Equilibrium Investment, Share Retention and Voluntary Disclosure Decisions in IPOs in the Presence of Damages for Nondisclosure*

Ronald A. Dye
Kellogg School of Management
Northwestern University

May 27, 2014

1 Introduction

The present paper contains an economic model that examines the equilibrium investment, disclosure, and share retention decisions of an entrepreneur who launches an IPO. While models of IPOs have been a long-standing staple of the finance and accounting literatures (see, e.g., Leland and Pyle [1977], Downes and Heinkel [1986], Trueman [1986], Hughes [1988]), the present analysis contains the novel feature that many of the entrepreneur's pre-IPO optimal actions are driven in part by a concern over the possible liability the entrepreneur may be subject to in the event the entreprenuer fails to disclose value relevant information he knows about his firm at the time of the IPO. (There are other, more traditional forces that influence the entrepreneur's decisions as well, including entrepreneurial moral hazard in his investment choices and the entreprenuer's assumed higher rate of discounting cash flows than outside investors.) While examining penalties for nondisclosure in any market context characterized by information asymmetry is a natural object of study, examining such penalties is particularly appropriate

*This is an early draft. Comments welcome!

in the context of IPOs given the real resource allocation effects surrounding primary issues.

Our analysis yields a variety of comparative statics results, including results concerning how the entrepreneur's pre-IPO investment and share retention decisions are affected by each of: the quality of the private information he receives, as measured by the precision of his information; the probability he receives private information; and the severity of the entrepreneur's penalties for nondisclosure, as measured by the size of a "damages multiplier" (which determines what fraction of investor losses arising from investors having overpaid for the shares they bought as a consequence of the entreprenuer having deliberately withheld information from them must be reimbursed by the entreprenur).

While several of our findings are intuitive, at least three stand out as unexpected. First, we show that, for a broad range of damages multipliers, the entrepreneur's equilibrium disclosure policy is independent of whether he focuses on the "short run" and is only concerned about the immediate effects of his disclosure on the selling price of his shares in the firm or instead, whether he focuses on the "long run" and also accounts for the potential damage payments he may be liable for in the event he decides to withhold his information from investors and that withholding is subsequently detected. Second, we show that small increases in the damages multiplier always *increase* the probability the entrepreneur will withhold information. Third, we show that small (or local) increases in the damages multiplier always (weakly) *increase* the fraction of the firm the entrepreneur sells to outside investors, when the entrepreneur contemplates how is equilibrium share retention decisions affect the entrepreneur's incentives to investment in the firm prior to the IPO. An additional conclusion we obtain is that, contrary to the central finding of Leland and Pyle [1977], increases in the equilibrium fraction of shares retained by the entrepreneur need not imply higher firm value.

In the course of generating these and other results, we obtain what we believe to be new findings about the equilibrium probability that a value-maximizing manager of a firm will disclose an estimate of the firm's value that the manager occasionally privately receives: if this estimate is normally distributed and is an unbiased estimate of the firm's cash flows, we show that the equilibrium probability the manager will disclose this estimate is independent of each of: the mean and variance of his firm's cash flows, the precision of the estimate; and the manager's equilibrium ownership stake in the firm. As part of these findings, we show that no matter what the ex ante probability p the manager privately receives an estimate, the ex ante probability he will disclose the estimate in equilibrium always exceeds $\frac{1}{2} \times p$.

The paper proceeds as follows. The next section, Section 2, introduces the model. Section 3 describes features of the entrepreneur's preferences. Section 4 examines the entrepreneur's optimal disclosure policy, and it contains the general finding regarding the equilibrium disclosure probability mentioned above. Section 5 evaluates the entrepreneur's equilibrium investment decision, taking his share retention decision as given. Sections 6 and 7 respectively study the entrepreneur's optimal share retention decision. Section 8 contains a summary of some of our main findings, and the appendix contains proofs of all results not proven in the text or accompanying footnotes.

2 Base model setup

A risk-neutral entrepreneur ("E") initially owns 100% of a firm that consists of the following decreasing returns to scale production technology. If E privately selects investment $I \geq 0$ today at cost $.5I^2$, then his firm ultimately generates cash flows z, the realization of the random variable $\tilde{z}$, which is normally distributed with mean $m(I)$ and variance $\frac{1}{\tau}$, henceforth written as $\tilde{z}\tilde{\ } N(m(I), \frac{1}{\tau})$. Here, $m(I)$ is some strictly increasing, twice differentiable, weakly concave function representing the undiscounted expected cash flows E's firm generates conditional on E selecting investment I. For much of the analysis, we set $m(I) = w \times I$, for some positive contant w representing the marginal productivity of E's investment, but many of the results below generalize to arbitrary functions $m(\cdot)$ that satisfy these monotonicity, concavity, and differentiability conditions.

The cash flows z E's firm produces are realized some time in the future. E perceives the present value of cash flows z today to be $\delta \times z$ for some constant $\delta \in (0,1)$, E's discount rate. This discount rate reflects E's impatience. E is assumed to be more impatient than outside investors in that outside investors value the firm's ultimate cash flows z today at $\delta_i z$ for some constant δ_i where $\delta < \delta_i \leq 1$ reflecting the investors' lower discount rate. (This difference between E's and outside investors' assumed impatience is natural, and can be motivated by E's life cycle or liquidity demands.) To simplify notation, and with no loss of generality, we set $\delta_i = 1$ throughout the following.

In this base model, we assume that concurrently with selecting investment I, E chooses what fraction $f \in [0,1]$ of his firm to sell to outside investors.[1] E could set $f = 0$ and keep all the cash flows z his firm produces for himself, or alternatively, E could set f at some positive value, and thereby reliquish property rights over fraction f of the firm's cash flows in return for receiving an "up front" payment from outside investors, as described further below. When E sells fraction $f > 0$ of his firm to outside investors, we interpret this as constituting an IPO. The fraction f is public information.

Before E sells any fraction of the firm to outside investors, with probability $p \in (0,1)$ E privately receives an estimate $\tilde{v}$ of the firm's cash flows $\tilde{z}$. This estimate $\tilde{v}$ is taken to be unbiased and given by

$$\tilde{v} = \tilde{z} + \tilde{\varepsilon}, \text{ where } \tilde{\varepsilon}\,\tilde{}\,N(0, \frac{1}{r}) \text{ is independent of } \tilde{z}. \tag{1}$$

When E selects investment I, we write $g(v|I)$ and $G(v|I)$ for the density and cdf of $\tilde{v}$ respectively. Given I, it is apparent that the prior distribution of $\tilde{v}$ is

[1] In section 7 below, we evaluate the consequences of dropping the assumption that E selects f "up front."

$\tilde{v}\tilde{\ }N(m(I),\sigma^2)$, where $\sigma^2 \equiv \frac{1}{\tau}+\frac{1}{r}=\frac{r+\tau}{r\tau}$. With probability $1-p$, E receives no such estimate.

If E receives information v, E must decide whether to disclose v or not.[2] As is now standard in the voluntary disclosure literature (cf. Dye [1985]), we assume: 1. if E discloses information, it must be truthful; 2. if E receives no information, he necessarily makes no disclosure, and 3. E cannot truthfully disclose that he has received no information.

After the IPO is complete, if E disclosed nothing prior to the IPO, then a fact finder (investor, reporter, auditor, etc.) is presumed to undertake an investigation and, if the reason E disclosed nothing turns out to be that E withheld information, the fact finder with probability $q \in (0,1)$ both detects that E withheld information and what the withheld information was. In the latter event, E is forced to pay damages to investors who purchased shares in E's firm. These damage payments, as we describe in detail below, are taken to be proportional to the amount investors overpaid for their share of the firm, where "overpaid" is determined by the difference between what they actually paid and what they would have paid had E disclosed the information in his possession. The fact finder fails to discover that E withheld information (with probability $1-q$), but the fact finder never wrongly asserts E withheld information.

Finally, the firm's realized cash flows z occur and are apportioned between E and outside investors based on the fraction f of shares E sold to the investors.

In summary, the time line of the base model is as follows: E starts out by selecting investment I and what fraction f, if any, of his firm to sell to investors, and then waits to receive private information about the firm's value. If E receives no such information, he necessarily makes no disclosure; if E receives information, he decides whether he is better off disclosing the information to

[2] The model that follows could be augmented by positing that, in addition to selectively making voluntary disclosures, E's firm is obligated to make certain prescribed mandatory disclosures too, corresponding to disclosures in proxy statements, S-1 disclosures, and the like. We do not formally incorporate such mandatory disclosures in the model.

investors or withholding that information. Then the IPO takes place. If E withheld information from outside investors before the IPO, then E is subject to damage payments in the event the fact finder subsequently detects his withholding. Finally, the firm's realized cash flows occur, and both E and outside investors consume those cash flows in proportion to their previously determined ownership stakes in the firm.

3 E's preferences

To write E's preferences explicity, we start by distinguishing between E's actual investment choice and investors' conjecture about his investment choice; the notation I refers to an actual investment choice E makes, I^* refers to his optimal (actual) investment choice, and $\hat{I}$ refers to investors' conjecture about E's investment choice. We let $f \times P(v|\hat{I})$ represent the amount outside investors pay E in the event E sells investors fraction f of the firm to them after disclosing that $\tilde{v} = v$, when the investors believe E selected investment $\hat{I}$. Since we assume there are lots of risk neutral, competitive, and homogenously informed outside investors, $f \times P(v|\hat{I})$ can be described alternatively as the expected cash flows these outside investors expect to receive from buying fraction f of the firm. Hence, $P(v|\hat{I})$ is investors' assessments of the amount of the expected cash flows they expect the entire firm to generate. Similarly, we let $f \times P^{nd}(\hat{I})$ represent the amount investors pay E for fraction f of the firm when E makes no disclosure.

When E receives, but decides to withhold, v from investors, the difference $f \times (P^{nd}(\hat{I}) - P(v|\hat{I}))$ constitutes the losses investors incur in expectation as a consequence of E having withheld information v from them: it is the difference between what they paid E for fraction f of the firm and the expected cash flows they will actually get from purchasing this fraction of the firm, based on information in E's possession. In the event the fact finder discovers E withheld information from investors, we posit E must pay investors some multiple β of these damages, specifically, $f \times \beta \times (P(v|\hat{I}) - P^{nd}(\hat{I}))$, where $\beta \geq 0$ is a "damages multiplier" determining the fraction of damages E must reimburse investors for. In the following, we impose a bound how large the damages multiplier can be: we restrict β so that

$$q(1+\beta) < 1. \tag{2}$$

There is substantial documentation that β is much smaller than 1 in practice: investors who purchase a firm's shares while its managers withheld unfavorable information from them are far from being "made whole," so the bound on β implied by (2) will not be binding in practice.[3]

With these prefatory comments complete, we can now describe E's preferences explicitly. We temporarily simplify notation by suppressing reference to investors' conjectures about the firm's investment $\hat{I}$ and hence write P^{nd} in place of $P^{nd}(\hat{I})$ and $P(v)$ in place of $P(v|\hat{I})$. If E's actual initial investment is I, E sells fraction f of his firm to outside investors, the realized value of $\tilde{v}$ turns out to be v, and E discloses v; then (the present value of) E's utility is given by:

$$\delta(1-f) \times E[\tilde{z}|v, I] + f \times P(v) - .5I^2. \tag{3}$$

This is the sum of: (1) the present value of the expected cash flows of the firm (evaluated at E's discount rate) for the fraction of the firm E retains conditional on E's information (v and I), $\delta(1-f) \times E[\tilde{z}|v, I]$; 2. the payment E receives from outside investors for selling fraction f of the firm to them after disclosing v, $f \times P(v)$, net of: 3. the cost of E's initial investment, $.5I^2$. In this expression, note that we do not discount the payment $f \times P(v)$ E receives from outside investors, since we assume this payment (and hence the IPO) occurs sufficiently soon after E's initial investment (and sufficiently before the firm's cash flows realize its value) that no discounting is warranted.

