathbb{E}[V|\epsilon_i + \epsilon_s \geq -w - \theta]$. Write the conditional expected values of spec and pitch in the form of the expectation of $\delta(w,\theta)$ over their corresponding domain of θ's.

$$\begin{aligned}
\mathbb{E}[V|\theta \geq r_0(w), V \geq 0] &= \frac{\int_{r_0(w)}^{\infty} \delta(w,\theta)(1 - F(-w-\theta))g(\theta)d\theta}{\int_{r_0(w)}^{\infty}(1 - F(-w-\theta))g(\theta)d\theta} \\
\mathbb{E}[V|\theta < r_0(w), V \geq 0] &= \frac{\int_{-\infty}^{r_0(w)} \delta(w,\theta)(1 - F(-w-\theta))g(\theta)d\theta}{\int_{-\infty}^{r_0(w)}(1 - F(-w-\theta))g(\theta)d\theta}.
\end{aligned}$$

Because $\delta(w,\theta)$ is an increasing function of θ,

$$\mathbb{E}[V|\theta \geq r_0(w), V \geq 0] > \delta(w, r_0(w)) > \mathbb{E}[V|\theta < r_0(w), V \geq 0].$$

□ **Adding the extra evaluation effect of pitch.** Adding the extra evaluation effect increases the conditional expected value of pitch, i.e., $\mathbb{E}[V|\theta < r_0(w), \epsilon_i \geq \epsilon_i^*(w,\theta), V \geq 0] > \mathbb{E}[V|\theta < r_0(w), V \geq 0]$, where $\epsilon_i^*(w,\theta)$ is the solution of $v(w,\theta,\epsilon_i) = c_s$.

Let $\delta(w,\theta,\epsilon_i) = \mathbb{E}[V|\epsilon_s \geq -w - \theta - \epsilon_i]$. Again, the conditional expected value of pitch with and without considering the extra evaluation effect are

$$\begin{aligned}
&\mathbb{E}[V|\theta < r_0(w), \epsilon_i \geq \epsilon_i^*(w,\theta), V \geq 0] \\
= &\frac{\int_{-\infty}^{r_0(w)} \{\int_{\epsilon_i^*(w,\theta)}^{\infty}(1-F^s(-w-\theta-\epsilon_i))dF^i(\epsilon_i)\} \left(\frac{\int_{\epsilon_i^*(w,\theta)}^{\infty} \delta(w,\theta,\epsilon_i)(1-F^s(-w-\theta-\epsilon_i))f^i(\epsilon_i)d\epsilon_i}{\int_{\epsilon_i^*(w,\theta)}^{\infty}(1-F^s(-w-\theta-\epsilon_i))f^i(\epsilon_i)d\epsilon_i}\right) dG(\theta)}{\int_{-\infty}^{r_0(w)} \int_{\epsilon_i^*(w,\theta)}^{\infty}(1-F^s(-w-\theta-\epsilon_i))dF^i(\epsilon_i)dG(\theta)}
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}[V|\theta < r_0(w), V \geq 0] \\
= &\frac{\int_{-\infty}^{r_0(w)} \{\int_{-\infty}^{\infty}(1-F^s(-w-\theta-\epsilon_i))dF^i(\epsilon_i)\} \left(\frac{\int_{-\infty}^{\infty} \delta(w,\theta,\epsilon_i)(1-F^s(-w-\theta-\epsilon_i))f^i(\epsilon_i)d\epsilon_i}{\int_{-\infty}^{\infty}(1-F^s(-w-\theta-\epsilon_i))f^i(\epsilon_i)d\epsilon_i}\right) dG(\theta)}{\int_{-\infty}^{r_0(w)} \int_{-\infty}^{\infty}(1-F^s(-w-\theta-\epsilon_i))dF^i(\epsilon_i)dG(\theta)}.
\end{aligned}$$

Each expression above is the expectation of a function of θ over $(-\infty, r_0(w))$ with respect to its respective density. $\mathbb{E}[V|\theta < r_0(w), \epsilon_i \geq \epsilon_i^*(w,\theta), V \geq 0] > \mathbb{E}[V|\theta < r_0(w), V \geq 0]$, because 1) the integrand of the former is greater than that of the latter for all θ, and 2) the density of θ of the former is more favorable than that of the latter in terms of first-order stochastic dominance.

First, given (w,θ), $\frac{\int_{\epsilon_i^*(w,\theta)}^{\infty} \delta(w,\theta,\epsilon_i)(1-F^s(-w-\theta-\epsilon_i))f^i(\epsilon_i)d\epsilon_i}{\int_{\epsilon_i^*(w,\theta)}^{\infty}(1-F^s(-w-\theta-\epsilon_i))f^i(\epsilon_i)d\epsilon_i} > \frac{\int_{-\infty}^{\infty} \delta(w,\theta,\epsilon_i)(1-F^s(-w-\theta-\epsilon_i))f^i(\epsilon_i)d\epsilon_i}{\int_{-\infty}^{\infty}(1-F^s(-w-\theta-\epsilon_i))f^i(\epsilon_i)d\epsilon_i}$, because $\delta(w,\theta,\epsilon_i)$ is an increasing function of ϵ_i, and the density $\frac{(1-F^s(-w-\theta-\epsilon_i))f^i(\epsilon_i)}{\int_{\epsilon_i^*(w,\theta)}^{\infty} 1(1-F^s(-w-\theta-\epsilon_i))f^i(\epsilon_i)d\epsilon_i}$ for $\epsilon_i \in (\epsilon_i^*(w,\theta), \infty)$ and 0 elsewhere is more favorable than $\frac{(1-F^s(-w-\theta-\epsilon_i))f^i(\epsilon_i)}{\int_{-\infty}^{\infty} 1(1-F^s(-w-\theta-\epsilon_i))f^i(\epsilon_i)d\epsilon_i}$ in the sense of first-order stochastic dominance.

Second, the density $\frac{\int_{\epsilon_i^*(w,\theta)}^{\infty}(1-F^s(-w-\theta-\epsilon_i))dF^i(\epsilon_i)}{\int_{-\infty}^{r_0(w)}\int_{\epsilon_i^*(w,\theta)}^{\infty}(1-F^s(-w-\theta-\epsilon_i))dF^i(\epsilon_i)dG(\theta)}$ for $\theta \in (-\infty, r_0(w))$ and 0 elsewhere is more favorable than $\frac{\int_{-\infty}^{\infty}(1-F^s(-w-\theta-\epsilon_i))dF^i(\epsilon_i)}{\int_{-\infty}^{r_0(w)}\int_{-\infty}^{\infty}(1-F^s(-w-\theta-\epsilon_i))dF^i(\epsilon_i)dG(\theta)}$ for $\theta \in (-\infty, r_0(w))$ and 0 elsewhere in the sense of first-order stochastic dominance as well (the proof of the claim is omitted).

Finally, it can be shown that both conditional expected values increase with w. However, $\lim_{w\to\infty} \mathbb{E}[V|\theta \geq r_0(w), V \geq 0] = \infty$, and $\lim_{w\to\infty} \mathbb{E}[V|\theta < r_0(w), \epsilon \geq \epsilon_i^*(w,\theta), V \geq 0]$ is finite. The former is true because $\mathbb{E}[V|\theta \geq r_0(w), V \geq 0]$ is the expectation of $\delta(w,\theta)$, and $\delta(w,\theta)$ tends to infinity as $w \to \infty$. The latter is true because for all w, $\mathbb{E}[V|\theta < r_0(w), \epsilon \geq \epsilon_i^*(w,\theta), V \geq 0] < \mathbb{E}[V|\theta = r_0(w), \epsilon \geq \epsilon_i^*(w, r_0(w)), V \geq 0]$, which is a constant.

Thus, when the writer's observable quality w is sufficiently high, the writer selection effect dominates the extra evaluation effect. The conditional performance of spec is better. When w is sufficiently low, it is possible that pitch performs better, because the extra evaluation effect may dominate the writer selection effect. □

B. A variation of the model: buyer does not observe θ until the script is finished

In the paper, the buyer observes the writer's private signal, θ, as soon as she hears the story. This is reasonable for spec, because the qualities of the idea and the script are clear, given a complete script. For pitch, however, the writer may find it difficult to convey what he believes about the idea's potential. Here, I discuss another extreme where the buyer does not observe θ until the script is finished. The reality is likely to lie somewhere in between.