[3] See, e.g., Ryan and Simmons [2009].

Similarly, if E selects investment I and $\tilde{v}$ turns out to have the value v, and either E does not learn v, or E learns, but withholds, v from investors and the fact finder fails to subsequently detect E's withholding (so no damage payments are assessed), then (the present value of) E's utility is given by:

$$\delta(1-f) \times E[\tilde{z}|v, I] + f \times P^{nd} - .5I^2. \quad (4)$$

Finally, if E selects investment I, $\tilde{v}$ turns out to have the value v; E learns v and withholds that information from investors, and this withholding is subsequently detected by the fact finder (so E is assessed damage payments), then (the present value of) E's utility is given by:

$$\delta(1-f) \times E[\tilde{z}|v, I] + f \times P(v) - f \times \beta(P^{nd} - P(v)) - .5I^2. \quad (5)$$

If E elects to withhold v from investors, E will not know at that time whether the fact finder will subsequently detect E's withholding. So, at that time E's expected discounted utility will be a weighted average of (4) and (5), with weights $1-q$ and q respectfully, leading to:

$$\delta(1-f) \times E[\tilde{z}|v, I] + (1-q)fP^{nd} + q(fP(v) - f\beta(P^{nd} - P(v))) - .5I^2. \quad (6)$$

4 Preliminary analysis of E's decision to withhold or disclose v

When E learns $\tilde{v} = v$, whether E is better off withholding or disclosing v is determined by whether (6) is bigger or smaller than (3). Since the both of the terms $\delta(1-f) \times E[\tilde{z}|v, I]$ and $-.5I^2$ appear in each of (3) and (6), E is better off withholding v iff the following inequality holds:[4]

$$(1-q) \times fP^{nd} + qf \times [P(v) - \beta(P^{nd} - P(v))] > fP(v). \quad (7)$$

[4] For the sake of specificity, we assume that if E is indifferent between withholding and disclosing, E discloses (though, since $\tilde{v}$ is continuously distributed, such cases of indifference are zero probability events, and hence are irrelevant to any "ex ante" expected utility calculations).

This last inequality can be rewritten as:

$$(1 - q - q\beta)fP^{nd} + (q + q\beta)fP(v) > fP(v).$$

When the bound (2) holds, it is clear that this last inequality is equivalent to:

$$P^{nd} > P(v). \tag{8}$$

This result is surprising. It shows that even though we take E's objective function to be "long run," in so far as E accounts for the expected cost of the damage payments he may be liable for at the time he makes his disclosure decision, E's optimal decision to disclose or withhold v is the same as if E's objective function were "short run," with E being concerned only with maximizing the price of the firm circa the time he makes his disclosure decision - without concern for the possible damage payments he may be subject to if he withholds information. The result obtains because both "10b-5 like" damage payments and the optimal short run disclosure decision are based on the difference between the "no disclosure" price of a share and the price of a share had E disclosed his information. We summarize this finding in the following lemma.

Lemma 1 *When (2) holds, E's optimal "long run" disclosure policy coincides with E's optimal "short run" disclosure policy.*

To proceed, we next provide an explicit expression for $P(v) = P(v|\hat{I})$, the value outside investors attach to the whole firm when E discloses v. Note that, before v is disclosed, $N(m(\hat{I}), \frac{1}{\tau})$ constitutes investors' prior beliefs about the distribution of $\tilde{z}$ given $\hat{I}$. Recalling (1), we know that, given $\tilde{z} = z$, the distribution of the estimate $\tilde{v}$ may be expressed as $\tilde{v}|_z \tilde{} N(z, \frac{1}{r})$. So, according to standard updating formulas for normal distributions (see, e.g., DeGroot [1970]), investors' posterior assessment of $\tilde{z}$ after seeing v is also normally distributed with mean $E[\tilde{z}|v, \hat{I}] = \frac{\tau\hat{m}(\hat{I})+rv}{\tau+r}$. Hence, given investors' assumed risk-neutrality and recalling the normalization $\delta_i = 1$, we conclude:

$$P(v|\hat{I}) = E[\tilde{z}|v, \hat{I}] = \frac{\tau m(\hat{I}) + rv}{\tau + r}. \tag{9}$$

Combining this with the inequality (8) where $P^{nd} = P^{nd}(f)$, we conclude that E's optimal "no disclosure" set is given by:

$$ND(f) = \{v | P^{nd}(f) > \frac{\tau m(\hat{I}) + rv}{\tau + r}\},$$

or equivalently, there exists a cutoff $v^c = v^c(f)(= \frac{P^{nd}(f)(\tau+r) - \tau m(\hat{I})}{r})$ so that the "no disclosure" set is described by:

$$ND(f) = \{v | v^c(f) > v\}. \tag{10}$$

5 Characterization of the equilbrium disclosure policy

We next address the issue of: if E makes no disclosure and investors believe E selects investment $\hat{I}$ and employs the disclosure cutoff $v^c(f)$ when he decides to sell fraction f of the firm to them, what are the expected cash flows, inclusive of expected damage payments, investors expect to receive? To answer this question, we note that, given E's nondisclosure, there are three distinct possible events of interest: 1. the reason E made no disclosure was that E received no information; 2. the reason E made no disclosure was that E withheld information, and subsequently the fact finder fails to detect that E withheld information; 3. the reason E made no disclosure was that E withheld information, and subsequently the fact finder detects that E withheld information - in which case investors who bought shares in the firm at the time of the IPO are are entitled to damage payments.

To analyze these three possibilities, we first simplify notation by once again holding investors' perceptions $\hat{I}$ of E's investment temporarily fixed, and so we write the cdf and density of of $\tilde{v}$ as $G(v)$ and $g(v)$ in place of $G(v|\hat{I})$ and $g(v|\hat{I})$ respectively, and we also write $\hat{m}$ in place of $m(\hat{I})$.

It follows that investors' ex ante perceptions of the probability E will make no disclosure is given by $1 - p + pG(v^c)$. This is the probability E receives no information plus the probability that E receives information that is below the cutoff v^c. Hence, applying Bayes' Rule, we conclude that the probability of the event that E received no information, conditonal on E making no disclosure, i.e., the probability of event (1) as described in the opening paragraph of this section conditional on E making no disclosure, is given by:

$$\frac{1-p}{1-p+pG(v^c)}. \tag{11}$$

Likewise, the probability of events (2) and (3) above as perceived by investors, given E makes no disclosure, are respectively:

$$\frac{p(1-q)G(v^c)}{1-p+pG(v^c)} \tag{12}$$

and

$$\frac{pqG(v^c)}{1-p+pG(v^c)}. \tag{13}$$

Next notice that, conditional on no disclosure and also conditional on event (1), investors expect to receive cash flows obtained from buying fraction f of the firm in the amount:

$$fE[\tilde{z}|\hat{I}] = f \times \hat{m}. \tag{14}$$

Investors' perceptions of the expected cash flows they will receive conditional on no disclosure and event (2) are given by:

$$fE[\tilde{z}|\tilde{v} < v^c] = f \times E[\frac{\tau\hat{m} + r\tilde{v}}{\tau + r}|\tilde{v} < v^c]. \tag{15}$$

To calculate investors' perceptions of the expected cash flows they will receive conditional on no disclosure and event (3), first fix a particular $v < v^c$ and note that if the fact finder subsequently discovers that E withheld this v, then the cash flows investors expect to receive, net of damage payments, are given by:

$$fE[\tilde{z}|v,\hat{I}] + f\beta(P^{nd} - E[\tilde{z}|v,\hat{I}]) = f \times \left[\frac{\tau\hat{m} + rv}{\tau + r} + \beta(P^{nd} - \frac{\tau\hat{m} + rv}{\tau + r})\right].$$

So, conditional *only* on no disclosure and event (3) (but not also conditioning on a particular $v < v^c$), the cash flows investors expect to receive, net of expected damage payments, are given by:

$$f \times E[\frac{\tau\hat{m} + r\tilde{v}}{\tau + r} + \beta(P^{nd} - \frac{\tau\hat{m} + r\tilde{v}}{\tau + r})|\tilde{v} < v^c]. \tag{16}$$

Thus, the total expected cash flows investors expect to obtain, conditional only on no disclosure by E, consists of the sum of the products of (11) and (14), (12) and (15), (13) and (16), i.e., consists of:

$$f \times \frac{(1-p)\hat{m} + p(1-q)G(v^c)E[\frac{\tau\hat{m}+r\tilde{v}}{\tau+r}|\tilde{v} < v^c] + pqG(v^c)E[\frac{\tau\hat{m}+r\tilde{v}}{\tau+r} + \beta(P^{nd} - \frac{\tau\hat{m}+r\tilde{v}}{\tau+r})|\tilde{v} < v^c]}{1 - p + pG(v^c)},$$

which may be rewritten as:

$$f \times \frac{(1-p)\hat{m} + pG(v^c)E[\frac{\tau\hat{m}+r\tilde{v}}{\tau+r}|\tilde{v} < v^c] + pqG(v^c)\beta E[P^{nd} - \frac{\tau\hat{m}+r\tilde{v}}{\tau+r}|\tilde{v} < v^c]}{1 - p + pG(v^c)}. \tag{17}$$

Competition among investors to buy fraction f of the firm's shares when E makes no disclosure will drive the price $f \times P^{nd}$ E receives for those shares to (17). That is:

$$f \times P^{nd} = (17). \tag{18}$$

Moreover, in order for the value $v = v^c$ to qualify as the cutoff defining E's no disclosure set (10), it must be true that when $\tilde{v} = v^c$, E is indifferent between disclosure and nondisclosure, implying from (9) that v^c and P^{nd} must be related through:

$$f \times \frac{\tau\hat{m} + rv^c}{\tau + r} = f \times P^{nd}. \tag{19}$$

Putting (18) and (19) together, it follows that in equilibrium, the cutoff v^c must satisfy the equation:

$$f \times \frac{\tau\hat{m} + rv^c}{\tau + r} = f \times \frac{(1-p)\hat{m} + pG(v^c)E[\frac{\tau\hat{m}+r\tilde{v}}{\tau+r}|\tilde{v} < v^c] + pqG(v^c)\beta E[\frac{\tau\hat{m}+rv^c}{\tau+r} - \frac{\tau\hat{m}+r\tilde{v}}{\tau+r}|\tilde{v} < v^c]}{1 - p + pG(v^c)}. \tag{20}$$

Rewrite (20) using the definition of conditional expectations and the obvious fact that $\hat{m} = \frac{\tau\hat{m}+r\hat{m}}{\tau+r}$ as:

$$\frac{\tau\hat{m} + rv^c}{\tau + r} = \frac{(1-p)\frac{\tau\hat{m}+r\hat{m}}{\tau+r} + p\int^{v^c} \frac{\tau\hat{m}+rv}{\tau+r} g(v)dv + p\frac{rq\beta}{\tau+r}\int^{v^c}(v^c - v)g(v)dv}{1 - p + pG(v^c)},$$

or alternatively as:

$$\frac{\tau\hat{m}+rv^c}{\tau+r}=\frac{(1-p+pG(v^c))\frac{\tau\hat{m}}{\tau+r}+\frac{r}{\tau+r}\left((1-p)\hat{m}+p\int^{v^c}vg(v)dv+pq\beta\int^{v^c}(v^c-v)g(v)dv\right)}{1-p+pG(v^c)}. \tag{21}$$

The following theorem links the equilibrium cutoff v^c that solves (21) to a "standardized" cutoff x^c defined in terms of the density $\phi(\cdot)$ and cdf $\Phi(\cdot)$ of a standard normal random variable, call it $\tilde{x}$. In the statement of the theorem, we let $\alpha \equiv 1-q\beta$.