The sequence of events is similar to that in the model, except that if the writer has chosen to pitch, the buyer does not observe the writer's private signal, θ, until the script is finished. The buyer still observes the writer's observable quality, w, and the quality of the abstract idea, ϵ_i, upon meeting. To avoid solving a bargaining problem under asymmetric information, I make two simplifying assumptions about the negotiation stage, given a pitch. First, the expected value of the idea, $\tilde{v}(w, \epsilon_i)$, is calculated based on what the buyer's information — i.e.,

$$\tilde{v}(w, \epsilon_i) = \int_{-\infty}^{\infty} v(w, \theta, \epsilon_i) dH(\theta | w, P), \tag{A9}$$

where the density of the buyer's posterior given a pitch from a writer of quality w is $h(\theta|w, P) = \frac{g(\theta)z(P|w,\theta)}{\int_{-\infty}^{\infty} g(\theta')z(P|w,\theta)d\theta'}$, and $z(P|w, \theta)$ is the probability that the writer chooses to pitch given (w, θ). Second, the writer accepts the offer if and only if his payment is greater than the writing cost he is to incur.

Analogous to a generalized Nash bargaining solution, the pitch is sold if and only if its expected surplus is positive — i.e., $\tilde{v}(w, \epsilon_i) \geq c_s$. Given a sale, the writer's and the buyer's payoffs at the negotiation stage are, respectively, $\alpha_p(\tilde{v}(w, \epsilon_i) - c_s)$ and $(1 - \alpha_p)(\tilde{v}(w, \epsilon_i) - c_s)$.

At Stage 1, the writer's expected payoffs from spec and pitch are

$$\begin{aligned} \tilde{S}^{\mathrm{W}}(w, \theta) &= \alpha_s \mathbb{E}[v(w, \theta, \epsilon_i)] - c_s \\ \tilde{P}^{\mathrm{W}}(w) &= \alpha_p \mathbb{P}(\tilde{v}(w, \epsilon_i) \geq c_s) \mathbb{E}[\tilde{v}(w, \epsilon_i) - c_s | \tilde{v}(w, \epsilon_i) \geq c_s]. \end{aligned}$$

The writer's expected payoff from spec is the same in the benchmark case. His expected payoff from pitch does not depend on θ now, because the buyer does not observe it.

At Stage 2, the buyer decides whether or not to meet the writer, updating her belief, $H(\theta|w, S)$ and $H(\theta|w, P)$, about the writer's private signal θ. The buyer's expected payoffs from meeting the writer for spec and for pitch are

$$\begin{aligned} \tilde{S}^{\mathrm{B}}(w) &= \textstyle\int_{-\infty}^{\infty} \{(1 - \alpha_s) \mathbb{E}[v(w, \theta, \epsilon_i)] - c_m\} dH(\theta | w, S) \\ \tilde{P}^{\mathrm{B}}(w) &= (1 - \alpha_p) \mathbb{P}(\tilde{v}(w, \epsilon_i) \geq c_s) \mathbb{E}[\tilde{v}(w, \epsilon_i) - c_s | \tilde{v}(w, \epsilon_i) \geq c_s] - c_m. \end{aligned}$$

Let $\tilde{m}_p$ and $\tilde{r}_0(w)$ denote the new thresholds.

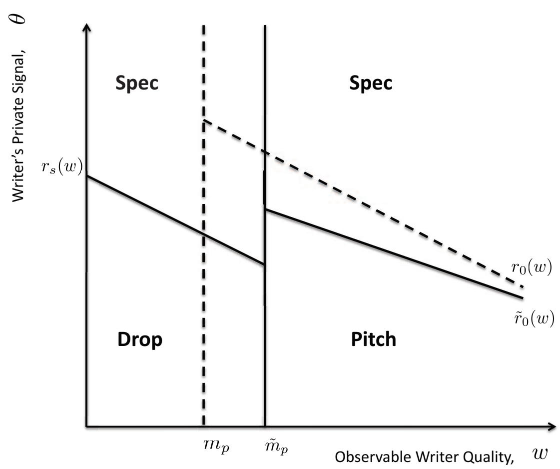
Proposition A1. (buyer does not observe θ until the script is finished) *The writer's choice in equilibrium is qualitatively similar to Proposition 1; the differences are*