Theorem 2 *If $v^c = v^c(\hat{m},\sigma)$ solves (21), i.e., if $v^c(\hat{m},\sigma)$ defines the equilbrium cutoff when investors perceive the mean of the firm's cash flows (and also the mean of $\tilde{v}$) to be $\hat{m}$ and the variance of $\tilde{v}$ is σ^2, then*

$$v^c(\hat{m},\sigma)=\sigma x^c+\hat{m}, \tag{22}$$

where $x = x^c$ is the unique solution to the equation

$$x(1-p+\alpha p\Phi(x))+\alpha p\phi(x)=0. \tag{23}$$

This theorem has several consequences. First, (22) provides two immediate and general comparative statics: the equilibrium cutoff v^c increases in lock step with investors' perceptions of the mean of the firm's cash flows, i.e., $\frac{\partial v^c(\hat{m},\sigma)}{\partial \hat{m}}=1$, and the equilibrium cutoff v^c increases linearly in the standard deviation σ of $\tilde{v}$, specifically: $\frac{\partial v^c(\hat{m},\sigma)}{\partial \sigma}=x^c$. But, perhaps more importantly, the theorem establishes that the normalized equilbrium disclosure cutoff x^c is independent of each of: E's equilibrium investment choice $I^*(f)$, the variance $\frac{1}{\tau}$ of the firm's cash flows $\tilde{z}$, the precision r of the estimate $\tilde{v}$, E's discount rate δ, and the fraction f of the firm E retains. Therefore, the equilibrium probability $\Phi(x^c)$ that E will withhold (and hence also the equilibrium probability $1-\Phi(x^c)$ that E will disclose) the information he receives is also independent of all these factors.[5]

[5] Note that there is no contradiction between the assertion that the *probability* E discloses information is independent of each of these factors and the observation that the equilibrium cutoff v^c *changes* with these factors, since the cdf $G(v|I)$ of the random variable $\tilde{v}$ changes in a way that, combined with the equilibrium change in the cutoff v^c, leaves the probability E will withhold information unchanged.

This robustness of the equilibrium disclosure probability to all these factors (in the presence of unbiased and normally distributed estimates of the firm's cash flows) is, as far as we are aware, a new result in the literature on voluntary disclosures.

The following corollary records several additional findings regarding the normalized equilibrium cutoff x^c.

Corollary 3 *The unqiue x^c that solves (23):*

(i) is negative: $x^c < 0$;

(ii) is strictly decreasing in $p : \frac{\partial x^c}{\partial p} < 0$;

(iii) is strictly increasing in $\beta : \frac{\partial x^c}{\partial \beta} > 0$;

(iv) is strictly increasing in $q : \frac{\partial x^c}{\partial q} > 0$;

(v) is strictly decreasing in α*:* $\frac{\partial x^c}{\partial \alpha} < 0$;

(vi) implies that the ex ante probability E will make a disclosure is at least $p/2$.

Before discussing these comparative statics individually, we make two observations that apply to all of them. First, since the cdf Φ is monotone increasing, each of these comparative statics is also a comparative static about the equilibrium probability that E does *not* make a disclosure in equilibrium. Thus, for example, since the corollary asserts that x^c increases in β, it follows that the probability that E makes no disclosure in equilibrium also increases in β.[6] Second, each of these comparatives statics is also a comparative static about how the equilibrium cutoff v^c changes in any parameter, in view of the monotone relationship between x^c and v^c described in Theorem 2, namely that $v^c(\hat{m}, \sigma) = \sigma x^c + \hat{m}$. Thus, for example, since the corollary asserts that in equi-

[6] In the case of the parameter p, the ex ante probability E receives information, this last statement in the text is also true, but it requires the following extra observation to confirm: the ex ante probability that E does not make a disclosure is $1 - p + pG(x^c(p))$. Thus, the marginal effect on this probability of a small increase in p is: $\frac{\partial}{\partial p}(1 - p + pG(x^c(p))) = -1 + G(x^c(p)) + pg(x^c(p))\frac{\partial x^c(p)}{\partial p} = -(1 - G(x^c(p))) + pg(x^c(p))\frac{\partial x^c(p)}{\partial p}$. Since the corollary asserts that $\frac{\partial x^c(p)}{\partial p} < 0$, we conclude that $\frac{\partial}{\partial p}(1 - p + pG(x^c(p))) < 0$ too.

librium x^c increases in β, it follows that the equilibrium cutoff $v^c(\hat{m}, \sigma)$ also increases in β.

Now, as to the individual conclusions of the corollary: part (i) implies that the equilibrium cutoff $x^c$is below the prior mean of $\tilde{x}$ (and hence the equilibrium cutoff v^c is below the prior mean of $\tilde{v}$) and part (ii) shows that these cutoffs are decreasing in the prior probability E receives information. Thus, from these results, we know that any estimate v E receives above the prior mean will be disclosed, and the higher the ex ante probability p E receives information, the higher the probability E will disclose the information E receives *conditional* on receiving the information. These two results are generalizations of previously obtained results concerning equilibrium cutoff disclosure policies obtained by Dye [1985] and Jung and Kwon [1988], now extended to situations where the value-maximizing firm (or E) is subject to potential damage payments in the presence of a normally distributed estimate of the firm's cash flows.

The result of part (iii) of the corollary that x^c increases in β might at first blush seem so counterintuitive as to be wrong, as it implies that an increase in the damages multiplier β *reduces* the probability E will disclose the information he receives. Since the damages payment $\beta f(P^{nd} - P(v))$ can be viewed as a penalty for nondisclosure, and the damages multiplier β can be viewed as a parameter for this penalty, the result runs contrary to the intuition that an increase in a penalty for nondisclosure should encourage more, not less, disclosure. But, the result has an easy explanation. Recall from Lemma 1 that, as long as the inequality in (2) is maintained, E's optimal (long run) disclosure policy is determined by the (short) run comparison of P^{nd} and $P(v)$ in (8). Clearly, the parameter β does not appear explicitly in (8); but β does appear implicitly in (8) through P^{nd}. P^{nd} increases in β because as β increases, outside investors who purchase shares in the firm when E makes no disclosure purchase shares of a firm that will give them bigger damages payments in the event the fact finder catches E withholding information from investors. Since

P^{nd} increases in β, it follows that inequality (2) will hold for more values of v as β increases, and hence E will withhold the information he receives more often.

This same logic applies to explain part iv of the corollary, i.e., why an increase in the efficacy of the fact finder, in the form of an increase in the probability q that the fact finder detects that E withheld information, also results in an increase in the probability E withholds information: increasing q has the effect of increasing the no disclosure price because it increases the probability investors will receive damages.

Part (v) of the Corollary are generalizations of parts iii and iv, given $\alpha's$ definition (recall $\alpha = 1 - q\beta$).

Part (vi) yields the robust conclusion that, regardless of the probability p the firm receives information, the ex ante probability E will disclose the information exceeds $p/2$. This follows directly from the observation of part i of the corollary that $x^c < 0$, along with the symmetry of the density of a standard normal random variable around $z = 0$ (which implies that $p(1 - \Phi(x^c)) > p/2$).

6 E's equilibrium investment choice for a *fixed* share retention level f

Next, we concern ourselves with how E makes his initial investment choice, when investors believe E has selected investment $\hat{I}$ upon deciding to sell fraction f of his firm to them. We temporarily write m in place of $m(I)$ and and $\hat{m}$ in place of $m(\hat{I})$, and we recall that the value $f \times P^{nd}(\hat{I})$ investors attach to buying fraction f of the firm when E makes no disclosure is also equal to $f \times \frac{\tau\hat{m}+rv^c}{\tau+r}$ for the equilibrium cutoff $v^c = v^c(\hat{I}(f))$ that solves (20).

We assert that E's "ex ante" objective function is given by:

$$OBJ \equiv (1 - f)\delta m - .5I^2 + f \times \Psi(I, \hat{I}), \tag{24}$$

where

$$\Psi(I,\hat{I}) \equiv (1-p+pG(v^c|I))\frac{\tau\hat{m}+rv^c(\hat{I})}{\tau+r} + \quad (25)$$
$$p\int_{v^c(\hat{I})}^{\infty}\frac{\tau\hat{m}+rv}{\tau+r}g(v|I)dv - pq\beta\int_{-\infty}^{v^c(\hat{I})}(\frac{\tau\hat{m}+rv^c(\hat{I})}{\tau+r}-\frac{\tau\hat{m}+rv}{\tau+r})g(v|I)dv.$$

Here, $f \times \Psi(I,\hat{I})$ is the price E anticipates receiving from investors when they purchase fraction f of the firm from him net of the expected damage payments he subsequently expects to pay investors. To see this, note that if E receives no information, or if he receives information v below v^c, he makes no disclosure and so collects the "no disclosure" price $f \times P^{nd} = f \times \frac{\tau\hat{m}+rv^c}{\tau+r}$ from investors; since the ex ante probability of no disclosure is $1-p+pG(v^c|I)$, the contribution of $f \times \frac{\tau\hat{m}+rv^c}{\tau+r}$ to E's expected utility is:

$$f(1-p+pG(v^c|I)) \times \frac{\tau\hat{m}+rv^c(\hat{I})}{\tau+r} \quad (26)$$

If E receives information $v \geq v^c$, E is better off disclosing it rather than staying silent. Since the conditional expected value of such good news is $fE[\frac{\tau\hat{m}+r\tilde{v}}{\tau+r}|\tilde{v} > v^c, I]$, and since the ex ante probability of getting such good news is $p(1-G(v^c|I))$, it follows that this contribution to E's expected utility is:

$$p(1-G(v^c|I))fE[\frac{\tau\hat{m}+r\tilde{v}}{\tau+r}|\tilde{v} > v^c, I] = p\int_{v^c}^{\infty}\frac{\tau\hat{m}+rv}{\tau+r}g(v|I)dv. \quad (27)$$

Finally, recall that the expected cost of E's damage payments, conditional on having to pay them, is given by $f\beta E[\frac{\tau\hat{m}+rv^c}{\tau+r} - \frac{\tau\hat{m}+rv}{\tau+r}|\tilde{v} < v^c, I]$. Since the ex ante probability E will be obliged to make these damage payments is $pqG(v^c|I)$, the contribution of these damage payments to E's expected utlity is:

$$fpqG(v^c|I)\beta E[\frac{\tau\hat{m}+rv^c}{\tau+r} - \frac{\tau\hat{m}+rv}{\tau+r}|\tilde{v} < v^c, I] = fpq\beta\int_{-\infty}^{u^c}(\frac{\tau\hat{m}+rv^c}{\tau+r} - \frac{\tau\hat{m}+rv}{\tau+r})g(v|I)dv. \quad (28)$$

Summing (26) and (27) and subtracting (28) leads to expression (25).

To characterize E's optimal investment choice, consider the first-order condition associated with maximizing E's objective function (24) with respect to I.

This first-order condition is given by:

$$0 = (1-f)\delta m'(I) - I + f \times \frac{\partial \Psi}{\partial I}. \tag{29}$$

Here, $f \times \frac{\partial \Psi}{\partial I}$ is E's perception of the sensitivity of the net ("net" of damage payments) proceeds E expects to receive from selling fraction f of the firm to a marginal increase in I. Recalling $\alpha \equiv 1 - q\beta$, one can show (see the accompanying footnote[7]) that, in equilibrium - when $I = \hat{I}$—this sensitivity is

[7] To see this, we start by showing that Ψ can be rewritten as:

$$\begin{aligned} \Psi(I,\hat{I}) \quad = \quad & (1-p)\frac{\tau\hat{m}+rv^c(\hat{I})}{\tau+r} + \\ & p(1-q\beta)\int_{-\infty}^{\infty} \max\{\frac{\tau\hat{m}+rv^c(\hat{I})}{\tau+r}, \frac{\tau\hat{m}+rv}{\tau+r}\}g(v|I)dv + pq\beta\frac{\tau\hat{m}+rm}{\tau+r}. \end{aligned} \tag{30}$$

To see this, first note that:

$$\begin{aligned} \int_{-\infty}^{v^c}(\frac{\tau\hat{m}+rv^c}{\tau+r} - \frac{\tau\hat{m}+rv}{\tau+r})g(v|I)dv &= \int_{-\infty}^{\infty}\max\{\frac{\tau\hat{m}+rv^c}{\tau+r} - \frac{\tau\hat{m}+rv}{\tau+r}, 0\}g(v|I)dv \\ &= \int_{-\infty}^{\infty}\max\{\frac{\tau\hat{m}+rv^c}{\tau+r} - \frac{\tau\hat{m}+rv}{\tau+r}, \frac{\tau\hat{m}+rv}{\tau+r} - \frac{\tau\hat{m}+rv}{\tau+r}\}g(v|I)dv \\ &= \int_{-\infty}^{\infty}\left(\max\{\frac{\tau\hat{m}+rv^c}{\tau+r}, \frac{\tau\hat{m}+rv}{\tau+r}\} - \frac{\tau\hat{m}+rv}{\tau+r}\right)g(v|I)dv \\ &= \int_{-\infty}^{\infty}\max\{\frac{\tau\hat{m}+rv^c}{\tau+r}, \frac{\tau\hat{m}+rv}{\tau+r}\}g(v|I)dv - \int_{-\infty}^{\infty}\frac{\tau\hat{m}+rv}{\tau+r}g(v|I)dv \\ &= \int_{-\infty}^{\infty}\max\{\frac{\tau\hat{m}+rv^c}{\tau+r}, \frac{\tau\hat{m}+rv}{\tau+r}\}g(v|I)dv - \frac{\tau\hat{m}+rm}{\tau+r}. \end{aligned} \tag{31}$$

Also notice that:

$$\begin{aligned} \frac{\tau\hat{m}+rv^c}{\tau+r}G(v^c|I) + \int_{v^c}^{\infty}\frac{\tau\hat{m}+rv}{\tau+r}g(v|I)dv &= \int_{-\infty}^{v^c}\frac{\tau\hat{m}+rv^c}{\tau+r}g(v|I)dv + \int_{v^c}^{\infty}\frac{\tau\hat{m}+rv}{\tau+r}g(v|I)dv \\ &= \int_{-\infty}^{\infty}\max\{\frac{\tau\hat{m}+rv^c}{\tau+r}, \frac{\tau\hat{m}+rv}{\tau+r}\}g(v|I)dv, \end{aligned}$$

so upon rearrangment, we conclude:

$$\int_{v^c}^{\infty}\frac{\tau\hat{m}+rv}{\tau+r}g(v|I)dv = \int_{-\infty}^{\infty}\max\{\frac{\tau\hat{m}+rv^c}{\tau+r}, \frac{\tau\hat{m}+rv}{\tau+r}\}g(v|I)dv - \frac{\tau\hat{m}+rv^c}{\tau+r}G(v^c|I). \tag{32}$$

Substituting the expressions (31) and (32) into (25) establishes (30).

Next, by appealing to Lemma 8(vi) in the appendix, we observe that: $\int_{-\infty}^{\infty}\max\{v^c, v\}g_I(v|I)dv = m'(I)(1-G(v^c|I))$. Since $\int_{-\infty}^{\infty} g_I(v|I)dv = 0$, it follows that:

$$\begin{aligned} & \int_{-\infty}^{\infty}\max\{\frac{\tau\hat{m}+rv^c}{\tau+r}, \frac{\tau\hat{m}+rv}{\tau+r}\}g_I(v|I)dv = \int_{-\infty}^{\infty}\left(\frac{\tau\hat{m}}{\tau+r} + \frac{r}{\tau+r}\max\{v^c, v\}\right)g_I(v|I)dv \\ = \; & \int_{-\infty}^{\infty}\frac{\tau\hat{m}}{\tau+r}g_I(v|I)dv + \frac{r}{\tau+r}\int_{-\infty}^{\infty}\max\{v^c, v\}g_I(v|I)dv = \frac{r}{\tau+r}\int_{-\infty}^{\infty}\max\{v^c, v\}g_I(v|I)dv \\ = \; & \frac{rm'(I)(1-G(v^c|I))}{\tau+r}. \end{aligned}$$

given by

$$\frac{\partial \Psi(I, I)}{\partial I} = m'(I) \times X,$$

where

$$X \equiv \frac{pr}{\tau + r}(1 - \alpha \Phi(x^c)). \tag{33}$$

Hence, the first-order condition (29) can be rewritten as:

$$0 = (1 - f)\delta m'(I) - I + f m'(I) X.$$

Hence, when $m(I) = w \times I$, we conclude:

Theorem 4 *For given $f \in [0, 1]$, E's equilibrium investment level (when $\hat{I}(f) = I^*(f)$) is given by:*

$$I^*(f) = \delta w + f w \times (X - \delta). \tag{34}$$

The theorem is best understood in terms of its comparative statics implications, summarized in the following cororllary.

Corollary 5 *(i) $I^*(f)$ is strictly linearly increasing (decreasing) in f iff $X > \delta$ (resp., $X < \delta$);*

(ii) for fixed f, $I^*(f)$ is:

(iia) increasing in r;

(iib) decreasing in τ;

(iic) increasing in β;

(iid) increasing in q; and

(iie) decreasing in α;

(iif) increasing in δ.

Hence, differentiating (30) with respect to I yields:

$$\begin{aligned} \frac{\partial \Psi}{\partial I} &= p(1 - q\beta)\frac{rm'(I)(1 - G(v^c|I))}{\tau + r} + pq\beta\frac{rm'(I)}{\tau + r} = \frac{prm'(I)}{\tau + r}\left((1 - q\beta)(1 - G(v^c|I)) + q\beta\right) \\ &= \frac{prm'(I)}{\tau + r}\left(1 - q\beta - (1 - q\beta)G(v^c|I) + q\beta\right) = \frac{prm'(I)}{\tau + r}\left(1 - (1 - q\beta)G(v^c|I)\right) \\ &= \frac{prm'(I)}{\tau + r}\left(1 - \alpha\Phi(x^c)\right) \end{aligned}$$

(this last equality follows from Theorem 2 when $I = \hat{I}$, i.e., when investors' conjecture about E's effor choice is correct), as claimed in the text.

Part i of the corollary is intuitive: whether E's optimal investment increases or decreases in f is determined by how large the sensitivity of the present value of the portion of the firm's cash flows E retains is to his investment choice, $\frac{\partial}{\partial I}(1-f)\delta wI = (1-f)\delta w$, relative to how large is the sensitivity of E's net proceeds from the fraction of the firm he sells is to his investment choice, $\frac{\partial \Psi(I,I)}{\partial I} = fwX$. If the former is larger (smaller) than the latter, E's investment decreases (increases) with f. The rest of the results in the corollary are equally intuitive: consider e.g., the result in (iia) that E's investment optimally increases in the precision r of the estimate $\tilde{v}$. The amount investors will pay for the fraction of the firm they buy will change more with the realized (and disclosed) value v of the estimate $\tilde{v}$ the greater is the precision r of this estimate. Realizing this, E naturally works harder as r increases. We do not take up the space to explain the other comparative statics, but they are all as intuitive as these two results.

7 E's equilibrium share retention decision

In this section, we study E's choice of what fraction f of the firm to sell in the IPO. The tension in selecting f is clear: since E discounts future cash flows more than outside investors do, other things equal E has an incentive to sell most or all of the firm to arbitrage this difference in discount rates, but the larger the fraction of the firm E sells, he has less of an incentive to invest in the firm "up front." Optimally choosing f just entails figuring out the optimal trade off between these two forces.

Taking investors' conjectures about his investment choice $\hat{I}(f)$ as given and then choosing $I^*(f)$ optimally, E's objective function can be written as:

$$OBJ = (1-f)\delta m(I^*(f)) - .5(I^*(f))^2 + f \times \Psi(I^*(f), \hat{I}(f)). \tag{35}$$

This is the expected discounted value to E of the fraction of cash flows E retains evaluated at E's optimal investment level choice, $(1-f)\delta m(I^*(f))$, net of the cost to E of that investment, $.5(I^*(f))^2$, plus the expected proceeds E receives for selling fraction f of the firm to outside investors net of the expected damage payments he may owe them, conditional on investors believing E has adopted investment level $\hat{I}(f)$, namely $f \times \Psi(I^*(f), \hat{I}(f))$. Invoking the envelope theorem,[8] and evaluating the derivative at the equilibrium investment level where $I^*(f) = \hat{I}(f)$, we observe that the derivative of OBJ in (35) with respect to f is given by:

$$\frac{\partial}{\partial f} OBJ = -\delta m(I^*(f)) + \Psi + f \times \frac{\partial \Psi}{\partial \hat{I}}|_{I=\hat{I}=I^*(f)} \times \frac{\partial \hat{I}(f)}{\partial f}. \tag{36}$$

For purposes of evaluating E's optimal choice of f, it is desirable to compute this last derivative, $\frac{\partial}{\partial f}OBJ$, evaluated at the equilibrium investment choice $I = I^*(f) = \hat{I}(f)$, explicitly. To that end we first notice that (see the accompanying footnote for details[9]) when investors are correct about E's investment choice:

$$f \times \Psi(I, I) = f \times m(I). \tag{37}$$

That is, the net proceeds E receives for selling fraction f of the firm to investors, $f \times \Psi(I, I)$, equals fraction f of the firm's total expected cash flows $m(I)$. The damage payments constitute, in expectation, a net wash: what E receives "up front" as payments from investors is what in expectation E, or rather, E's firm, expects to pay back to investors later. Notwithstanding this observation, it does not follow that the damage payments are irrelevant, as they affect what equilibrium investment choice E adopts, as we saw above in connection with the corollary to Theorem 4 (where it was shown that E's equilibrium investment choice $I^*(f)$ increases in the damages multiplier β).

[8] Which, when considering small changes in f, allows us to ignore the effect of changes in f on $I^*(f)$, but it does not allow us to disregard the effect of changes in f on $\hat{I}(f)$, since E controls (and can adjust) $I^*(f)$ as he f changes, but he does not control how investors' conjectures $\hat{I}(f)$ change with f.