(i) *the buyer's meeting threshold for pitch is more strict — i.e., $\tilde{m}_p > m_p$;*

(ii) *for all $w \geq \tilde{m}_p$, the writer chooses to spec more — i.e., $\tilde{r}_0(w) < r_0(w)$. The difference decreases with w — i.e., $\partial(r_0(w) - \tilde{r}_0(w))/\partial w < 0$, and $\lim_{w \to \infty} \tilde{r}_0(w) = r_0(w)$.*

Proof. Proof is omitted and is available upon request. □

FIGURE 7. **Writer's Choice in Equilibrium (buyer does not observe θ until the script is finished)**



Notes: The solid lines describe the writer's equilibrium choice when the buyer does not observe θ until the script is finished. The dashed lines describe the writer's equilibrium choice in the model (the same as Figure 2).

The differences are illustrated in Figure 7. First, given any observable quality, w, the writer prefers to spec (relative to pitch) more — i.e., $\tilde{r}_0(w) < r_0(w)$. This is the classic adverse selection problem: When asymmetric information does not resolve until a script is finished, the writer has an extra incentive to spec in order to differentiate his idea from worse ones. The difference, however, decreases as w increases, and disappears at the limit. In other words, writers of better observable quality suffer less from asymmetric information. Second, the barrier to pitch is higher — i.e., $\tilde{m}_p > m_p$. There are two reasons for this: One is that the average quality of the pitch offered for sale decreases as a result of the adverse selection problem; the other is that there is efficiency loss, because the buyer does not observe the true θ when deciding whether to buy a pitch upon meeting.

C. Data construction

The data were collected from various internet databases: *Done Deal Pro*, *IMDb*, *TheNumbers*, *Hollywood Literary Sale*, and *Who's Buying What*. Data obtained from these various sources were matched using Stata programs or by hand. The following documents in details how the data were constructed.

Step 1: sale data.

(1) I used *Done Deal Pro* as the main source for idea sales. Its "Sales Archive" contains sales from 9/25/1997 until now. Below is an example.

Title:	**Billy Two Sugars**
Logline:	Centers on a low-man-on-the-totem-pole driver for a huge Cuban drug lord in Miami. This man becomes sick and tired of his abusive boss and decides to change his fate by putting together a gang of other lowly down-and-outs in order to rip off the drug lord in order to gain riches and respect.
Writer:	Seth Pearlman
Agency:	Creative Artists Agency
Studio:	New Line Cinema
Price:	Mid six figures
Genre:	Comedy
Logged:	12/8/1999
More:	New Line made a preemptive bid on this spec in which Brett Ratner is attached to direct and produce with Antonio Banderas to star. Bandera's company, Green Moon Prods' Diane Sillan Isaacs will produce with Pearlman co-producing.

(2) I obtained sales from 1/1/1998 to 2/26/2008. About 55% of the sales are adaptations of a source material, such as a novel or a magazine article, or assignment jobs for which the writer is hired to rewrite existing screenplays or to adapt the producer's ideas into a screenplay. These cases can be identified using information in section *More*. For the current paper, I excluded these cases, because the writer(s) listed are not selling their original ideas. This left a sample of approximately 3,522 original ideas sold by writers.

(3) Of the original idea sales, I used sales that took place by 12/31/2003. The reason is that I wanted to leave enough time for the project to be released, so that I could obtain the final outcome of the sales. This resulted in a sample with about 2,019 sales.

(4) As shown in the example above, I could obtain the following information for each sale: (a) title; (b) date of the updates (*Logged*). I define this date as the SALE DATE. The database is updated daily, as soon as the relevant parties make the announcement in the trade press. The announcement in the market is usually very timely, either because the studio wants to preempt a project, or the writer (and his agent) want the positive advertising effect. Thus, it is reasonable to use the date as a proxy for the time of sale; (c) name of the writer(s); (d) name of the intermediaries that represent each writer. More on this later in Step 4; (e) name

of the buyer(s). These can be a studio or (and) a production company; (f) whether the sale is a spec or a pitch. This information is identified using information in section *More*. For example, the sale I gave above says that it is a *spec*. Not all sales have a clear indication of whether it is a spec or a pitch. I complemented the information using two other sources: *Hollywood Literary Sales* and *Who's Buying What*.[54] I identified the same sale by the title, the date of updates, and the writers. Eventually, I was able to complete the sale stage for 80% of the sales, totaling 1,638 sales; (g) other information about the sale: genre, rough price range for part of the sales, whether there are director or actors/actresses attached (in section *More*).