[9] This follows from (21), the defining condition for the equilibrium value of v^c, as (21) implies, when $I = I^*(f) = \hat{I}(f)$ (and hence $\hat{m} = m = m(I^*(f))$), using the shorthand $G(v^c)$ for $G(v^c|I)$ and $g(v^c)$ for $g(v^c|I)$) :

$$\frac{\tau m + r v^c}{\tau + r} = \frac{(1 - p + pG(v^c))\frac{\tau m}{\tau + r} + \frac{r}{\tau + r}\left((1-p)m + p\int^{v^c} v g(v)dv + pq\beta \int^{v^c}(v^c - v)g(v)dv\right)}{1 - p + pG(v^c)}$$

and so, upon multiplying both sides of this last equation by the denominator of its RHS and

Next, recalling the definition of X in (33) above, we can also show that (see the accompanying footnote[10]) when $\hat{I} = I = I^*(f)$:

$$\frac{\partial}{\partial \hat{I}} \Psi(I, \hat{I})|_{I=\hat{I}=I^*(f)} = (1 - X)w. \tag{38}$$

Since from (34), we know when $m(I) = wI$, that $I^*(f)$ is linear in f and, related,

then adding $p \int_{v^c}^{\infty} \frac{\tau m + rv}{\tau + r} g(v)dv$ to both sides, we get:

$$\begin{aligned}
& \frac{\tau m + rv^c}{\tau + r} \times (1 - p + pG(v^c)) + p \int_{v^c}^{\infty} \frac{\tau m + rv}{\tau + r} g(v)dv \\
= \; & (1 - p + pG(v^c)) \frac{\tau m}{\tau + r} + \frac{r}{\tau + r} \left((1 - p)m + p \int^{v^c} vg(v)dv + pq\beta \int^{v^c} (v^c - v)g(v)dv \right) \\
& + p \int_{v^c}^{\infty} \frac{\tau m + rv}{\tau + r} g(v)dv \\
= \; & (1 - p)m + pG(v^c) \frac{\tau m}{\tau + r} + \frac{r}{\tau + r} \left(p \int^{v^c} vg(v)dv + pq\beta \int^{v^c} (v^c - v)g(v)dv \right) \\
& + p \int_{v^c}^{\infty} \frac{\tau m + rv}{\tau + r} g(v)dv \\
= \; & (1 - p)m + p \int^{v^c} \frac{\tau m + rv}{\tau + r} g(v)dv + \frac{r}{\tau + r} pq\beta \int^{v^c} (v^c - v)g(v)dv + p \int_{v^c}^{\infty} \frac{\tau m + rv}{\tau + r} g(v)dv \\
= \; & (1 - p)m + p \int_{-\infty}^{\infty} \frac{\tau m + rv}{\tau + r} g(v)dv + \frac{r}{\tau + r} pq\beta \int^{v^c} (v^c - v)g(v)dv \\
= \; & (1 - p)m + pm + \frac{r}{\tau + r} pq\beta \int^{v^c} (v^c - v)g(v)dv \\
= \; & m + pq\beta \int^{v^c} (\frac{\tau m + rv^c}{\tau + r} - \frac{\tau m + rv}{\tau + r} v)g(v)dv.
\end{aligned}$$

Hence, when $I = I^*(f) = \hat{I}(f)$, (25) can be expressed as:

$$\begin{aligned}
\Psi \; = \; & (1 - p + pG(v^c)) \frac{\tau m + rv^c(\hat{I})}{\tau + r} + p \int_{v^c(\hat{I})}^{\infty} \frac{\tau m + rv}{\tau + r} g(v|I)dv \\
& - pq\beta \int_{-\infty}^{v^c(\hat{I})} (\frac{\tau m + rv^c(\hat{I})}{\tau + r} - \frac{\tau m + rv}{\tau + r}) g(v|I)dv \\
= \; & m + pq\beta \int^{v^c} (\frac{\tau m + rv^c}{\tau + r} - \frac{\tau m + rv}{\tau + r} v)g(v)dv - pq\beta \int_{-\infty}^{v^c(\hat{I})} (\frac{\tau m + rv^c}{\tau + r} - \frac{\tau m + rv}{\tau + r}) g(v|I)dv \\
& m,
\end{aligned}$$

as claimed in the text.

[10] Rewrite $\Psi(I, \hat{I})$ as

$$\begin{aligned}
\Psi(I, \hat{I}) \; = \; & (1 - p) \frac{\tau \hat{m} + rv^c(\hat{I})}{\tau + r} + p(1 - q\beta) \left(\int_{-\infty}^{v^c(\hat{I})} \frac{\tau \hat{m} + rv^c(\hat{I})}{\tau + r} g(v|I)dv + \int_{v^c(\hat{I})}^{\infty} \frac{\tau \hat{m} + rv}{\tau + r} g(v|I)dv \right) \\
& + pq\beta \frac{\tau \hat{m} + rm}{\tau + r}.
\end{aligned}$$

that $I^*(f)$ can be written as:

$$I^*(f) = \delta w + f \times \frac{\partial I^*(f)}{\partial f} \tag{39}$$

where

$$\frac{\partial I^*(f)}{\partial f} = w(X - \delta). \tag{40}$$

But, since an equilbrium requirement is that $I^*(f) \equiv \hat{I}(f)$ is an identity in f, it follows that:

$$\frac{\partial I^*(f)}{\partial f} = \frac{\partial \hat{I}(f)}{\partial f}. \tag{41}$$

Combining this expression for $\Psi(I, \hat{I})$ with $m(\hat{I}) = w\hat{a}$, we conclude that

$$\begin{aligned}
\frac{\partial}{\partial \hat{I}}\Psi(I,\hat{I}) &= \frac{\partial}{\partial v^c}\Psi(I,\hat{I})\frac{\partial v^c}{\partial \hat{I}} + \frac{\partial}{\partial \hat{m}}\Psi(I,\hat{I})\frac{\partial \hat{m}}{\partial \hat{I}} \\
&= \{(1-p)\frac{r}{\tau+r} + p(1-q\beta)G(v^c(\hat{I})|I)\frac{r}{\tau+r}\}\frac{\partial v^c}{\partial \hat{I}} \\
&\quad +\{(1-p)\frac{\tau}{\tau+r} + p(1-q\beta)\frac{\tau}{\tau+r} + pq\beta\frac{\tau}{\tau+r}\}\frac{\partial \hat{m}}{\partial \hat{I}} \\
&= \{1-p+p(1-q\beta)G(v^c(\hat{I})|I)\}\frac{r}{\tau+r}\frac{\partial v^c}{\partial \hat{I}} + \{1-p+p(1-q\beta)+pq\beta\}\frac{\tau}{\tau+r}\frac{\partial \hat{m}}{\partial \hat{I}} \\
&= \{1-p+p(1-q\beta)G(v^c(\hat{I})|I)\}\frac{r}{\tau+r}\frac{\partial v^c}{\partial \hat{I}} + \frac{\tau}{\tau+r}\frac{\partial \hat{m}}{\partial \hat{I}} \\
&= \{1-p+p(1-q\beta)G(v^c(\hat{I})|I)\}\frac{rw}{\tau+r} + \frac{\tau w}{\tau+r}.
\end{aligned}$$

In this last line, we utilized Theorem 2 to conclude $\frac{\partial v^c}{\partial \hat{I}} = m'(\hat{I}) = w$. Hence, when $\hat{I} = I$:

$$\begin{aligned}
\frac{\partial}{\partial \hat{I}}\Psi(I,I) &= (\{1-p+p(1-q\beta)G(v^c(I)|I)\}\frac{r}{\tau+r} + \frac{\tau}{\tau+r})w \\
&= (\frac{r}{\tau+r} - \frac{pr}{\tau+r}(1-(1-q\beta)G(v^c(I)|I)) + \frac{\tau}{\tau+r})w \\
&= (1 - \frac{pr}{\tau+r}(1-\alpha G(v^c(I)|I)))w \\
&= (1 - \frac{pr}{\tau+r}(1-\alpha\Phi(x^c)))w \\
&= (1-X)w.
\end{aligned}$$

In the second to last line, we employed Theorem 2 to conclude that the equilibrium probability of no disclosure is given by $\Phi(x^c)$.

Putting (38) to (41) together, we can express the derivative (36) as:

$$\begin{aligned}\frac{\partial}{\partial f}OBJ &= -\delta m(I^*(f)) + \Psi + f \times \frac{\partial \Psi}{\partial \hat{I}}|_{I=\hat{I}=I^*(f)} \times \frac{\partial \hat{I}(f)}{\partial f} && (42)\\ &= (1-\delta)wI^*(f) + f \times \frac{\partial \Psi}{\partial \hat{I}}|_{I=\hat{I}=I^*(f)} \times \frac{\partial \hat{I}(f)}{\partial f}\\ &= (1-\delta)w(\delta w + f \times \frac{\partial I^*(f)}{\partial f}) + f \times (1-X)w \times \frac{\partial I^*(f)}{\partial f}\\ &= (1-\delta)\delta w^2 + (1-\delta+1-X)\, wf \frac{\partial I^*(f)}{\partial f}\\ &= (1-\delta)\delta w^2 + (2-\delta-X)\, w^2 f(X-\delta). && (43)\end{aligned}$$

Working with (43), we obtain the following theorem. In the statement of the theorem, we make reference to the function

$$\psi(r) \equiv (1 - X(r))^2 \text{ for all } r > 0,$$

where $X = X(r)$, which we presently view as a function of r, is as defined in (33) above.

Theorem 6 *(A) If* $(1 - p(1 - \alpha\Phi(x^c)))^2 < 1 - \delta$*, then there exists a unique* $r > 0$*, call it* r^δ*, such that* $\psi(r^\delta) = 1 - \delta$*.*

(Ai) For all $r < r^\delta$*, the optimal* f *is given by*

$$f^* = \frac{(1-\delta)\delta}{(1-X)^2 - (1-\delta)^2}; \tag{44}$$

(Aii) for all $r > r^\delta$*, the optimal* f *is* $f = 1$*.*

(B) If $(1-p(1-\alpha\Phi(x^c)))^2 \geq 1-\delta$*, then for all* $r > 0$*, the optimal* f *is given by (44).*

Using the theorem, we can make specific predictions about how E's optimal share retention $1 - f^*$ varies with various parameters of the model, as summarized in the following corollary.

Corollary 7 *E's equilibrium ownership stake* $1 - f^*$ *always (at least weakly):*

(i) declines as the precision r *of the estimate* $\tilde{v}$ *increases;*

(ii) declines as the probability p E receives information increases;

(iii) declines as the damages multiplier β increases.

The first two comparative statics in the corollary are intuitive. Holding his investment choice fixed, E always has an incentive to sell 100% of his firm to outside investors, since they value the firm more than he does (because of their lower discount rates). But, since E's investment choice is endogenous, and depends in part on what fraction of the firm he retains, his ex ante investment choice will be inefficiently low unless he retains a substantial ownership stake in the firm. However, as the quality of the estimate $\tilde{v}$ E receives increases, as measured by the precision of the estimate, or as the probability E receives the estimate increases, the accounting estimate $\tilde{v}$ can be relied on more to ensure that E has good incentives to work hard to invest in the firm even if he sells some fraction of his original ownership stake. Hence, E can profitably sell a larger stake in the firm as either r or p increase. Perhaps somewhat surprisingly, according to part (iii) of the corollary E sells a larger stake in the firm to investors as the damages multiplier increases. This is somewhat surprising, because E incurs more costs in selling a larger fraction of the firm to investors as the damages multiplier increases.

8 Extensions

8.1 E's continued involvement in the firm after the IPO

In the base model, there were no actions that E could take after the IPO was completed that influenced the cash flows E's firm eventually generates. In practice, this may not be true, because E may make key operational and/or investment decisions after the IPO that affect the firm's cash flows. But, the base model can be easily extended to incorporate such effects. For example, suppose that after the IPO is completed, E makes a second personally costly investment decision I_2 that augments the cash flows the firm ultimately produces. Specifically, suppose that the firm's total cash flows z are increased in expectation by $w_2 I_2$ when E selects investment I_2 after the IPO at cost to E of $.5I_2^2$. If at the IPO, E sold fraction f of the firm to outside investors, then E's choice of I_2 will be determined by maximizing $(1-f)w_2 I_2 - .5I_2^2$, leading E to choose optimal second investment $I_2^* = (1-f)w_2$, and thereby augmenting the firm's expected cash flows by $(1-f)w_2 I_2^* = (1-f)^2 w_2^2$.

Expanding the base model to include this second, post IPO, investment by E, it is clear that: (a) E's optimal "long run" disclosure choice, as in (7), will continue to be the same as the optimal "short run" disclosure choice, as in (8), provided the bound (2) on β continues to hold; (b) Theorem 2 will continue to characterize the equilibrium cutoff for disclosure; (c) the equilibrium "no disclosure" price of the firm will now be the sum of $\frac{\tau \hat{m} + r v^c}{\tau + r}$, where the cutoff is as just described in (b) augmented by the extra cash flows $(1-f)^2 w_2^2$ generated by the second investment (as calculated in the preceding paragraph); (d) the comparative statics summarized in Corollary 3 will hold without change; and (e) the optimal initial investment for a fixed share retention decision will continue to be as described in Theorem 4, and the comparative statics concerning that investment will be as described in Corollary 5. What will change is E's preferred choice regarding the fraction of the firm E sells in the IPO. Naturally, this fraction optimally will be adjusted downwards to reflect the favorable impact of increasing E's share retention on the size of his optimal second investment I_2^*. This change will require adjustments to the statements of Theorem 6 and Corollary 7 concerning E's optimal share retention. As these adjustments are straightforward, we do not report them here.