(5) The final sample has 1,638 original ideas sold by writers from 1/1/1998 to 12/31/2003.

Step 2: match writers in the sale data to people on *IMDb*. *Done Deal Professional* has no information about the writers except for their names. I matched the writers to people on *IMDb*, which contains rich information on people's experiences in the industry. I matched the writers in two steps:

(1) I first matched the names by writing Ruby codes. *IMDb* assigns each person a unique identifier. To maximize the accuracy of the matching, I first deleted people who share the same name. In other words, if there were two or more writers with the same name, but different identifiers, I deleted them. Then, I matched the writers in the sale data set with the names on the list from *IMDb*. If there was a match, I identified them as the same person.[55]

(2) For writers in the sale data who were not matched automatically, I matched them by hand.[56] I made the judgement of whether they were the same person according to a couple of criteria. For example, if there were two writers in the team and the other was identified. I searched the films written by the other writer and tracked down the the other writers of the movie. This worked fairly well because writing teams usually tend to be pretty stable.

Eventually, many writers do not match to anyone on *IMDb*. For these writers, I treated them as inexperienced writers, and assigned zero to their writing experience. Though not a hundred percent accurate, I have done my best to guarantee accuracy. For writers that are matched to

[54] *Hollywood Literary Sales* has higher data quality than *Who's Buying What*. However, the updates stop at Dec. 2003. This is one of the reasons why the final sample I use stops at Dec. 2003.

[55] Because *Done Deal Pro* does not assign identifiers to people, I could not tell whether there are different writers that share the same name. There is nothing that I could do about this. However, judging by my experience with the data, it is reasonable to believe that these cases are rare. They are usually differentiated by adding different middle names.

[56] Writers may not be matched automatically because the middle name may or may not be present, or the writers may be deleted because they share the same names with others.

people on *IMDb*, I assigned them the unique identifiers that *IMDb* uses. For the rest, I assigned my own unique identifier to them.

Step 3: match sales to movies on *IMDb*. To obtain the final outcome of the sales, I matched sales to movies on *IMDb*, which lists almost all movies that are released — at least all movies theatrically released in the U.S. The match was done mostly by hand. For each of the 1,638 sales:

(1) I searched *IMDb* by the Title first. If I found movies of the same name that were released after the sale date, I further checked the writer's name. If all matched, I identified them as the same movie. A good thing about *IMDb* is that it lists all the working titles that the project has, which includes the title that the sale is named after.

(2) If I could not match the sale using the Title, I then searched by writer name because the original idea sellers are typically guaranteed a writing credit. Thus, if I could find the writer, I went through all movies that the writer has written and that were released after the sale date, and I identified the movie by similarity of the log-line.

If I could not match the sale by the above two steps, I identified the movies as not released.

Step 4: writer, intermediary, and movie characteristics. For writers that are matched to people on *IMDb*, I obtained information about them as described below. All steps were achieved by Ruby or Stata programs.

Writer characteristics

(1) Downloaded the *filmography by type* webpages for all writers that were matched into html pages. Read the unique identifiers assigned by *IMDb* of all movies in which these writers were involved as a writer, a director, an actor, or a producer.

(2) Downloaded the *combined details* webpages for all movies identified in (1) into html files. Read the following characteristics of each movie: type (such as whether it is a feature-length film),[57] genre, distributor, and release date.

(3) For each writer, made a list of all movies that she/he has been involved in by job (e.g., writer, director). For each of these movies, matcheded the information on type, distributor, genre, and release date obtained in (2). For each movie in which the writer was involved as a writer, created a variable differentiating the screenwriting credits versus credits for source materials.[58]

[57]Type includes "Feature," "Short," "TV" (TV movies, series, and programs), "V" (Video), and "VG" (Video Game).