8.2 E's equilibrium share retention with a delayed choice of share retention

Another possible extension of some interest is a variant of the base model studied above in which E is assumed to delay making a choice regarding what fraction

f of the firm to sell to outside investors until the time of the IPO, i.e., after he either learns the estimate v about the cash flows of the firm or learns that he will not obtain such an estimate. We offer a brief discussion of this extension in the paragraphs that follow.

The first observation we make regarding this "delayed f" extension is that if E receives, and decides to disclose, information v, then E's subsequent equilibrium choice of the fraction f of the firm to sell to outside investors is uniquely given by $f = 1$, that is, E will sell his entire firm to investors. The technical details of the argument are described in the accompanying footnote,[11] but the economics are simple: at the time of the IPO in the base model, the distribution of the firm's cash flows is fixed, so it is sensible at that point when v is public to allocate the firm's cash flows to the parties that value those cash flows the most. Since E has a smaller discount factor than the outsider investors do, it is efficient to transfer all of the firm's cash flows to the outside investors, by having them pay a lump sum (their present value of these cash flows) to E. Hence, if I denotes his initial investment level, E's expected utility conditional

[11] To see this, note first that if E discloses v, then investors' perceptions of the expected value of the firm, $P(v|\hat{I}) = \frac{\tau m(\hat{I})+rv}{\tau+r}$, cannot, in any equilibrium, vary with the fraction f of the firm E proposes to sell them. Equivalently, if E discloses v, investors' conjecture of E's investment choice $\hat{I}$ cannot vary with f in any equilibrium. This follows directly from an argument by contradiction: if investors' equilibrium conjecture $\hat{I} = \hat{I}(f)$ did vary with f, then when E actually makes his investment choice at the start of the model, E should anticipate, and take account of, that variation. Since E makes his investment choice "up front," then E ultimately must decide on a *single* investment choice. Hence, by the time E selects f at the time of the IPO, his actual investment choice will be sunk, and hence will not vary with f. So, his actual investment choice will not be consistent with investors' conjecture that his investment choice varies with f. Hence, a conjecture that I varies with f is not consistent with any equilibrium.

Hence, in equilibrium investors' conjecture $\hat{I}$ must be a constant independent of f. The same must also be true of $m(\hat{I})$: investors' conjecture about the firm's expected cash flows must also be independent of f.

Furthermore, in any equilibrium, investors' conjecture about $\hat{I}$ will coincide with E's actual choice, and hence $m(\hat{I})$ will coincide with $m(I)$. Hence, if E sells fraction f of the firm to investors when he received, and disclosed v, he will get expected utility at that point of $(1-f)\delta\frac{\tau m+rv}{\tau+r} + f\frac{\tau m+rv}{\tau+r} - .5I^2$. Since $\delta < 1$, and - by the reasoning just given, since m does not vary with f - E's expected utility is clearly strictly increasing in f for all $f \leq 1$.

This proves the optimality for E of selling all of his firm to investors whenever he discloses v.

on learning $\tilde{v} = v$ will be

$$P(v|I) - .5I^2 = \frac{\tau m(I) + rv}{\tau + r} - .5I^2. \tag{45}$$

Next, we discuss whether, when E receives information v that he decides not to disclose, E in equilibrium will choose to sell outside investors a fraction f of the firm that differs from the fraction, call it $f^{no\inf o}$, E would have sold investors in the event he received no information. Were E to make such a choice, then both outside investors and the fact finder would know that E received information that he did not disclose. While, theoretically, one could entertain the possibility that E signals, through his choice of f, that he obtained information he did not disclose at the same time the fact finder might not be able to detect E's withholding, it seems implausble to us that a fact finder in this situation would not continue to search for evidence of E's withholding until he found it.[12] If E anticipated such dilligence on the part of the fact finder, E would not choose to signal that he received information that he did not subsequently disclose. Hence, it is reasonable to conclude that for any v E received but did not disclose, E chooses the same fraction $f^{no\inf o}$ of the firm to sell to investors that he would have sold to them had he not received any information. Relying on this intuition, we write f^{nd} to refer to *the* equilibrium fraction of the firm E sells to investors when he makes no disclosure, independent of whether or not E withholds information.

The determination of this f^{nd} is not unique when E delays his choice of f to the IPO. While we do not analyze here the full spectrum of equilibria that can emerge in the "delayed" case, we can use the analysis of the preceding sections to characterize one of these "delayed f" equilibria. We assert that the following is an equilibrium: when E makes no disclosure, $f^{nd} = 1$ accompanied by off equilibrium beliefs that stipulate that in the event E makes no disclosure

[12] This is unlike the situation the fact finder faced in the previous sections of the paper, where the fact finder did not know, at the time E made no disclosure, whether E withheld information he received or whether E did not receive any information.

and chooses an f other than f^{nd}, then investors conclude that E observed and withheld a "very low" value of v, along with: $P^{nd}(f^{nd} = 1) = v^c(m(I^*), \sigma)$ where $v^c(m(I^*), \sigma)$ is defined by (22) above and I^* as defined in (34) above with $f = 1$, i.e., with $I^*(f = 1) = wX$.

We sketch here why, with these specifications, the preceding is an equilibrium in the "delayed f" case. If E receives information v, then given the above specification of off equilibrium beliefs, E will choose $f = 1$ regardless of whether he discloses or withholds v, and E will make a decision to disclose or withhold v based on a comparision of his expected utility conditional on receiving v from disclosing v, as described in (45) above (with I^* replacing I), to his expected utility (also conditional on receiving v) from withholding v, given by:

$$\begin{aligned} & (1-q)f^{nd}P^{nd} + q(f^{nd}P(v|I^*) - \beta(P^{nd} - P(v|I^*)) + (1-f^{nd})\delta E[\tilde{z}|I^*] - .5I^{*2} \\ = \quad & (1-q)P^{nd} + q(P(v|I^*) - \beta(P^{nd} - P(v|I^*)) - .5I^{*2}. \end{aligned}$$

This comparsion is easily shown to amount to comparing P^{nd} to $P(v|I^*)$, just as the desirability of disclosing or withholding v in the case studied in the preceding sections started by comparing LHS(7) to RHS(7) and finished by comparing LHS(8) to RHS(8). Because E's share retention is the same with and without disclosure, and because E's disclosure decision is the same in this case as that previously studied, it follows that E's no "disclosure price" and hence E's equilibrium "no disclosure" set is the same as that previously studied as well.

What is not the same in this case as that studied in previous sections is that no extra conditions are needed now (in the "delayed f" case) to ensure the optimality of $f^{nd} = 1$. We showed in Theorem 6 above that when E chooses f "up front," it is an equilbrium for E to set $f = 1$ only when the conditions of Theorem 6 A(ii) held. In contrast, note that the preceding argument for $f^{nd} = 1$ in the "delayed f" case holds without imposing such extra conditions as those described in Theorem 6 A(ii). This is to be expected. If E selects his share retention decision "delayed," then E perforce disregards the effects of his choice of f on his initial investment choice, as that choice is now sunk. Hence, the set of circumstances under which E is inclined to sell his entire firm at the time of the IPO will be considerably broader than when he chose f earlier and took into account the incentive effects of his share retention decision on his initial investment choice.

Finally we note that were we to combine the present "delayed f" extension with the previous extension considered above where E selects a second post-IPO investment level after the IPO is completed, then the optimal "delayed f" share retention typically will not shrink to $1 - f = 0$, in view of how damaging such a choice would be for E's optimal second investment decision. Thus, while E's optimal share retention level in the "delayed f" situation is still likely to be lower than it would be were E to choose the size of the IPO f "up front" in the presence of a second investment choice by E, it would typically not degenerate to E eliminating entirely his ownership stake in the firm.

9 Summary

We have studied a moral hazard-based model of an entrepreneur who develops a firm which engages in an IPO. Increases in the entrepreneur's equilibrium investment ex ante, before the IPO occurs, increase the post IPO firm's expected cash flows. The entrepreneur's partial, and sometimes complete, unwinding of his ownership stake in the firm through the IPO is efficient because the entrepreneur is posited to discount the firm's future cash flows more than do outside investors. We show how the entrepreneur's equilibrium investment decision, share retention decision, and voluntary disclosure decision are affected by each of: the entrepreneur's discount rate, the entrepreneur's cost of making the initial investment, the precision of the entrepreneur's estimate about the firm's future cash flows, the probability the entrepreneur receives such an estimate, and the entrepreneur's legal liability for withholding value relevant information from investors. Among our results, we show: 1. the robustness of the equilibrium probability the entreprenuer discloses the private information he receives to a variety of parameters affecting his information environment; 2. that the entrepreneur's disclosure decision is often the same whether he emphasizes "short run" or "long run" considerations in the presence of 10b-5 like damage payments; and 3. increases in the penalty imposed on the entrepreneur for withholding value-relevant information from investors often leads him to disclose less, not more, information to investors, and leads him to sell more, not less, of his firm to outside investors in the IPO.

When the entrepreneur chooses what fraction of the firm to retain just before the IPO, we show that he tends to choose to retain a smaller percentage of the firm (as compared to when he selects what fraction of the firm to retain "up front"). This outcome occurs because while retaining a higher fraction of the firm often has favorable effects on the entrepreneur's incentives to invest in his firm initially, such investments at the time the IPO takes place are sunk, and so are irrelevant to his optimal share retention decisions at that time. Maintaining significant post IPO share retention levels for the entrepreneur can be shown to be optimal even in those instances where the entrepreneur's initial investment become sunk, provided the entrepreneur's continued involvement in the firm sufficiently affects the realized values of the firm's post-IPO cash flows.

10 Appendix: Proofs (not proven in the text or footnotes)

Lemma 8 *If $\tilde{u}$ is normally distributed with mean $m(I)$ and variance σ^2, density $h(u|I)$ and cdf $H(u|I)$, then:*

(i) $\int^{u^c}(u^c - u)h(u|I)du = \int_{-\infty}^{\infty} \max\{u^c, u\}h(u|I)du - m(I) = H(u^c|I)(u^c - m(I)) + \sigma^2 h(u^c|I);$

(ii) $\int_{-\infty}^{\infty} \max\{u^c, u\} h(u|I) du = H(u^c|I)(u^c - m(I)) + \sigma^2 h(u^c) + m(I)$;

(iii) $\int^{u^c} u h(u|I) du = -\sigma^2 h(u^c|I) + m(I) H(u^c|I)$;

(iv) $\int_{u^c} u h(u|I) du = m(I)(1 - H(u^c|I)) + \sigma^2 \times h(u^c)$;

(v) $H_I(u^c|I) = -m'(I) h(u^c|I)$;

(vi) $\int_{-\infty}^{\infty} \max\{u^c, u\} h_I(u|I) du = m'(I)(1 - H(u^c|I))$.

Proof of Lemma 8 By definition: $h(u|I) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(u-m(I))^2}{2\sigma^2}}$ and $H(u^c|I) = \int_{-\infty}^{u^c} h(u|I) du$. Observe

$$\frac{dh(u|I)}{du} = \frac{d}{du} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(u-m(I))^2}{2\sigma^2}} = -\frac{u - m(I)}{\sigma^2} \times h(u|I).$$

Thus, for any u^c :

$$\int^{u^c} (u - m(I)) h(u|I) du = -\sigma^2 \times \int^{u^c} (-\frac{u - m(I)}{\sigma^2}) h(u|I) du = -\sigma^2 h(u^c|I),$$

so

$$\int^{u^c} u h(u|I) du = \int^{u^c} (u - m(I)) h(u|I) du + m(I) H(u^c|I) = -\sigma^2 h(u^c|I) + m(I) H(u^c|I).$$

This proves (iii).

Related:

$$\frac{dh(u|I)}{da} = \frac{u - m(I)}{\sigma^2} \times m'(I) \times h(u|I).$$

it follows that,

$$\begin{aligned} H_I(u^c|I) &= \int_{-\infty}^{u^c} h_I(u|I) du = \int_{-\infty}^{u^c} \frac{u - m(I)}{\sigma^2} \times m'(I) \times h(u|I) du \\ &= -m'(I) \int_{-\infty}^{u^c} (-\frac{u - m(I)}{\sigma^2}) h(u|I) du = -m'(I) h(u^c|I). \end{aligned}$$

This proves (v).

It is a standard observation concerning truncated normal random variables that:

$$E[\tilde{u}|\tilde{u} > u^c] = m(I) + \sigma \times \frac{\phi(\frac{u^c - m(I)}{\sigma})}{1 - \Phi(\frac{u^c - m(I)}{\sigma})}.$$

Hence, since $\phi(\frac{u^c - m(I)}{\sigma}) = \sigma h(u^c|I)$:

$$\begin{aligned}\int_{u^c} uh(u|I)du &= \Pr(\tilde{u} > u^c) \times E[\tilde{u}|\tilde{u} > u^c] = m(I)(1 - H(u^c|I)) + \sigma \times \phi(\frac{u^c - m(I)}{\sigma}) \\ &= m(I)(1 - H(u^c|I)) + \sigma^2 \times h(u^c|I).\end{aligned}$$

This proves (iv). This also shows that

$$\begin{aligned}\int_{-\infty}^{\infty} \max\{u^c, u\}h(u|I)du &= \Pr(\tilde{u} \leq u^c)u^c + \Pr(\tilde{u} > u^c) \times E[\tilde{u}|\tilde{u} > u^c] \\ &= H(u^c|I)u^c + m(I)(1 - H(u^c|I)) + \sigma^2 h(u^c) \\ &= H(u^c|I)(u^c - m(I)) + m(I) + \sigma^2 h(u^c).\end{aligned}$$

This proves (ii). This further implies:

$$\begin{aligned}\int^{u^c} (u^c - u)h(u|I)du &= \int_{-\infty}^{\infty} \max\{u^c - u, u - u\}h(u|I)du \\ &= \int_{-\infty}^{\infty} (\max\{u^c, u\} - u)h(u|I)du \\ &= \int_{-\infty}^{\infty} \max\{u^c, u\}h(u|I)du - \int_{-\infty}^{\infty} uh(u|I)du \\ &= \int_{-\infty}^{\infty} \max\{u^c, u\}h(u|I)du - m(I) \\ &= H(u^c|I)(u^c - m(I)) + \sigma^2 h(u^c|I).\end{aligned}$$

This proves (i). From this, it also follows that:

$$\begin{aligned}
\int_{-\infty}^{\infty} \max\{u^c, u\} h_I(u|I)du &= \frac{\partial}{\partial I}\int_{-\infty}^{\infty} \max\{u^c, u\} h(u|I)du \\
&= \frac{\partial}{\partial I}\left(H(u^c|I)(u^c - m(I)) + m(I) + \sigma\phi(\frac{u^c - m(I)}{\sigma})\right) \\
&= H_I(u^c|I)(u^c - m(I)) - m'(I)H(u^c|I) + m'(I) + \sigma\frac{\partial}{\partial I}\phi(\frac{u^c - m(I)}{\sigma}) \\
&= H_I(u^c|I)(u^c - m(I)) - m'(I)H(u^c|I) \\
&\quad + m'(I) + m'(I)\frac{u^c - m(I)}{\sigma^2}\sigma\phi(\frac{u^c - m(I)}{\sigma}) \\
&= H_I(u^c|I)(u^c - m(I)) - m'(I)H(u^c|I) + m'(I) \\
&\quad + m'(I)(u^c - m(I))h(u^c|I) \\
&= -m'(I)h(u^c|I)(u^c - m(I)) - m'(I)H(u^c|I) + m'(I) \\
&\quad + m'(I)(u^c - m(I))h(u^c|I) \\
&= m'(I)(1 - H(u^c|I)).
\end{aligned}$$

(The second to last line follows from part (v).) This proves (vi).■

Proof of Theorem 2 Multiply both sides of equation (21) by the denominator of RHS(21) to get:

$$\begin{aligned}
&\frac{\tau\hat{m} + rv^c}{\tau + r} \times (1 - p + pG(v^c)) \\
= &\quad (1 - p + pG(v^c))\frac{\tau\hat{m}}{\tau + r} + \frac{r}{\tau + r}\left((1 - p)\hat{m} + p\int^{v^c} vg(v)dv + pq\beta\int^{v^c}(v^c - v)g(v)dv\right),
\end{aligned}$$

or equivalently,

$$v^c \times (1 - p + pG(v^c)) = (1 - p)\hat{m} + p\int^{v^c} vg(v)dv + pq\beta\int^{v^c}(v^c - v)g(v)dv,$$

or equivalently

$$(v^c - \hat{m})(1 - p) + pv^cG(v^c) = p\int^{v^c} vg(v)dv + pq\beta\int^{v^c}(v^c - v)g(v)dv. \quad (46)$$

From Lemma 8 parts (i) and (iii) we know:

$$\int^{v^c} vg(v)dv = -\sigma^2 g(v^c) + \hat{m}G(v^c)$$

and

$$\int^{v^c} (v^c - v)g(v)dv = G(v^c)(v^c - \hat{m}) + \sigma^2 g(v^c),$$

so (46) can be written as:

$$(v^c - \hat{m})(1-p) + pv^c G(v^c) = -p\sigma^2 g(v^c) + p\hat{m}G(v^c) + pq\beta G(v^c)(v^c - \hat{m}) + pq\beta\sigma^2 g(v^c),$$

i.e., as:

$$(v^c - \hat{m})(1 - p + p(1 - q\beta)G(v^c)) + p(1 - q\beta)\sigma^2 g(v^c) = 0.$$

Dividing this last equation by σ and recalling the definition of α, we note that this last equation can be rewritten:

$$\left(\frac{v^c - \hat{m}}{\sigma}\right)(1 - p + \alpha p G(v^c)) + \alpha p \sigma g(v^c) = 0. \tag{47}$$

Now define x^c by

$$x^c \equiv \frac{v^c - \hat{m}}{\sigma}, \tag{48}$$

and substitute this x^c into (47), after observing that $G(v^c) = \Phi(x^c)$ and $\sigma g(v^c) = \phi(x^c)$, to conclude that (47) can be rewritten as:

$$x^c(1 - p + \alpha p \Phi(x^c)) + \alpha p \phi(x^c) = 0.$$

■

Proof of Corollary 3 part (i) First note that LHS(23) is positive for all $x \geq 0$, so if (23) has a solution, that solution must be negative. Also notice that LHS(23) is strictly increasing in x fo all x, since $\phi'(x) = -x\phi(x)$ and

$$\frac{\partial LHS(23)}{\partial x} = 1 - p + \alpha p \Phi(x) + x\alpha p \phi(x) - \alpha p \phi(x) x = 1 - p + \alpha p \Phi(x) > 0. \tag{49}$$

Also notice that LHS(23) goes to $-\infty$ (and hence, in particular, turns negative) as $x \to -\infty$, since $x(1-p) \to -\infty$ as $x \to -\infty$. Obviously, LHS(23) is continuous in x. Thus, (23) has a unique, negative solution for all $\alpha, p \in [0, 1]$. This proves (i). It also proves (vi), since $x^c < 0$ implies $\Phi(x^c) < .5$.

part (ii) Differentiate (23) totally with respect to p, using $\phi'(x) = -x\phi(x)$ to get:

$$\frac{\partial LHS(23)}{\partial x}\frac{\partial x^c}{\partial p} + \frac{\partial LHS(23)}{\partial p} = 0, \tag{50}$$

or equivalently, using (49):

$$[1 - p + \alpha p\Phi(x^c)]\frac{\partial x^c}{\partial p} = x^c(1 - \alpha\Phi(x^c)) - \alpha\phi(x^c).$$

Since $1 > \alpha$, we have $1 > \Phi(x^c) \geq \alpha\Phi(x^c)$. That, combined with $x^c < 0$, yields: $x^c(1 - \alpha\Phi(x^c)) - \alpha\phi(x^c) < 0$. From this, $\frac{\partial x^c}{\partial p} < 0$ follows.

part (v) Differentiate (23) totally with respect to α, in a fashion analogous to (50) above, to get:

$$[1 - p + \alpha p\Phi(x^c)]\frac{\partial x^c}{\partial \alpha} + x^c p\Phi(x^c) + p\phi(x^c) = 0,$$

or equivalently:

$$\frac{\partial x^c}{\partial \alpha} = -p\frac{x^c\Phi(x^c) + \phi(x^c)}{1 - p + \alpha p\Phi(x^c)}. \tag{51}$$

Now, we claim that $f(x)$ defined by:

$$f(x) \equiv x\Phi(x) + \phi(x) \tag{52}$$

is positive for all $x \in \mathbb{R}$. To see this, first note that notice that $\phi' = -x\phi$, so that $f'(x) = x\phi + \Phi - x\phi = \Phi(x) > 0$, so $f(\cdot)$ is strictly increasing in x for all x. Thus, if $\lim_{x\to-\infty} f(x) = 0$, we will be done, as this will show that $f(x) > 0$ for all x. But notice that $x\Phi(x)$ can be written as $x\Phi(x) = \frac{\Phi(x)}{\frac{1}{x}}$, and $\lim_{x\to-\infty}\Phi(x) = 0$, and $\lim_{x\to-\infty}\frac{1}{x} = 0$, so L'Hospitals rule applies to establish that

$$\lim_{x\to-\infty} x\Phi(x) = \lim_{x\to-\infty}\frac{\Phi(x)}{\frac{1}{x}} = \lim_{x\to-\infty}\frac{\phi(x)}{-\frac{1}{x^2}} = -\frac{1}{\sqrt{2\pi}}\lim_{x\to-\infty} x^2 e^{-x^2/2} = 0. \tag{53}$$

Since $\lim_{x\to-\infty}\phi(x) = 0$ too, it follows from (53) that $\lim_{x\to-\infty} x\Phi + \phi(x) = 0$, too. That is, $\lim_{x\to-\infty} f(x) = 0$. This completes the proof of the claim that $f(x) > 0$ for all finite x. Notice using the notation (52) that we can write $\frac{\partial x^c}{\partial \alpha}$ in

(51) as $\frac{\partial x^c}{\partial \alpha} = -p\frac{f(x^c)}{1-p+\alpha p\Phi(x^c)}$. Since $f(x)$ is now known to be positive for all x, it follows that $\frac{\partial x^c}{\partial \alpha} < 0$, as claimed in part (v). Parts (iii) and (iv) now follow immediately, since $\alpha = 1 - q\beta$, and hence $sgn\frac{\partial x^c}{\partial \beta} = sgn\frac{\partial x^c}{\partial \alpha}\frac{\partial \alpha}{\partial \beta} = -sgn\frac{\partial x^c}{\partial \alpha}$, and similarly $sgn\frac{\partial x^c}{\partial q} = sgn\frac{\partial x^c}{\partial \alpha}\frac{\partial \alpha}{\partial q} = -sgn\frac{\partial x^c}{\partial \alpha}$.■

Proof of Theorem 6

We begin by proving the following lemma.

Lemma 9 *Case 1 If* $\delta \leq X \equiv \frac{pr}{\tau+r}(1-\alpha\Phi(x^c))$*, then the optimal* f *is* $f = 1$

Case 2a If $\delta > X$ *and* $1-\delta < (1-X)^2$*, then the optimal* f *is*

$$f^* = \frac{(1-\delta)\delta}{(1-X)^2 - (1-\delta)^2}. \tag{54}$$

Case 2b $\delta > X$ *and* $1-\delta \geq (1-X)^2$*, then the optimal* f *is* $f = 1$.