[58] The former include writing credits with key words "Writer," "Written," "Screenplay," "Adaptation," ("Teleplay"), "Script," "Treatment," or "Story."

(4) Then, I was ready to define alternative measures of observable writer quality. For example, WRITEREXP is defined as the number of screenwriting credits for feature films released by major studios and big production companies in the previous five years

Intermediary characteristics. In Hollywood, a writer may be represented by three types of intermediaries: agents, managers, and lawyers. *Done Deal Pro* records the information in six separate fields if applicable: agent, agency, manager, management firm, lawyer, and law firm. Using the agent/agency as an example, the following paragraph describes how I treat this information in the analysis. The same applies to manager/management firms, and lawyer/law firm.

An agency is the private firm that an agent works for. In the analysis, if two writers are represented by the same agency, but different agent names are listed, I coded the writers as being represented by the same intermediary. The main reason is that an agency typically works as a team; thus, the reputation capital and network connection of individual agents are transferrable within a firm. In addition, an individual agent's name is usually missing from the database when the writer is represented by a sizable agency, as is the case in the given example. There are a few cases where only an agent, not the agency, is listed. This is usually because it is a small agency, consisting of a couple of agents and usually named after the agent. When this was the case, I coded the agent as an individual agency.

Movie characteristics. *IMDb* has comprehensive information on cast and crew. However, the information on the production budget and the box office are not readily available. *TheNumbers* has high-quality business-related data on almost all movies theatrically released in the U.S. after 1995. For movies in the sale data for which I could find a match with movies on *IMDb* (Step 3), I was able to complete 198 of them with positive U.S. box office information from *TheNumbers* or a couple of other sources, including *BoxofficeMojo*. These are the movies that I define as released. 149 of them have information on production budget.

D. Writer's compensation and pricing data

The typical compensation structure that I will explain below is based on a number of industry sources, including *Contracts for the Film & Television Industry* (Litwak, 1998), *The Movie Business Book* (edited by Squire, 2004), *Hollywood dealmaking: negotiating talent agreements* (Appleton and Yankelevits, 2002), and articles from trade magazines and screenwriter forums.

Compensation structure

If the writer sells a pitch, he is typically employed to write the first draft of the screenplay.[59] In most cases, compensation proceeds in steps. The typical price quote one sees in the trade magazines is in the format of "$A against $B." $A is called the front-end in the industry, and is roughly the price for two drafts and a polish. For each step, the corresponding payment is paid half upon commencement and half upon delivery. The buyer cannot deny the payment upon delivery as long as it is a craftsman-like work. The buyer has a period of time, which is typically about two weeks, to decide whether or not to proceed to the next step. During the reading period by the buyer, the writer cannot accept outside assignments that might prevent him from completing the remaining steps. $(B - A) is the credit bonus, which is paid only if the movie is produced. The amount is contingent on whether the writer receives a sole or shared credit.

Table 11 is an example given by John August in his article *Money 101 for screenwriters*.[60] The price of the pitch is "low six figures," which the article explains is $125,000 against $200,000. The table explains the break-down for each step. The delivery periods required for each step seem to be quite standard.[61]

In addition to the front-end and the credit bonus, the writer also gets profit sharing. Typically, the writer could receive between 1 and 5 percent of the net profit.[62] The magnitude of profit participation is also tied to the screenplay credit determined by the Guild. A shared credit generally reduces the writer's profit sharing by half. There are additional compensations for sequels, remakes, and derivative works in television.

Acquisition of a spec can be an outright purchase or an option/purchase agreement.[63] The basic structure of both is similar to a typical pitch contract, such as the purchasing price, which is similar

[59] If the buying company has an agreement with the Writers Guild of America (WGA), it is required that the writer be employed to write the first draft. All major studios and most production companies have agreements with WGA.

[60] John August is the writer of *Charlie and the Chocolate Factory* and *Charlie's Angels*.

[61] Romy Kaufman lists the same delivery periods in her article *The Story Editor* in the *The Movie Business Book* (2004) edited by Jason E. Squire. Mark Litwak, in *Contracts for the Film & Television Industry*, writes that it usually takes one to three months to write the first draft, and the reading period between each step is about 2 weeks.

[62] It has been estimated that only about 5% of movies officially show a net profit, and the "losers" include such blockbuster films as Rain Man, Forrest Gump, Who Framed Roger Rabbit, and Batman, which all took in huge amounts in box office and video sales.

[63] Option/purchase agreements are more common for new writers.

TABLE 11. **An Example of a Step Deal for a Pitch**

Step	Deliver periods	Payment
First draft - Commence	12 weeks	$36,000
First draft - Deliver		$36,000
Second draft - Commence	8 weeks	$9,000
Second draft - Deliver		$9,000
Polish - Commence	8 weeks	$17,500
Polish - Deliver		$17,500
Sole credit bonus	after filming completes	$75,000
Total		$200,000

Notes: This example is from John August's article *Money 101 for screenwriters*.

to $A, the credit bonus, which is $B - $A, and other compensations such as profit participation, additional payments for sequels, remake, and TV spin-offs. In an option/purchase agreement, for example, a writer sells a screenplay for $300,000 against $500,000, a possible break-down is that the writer gets paid immediately $100,000 as the option fee.[64] The rest of the purchasing price ($300,000 - $100,000 = $200,000) is paid when the option is exercised. In the case of rewrites, the writer of the original screenplay is guaranteed to perform the first rewrite. Whether the writer is paid an extra flat fee for the rewrite or the service is applicable to the purchasing price depends on the negotiation.[65] If the movie is produced, and the writer gets a sole screenplay credit, he gets $200,000 credit bonus. If the writer gets a shared screenplay credit, he gets paid half of the credit bonus. If it is an outright purchase, the purchasing price $300,000 is paid immediately upon agreement.

Pricing information in the data

About 57% of the sales have some price information. First, in most cases, the price reported is a range rather than an exact number, such as "low six figures." When prices are available, about 87% of pitches and 84% of specs are quoted with a range. Second, sometimes only the sum of the front-end and the credit bonus ($B) is reported. When prices are available, about 31% pitches and 26% of specs are quoted as one total amount. Finally, for specs, the data do not distinguish between an option/purchase agreement or an outright purchase.

For the purpose of the analysis, I code the price according to the following. First, I replace a price range with a numerical value. I try to be conservative because the agents have the reputation of

[64]There is a standard of one year for an initial option period and extensions of six months or a year for the second and subsequent optional periods. There are no standard as to price, which depends upon what the market is willing to pay. The price paid for extensions of the option period is usually not applicable to the eventual purchasing price.

[65]It seems that typically, the rewriting service is applicable to the purchasing price.

exaggerating the price within the range they can. Table 12 lists the numbers imputed for prices within the range of "six figures." Second, when only the total amount, $B, is reported, I define the front-end and the credit bonus as 60% and 40% of $B.[66]

TABLE 12. **Coding Mechanism for Price Data**

Original Quote	Replaced by
Low six figures	$200,000
Low mid six figures	$350,000
Mid six figures	$450,000
High mid six figures	$600,000
High six figures	$700,000

Using the previous pitch example. The trade magazines reports "low six figures," while the actual compensation is $125,000 against $200,000. According to the coding mechanism explained in the previous paragraph, "low six figures" is coded as $200,000. The front-end is $120,000, and the credit bonus is $80,000.

For quotes like "six figures" and "seven figures," it is hard to pin down a more specific range. Based on the conservative principle, I insert a lower number within the range as the number of digits increases. In particular, I assign 10,000,000 to "eight figures", 2,000,000 to "seven figures", 300,000 to "six figures", 40,000 to "five figures", and 5,000 to "four figures." It is important to note that the results are qualitative robust to alternative imputations based on the same principles.

Though not accurate, the price in the data provides valuable information on an idea's value. Moreover, this coding mechanism treats specs and pitches in the same way, thus, it is reasonable to believe that the coding does not create significant bias for the purpose of this dissertation.

[66]Some high profile sales also resemble the 60%/40% break-down. For example, Terry Rossio and Bill Marsilii are paid $3 million against $5 million for the spec script of *Deja Vu*. Bobby Florsheim and Josh Stolberg are paid $1.5 million against $2.5 million for *The Passion Of The Ark*. David Koepp is paid for $2 million against $3 million for *Panic Room*.