Proof of Lemma 9

First, we note that since

$$\begin{aligned}(2-\delta-X)(-\delta+X) &= -\delta(2-\delta) + X(2-\delta) + X\delta - X^2 \\ &= -\delta(2-\delta) + 2X - X^2 \\ &= 1-\delta)^2 - (1-X)^2,\end{aligned}$$

(43) can be rewritten as:

$$\frac{\partial}{\partial f}OBJ = (1-\delta)\delta w^2 + w^2 f\{(1-\delta)^2 - (1-X)^2\}.$$

Next, notice that since both δ and X are positive and less than 1, Case 1 of the lemma occurs iff $(1-\delta)^2 \geq (1-X)^2$. Thus, in view of (**??**), $\frac{\partial}{\partial f}OBJ$ is positive for all $f \geq 0$. Hence, the optimal f is $f^* = 1$.

Notice in Case 2 of the lemma, $(1-\delta)^2 < (1-X)^2$. Thus, in Case 2 the ratio in (54) is positive. In this case, this ratio is less than 1 iff

$$\begin{aligned}(1-\delta)\delta &< (1-X)^2 - (1-\delta)^2 \text{ iff} \\ (1-\delta)^2 + \delta - \delta^2 &< (1-X)^2 \\ 1-\delta &< (1-X)^2. \end{aligned} \tag{55}$$

When inequality (55) holds - Case 2(a) - $\frac{\partial}{\partial f}OBJ$ is positive for $f < f^*$, where f^* is as given in (54), and $\frac{\partial}{\partial f}OBJ$ is negative for $f > f^*$. Hence, in Case 2(a), the f^* as defined in (54) is the global maximum of OBJ defined in (35). When inequality (55) is reversed - Case 2(b) - $\frac{\partial}{\partial f}OBJ$ is positive for all $f < f^*$ where f^* is as defined in (54). In particular, since in Case 2(b) the f^* defined in (54) exceeds 1, it follows that for all $f < 1$ in Case 2b, $\frac{\partial}{\partial f}OBJ$ is positive. Hence, in Case 2(b), the optimal f is $f = 1$. This proves the lemma.$\blacksquare$

Now, to prove Theorem 6, recall the function $\psi(r)$ for $r > 0$ defined by $\psi(r) \equiv (1 - \frac{pr}{\tau+r}(1 - \alpha\Phi(x)))^2$.

We prove Case B (in the statement of Theorem 6) first. We begin by observing that the function $\psi(r)$ is strictly continuously decreasing in r for all $r > 0$, $\psi(0) = 1$, and $\lim_{r\to\infty}\psi(r) = (1 - p(1 - \alpha\Phi(x))).^2$

In Case B, $\psi(r) > 1 - \delta$ for all $r > 0$. In this case, $p(1 - \alpha\Phi(x)) \leq \delta$ must hold, since if $p(1 - \alpha\Phi(x)) > \delta$, then $1 - p(1 - \alpha\Phi(x)) < 1 - \delta$, and so $(1 - p(1 - \alpha\Phi(x)))^2 < 1 - \delta$, contrary to this case. Since $p(1 - \alpha\Phi(x)) \leq \delta$ obviously implies $X = \frac{pr}{\tau+r}(1 - \alpha\Phi(x)) < \delta$ holds for any $r > 0$, we have both $X < \delta$ and $(1 - X)^2 = \phi(r) > 1 - \delta$ for all $r > 0$ in this case. Thus, in Case A, we conclude that the conditions of Case 2(b) of the previous lemma are satisfied for all $r > 0$. By that lemma, we conclude that the optimal f is given by f^* as defined in (54) for all $r > 0$ in Case B.

In Case A (in the statement of Theorem 6), first note that by the properties of $\psi(\cdot)$ identified above, it is clear that there is a unique $r^\delta > 0$ as identified in the statement of the theorem. notice that since $1 - \delta = \psi(r^\delta) = (1 - \frac{pr^\delta}{\tau+r^\delta}(1-\alpha\Phi(x^c)))^2 < 1 - \frac{pr^\delta}{\tau+r^\delta}(1-\alpha\Phi(x^c))$, it follows that $\frac{pr^\delta}{\tau+r^\delta}(1-\alpha\Phi(x^c)) < \delta$. Hence, if $r < r^\delta$, then both $X = \frac{pr}{\tau+r}(1 - \alpha\Phi(x^c)) < \delta$ and $1 - \delta < \psi(r) = (1 - \frac{pr}{\tau+r}(1 - \alpha\Phi(x^c)))^2 = (1 - X)^2$. That is, if $r < r^\delta$, the conditions of Case 2A of the previous lemma are satisfied. Thus, f^* as defined in (55) is optimal. Finally, consider Case A with $r > r^\delta$. For all such r, we have $(1 - X)^2 = \psi(r) < 1 - \delta$. Now, in this case, either $\delta \leq X$ or $\delta > X$. If $\delta \leq X$, then the conditions of Case 1 are satisfied, so we conclude that the optimal f is $f = 1$. Alternatively, if $\delta > X$ then the conditions of Case 2a of the previous lemma are satisfied, and once again we conclude that the optimal f is $f = 1$. This completes the proof of Theorem 6.■

Proof of Corollary 7 The essential step is to show that as β increases, f^* as defined in (44) strictly increases. Note that f^* can be written as $f^* = -\frac{(1-\delta)\delta w^2}{w^2((1-\delta)^2-\psi)}$, where ψ was defined just before the statement of the theorem. We intend to prove that $\alpha\Phi(x(\alpha))$ always increases in α for the equilibrium cutoff $x(\alpha)$. This will show that $\alpha\Phi(x^c(\alpha))$ is decreasing in β (because $\frac{\partial}{\partial\beta}\alpha\Phi(x^c(\alpha)) = \frac{\partial}{\partial\alpha}\alpha\Phi(x^c(\alpha))\frac{\partial\alpha}{\partial\beta} = -q\frac{\partial}{\partial\alpha}\alpha\Phi(x^c(\alpha))$). Since X decreases in $\alpha\Phi(x^c(\alpha))$, it will follow that X increases in β. Since ψ decreases in X, this will in turn show that ψ decreases in β. Since f^* decreases in ψ, this will in turn show that f^* increases in β, as was to be shown.

We begin by computing the derivative

$$
\begin{aligned}
\frac{\partial}{\partial\alpha}\alpha\Phi(x^c(\alpha)) &= \Phi(x^c) + \alpha\phi(x^c)\frac{\partial x^c(\alpha)}{\partial\alpha} \\
&= \Phi(x^c) + \alpha\phi(x^c) \times \left(-p\frac{x^c\Phi(x^c)+\phi(x^c)}{1-p+\alpha p\Phi(x^c)}\right) \\
&= \Phi(x^c) + \alpha\phi(x^c) \times \left(p\frac{(x^c\Phi(x^c)+\phi(x^c))x^c}{\alpha p\phi(x^c)}\right) \\
&= \Phi(x^c) + (x^c\Phi(x^c)+\phi(x^c))x^c \\
&= \Phi(x^c) \times (1+(x^c)^2) + \phi(x^c)x^c.
\end{aligned}
$$

(The second line comes from the computation (51) in the Appendix; the third line comes from the equilibrium condition (23) that x^c satisfies, namely $x^c(1 - p + \alpha p\Phi(x^c)) + \alpha p\phi(x^c) = 0$ which - since we know $x^c < 0$ - can be written alternatively as $1 - p + \alpha p\Phi(x^c) = -\frac{\alpha p\phi(x^c)}{x^c}$, and so $-p\frac{x^c\Phi(x^c)+\phi(x^c)}{1-p+\alpha p\Phi(x^c)} = p\frac{(x^c\Phi(x^c)+\phi(x^c))x^c}{\alpha p\phi(x^c)}$). Now define $\Xi(x) \equiv \Phi(x) \times (1+(x)^2) + \phi(x)x$. We claim $\Xi(x)$ is positive for all $x \in \mathbb{R}$. To see that, note that, since $\frac{\partial\phi(x)}{\partial x} = -\xi\phi(x)$:

$$
\begin{aligned}
\frac{\partial\Xi(x)}{\partial x} &= \phi(x) \times (1+(x)^2) + 2\Phi(x)x + \phi(x) - \phi(x)(x)^2 \\
&= 2(\phi(x) + \Phi(x)x).
\end{aligned}
$$

We have shown in the Appendix, at line (52) that $\phi(x) + \Phi(x)x$ is positive for all x. Thus, $\Xi(x)$ is increasing in x. So, $\Xi(x)$ is positive for all x if $\lim_{x\to-\infty} \Xi(x) = 0$. Obviously, $\lim_{x\to-\infty} \phi(x)x = 0$ and $\lim_{x\to-\infty} \Phi(x) = 0$. So, it suffices to show $\lim_{x\to-\infty} \Phi(x)(x)^2 = 0$. Write $\Phi(x)(x)^2 = \frac{\Phi(x)}{\frac{1}{x^2}}$ and apply L'Hospital's rule to conclude $\lim_{x\to-\infty} \Phi(x)(x)^2 = \lim_{x\to-\infty} \frac{\Phi(x)}{\frac{1}{x^2}} = \lim_{x\to-\infty} \frac{\phi(x)}{-\frac{2}{x^3}} = -\frac{1}{2} \lim_{x\to-\infty}(x)^3\phi(x) = 0$. This completes the demonstration that $\lim_{x\to-\infty} \Xi(x) = 0$ and hence that $\Xi(x) > 0$ for all x and hence that $\alpha\Phi(x^c(\alpha))$ is increasing in α for all potential equilibrium cutoffs x^c.■

Proof of Corollary 7

(i) Notice that as r increases, f^* as defined in (44) always strictly increases. This follows from the following sequence of deductions:

$$r \quad \uparrow \quad \Longrightarrow (\psi = (1 - \frac{pr}{\tau + r}(1 - \alpha\Phi(x^c)))^2) \downarrow \Longrightarrow \left(w^2((1-\delta)^2 - (1 - \frac{pr}{\tau + r}(1 - \alpha\Phi(x^c)))^2)\right) \uparrow$$

$$\Longrightarrow \left(\frac{(1-\delta)\delta w^2}{w^2((1-\delta)^2 - (1 - \frac{pr}{\tau+r}(1 - \alpha\Phi(x^c)))^2)}\right) \downarrow \Longrightarrow \left(-\frac{(1-\delta)\delta w^2}{w^2((1-\delta)^2 - (1 - \frac{pr}{\tau+r}(1 - \alpha\Phi(x^c)))^2)}\right) \uparrow.$$

Also notice that in case A of the theorem as r increases from being below r^δ to being above r^δ the entrepreneur's optimal retention stake goes from a positive amount to zero.

(ii) The essential step of this demonstration is to show that (44) is strictly increasing in p. To see this, first recall from the corollary to Theorem 2 that $x^c(p)$ declines in p, so X increases in p, and hence ψ as defined prior to the statement of Theorem 6 decreases in p. Then, similar to the steps taken in the proof of part (i) above, it is easy to show that (44) is strictly increasing in r.

It is easy to check (using logic similar to that above) that the inequality defining case A of the theorem is more likely (and so the inequality defining case B of the theorem is less likely to occur) as β increases. Since when case B holds, f^* is always characterized by (44) whereas in case A, the optimal f is sometimes equal to one (and when the optimal f is not 1, it is also given by (44)), so this is yet another reason why, as β increases, the optimal f always weakly increases.

11 References

Downes, D. and R. Heinkel, "Signalling and the Valuation of Unseasoned New Issues," Journal of Finance 37, pp. 1-10.

Dye, Ronald, "Disclosure of Nonproprietary Information," Journal of Accounting Research Vol. 23, No. 1 (Spring, 1985), pp. 123- 145.

Hughes, Pat," Risk Sharing and Valuation under Moral Hazard," in . *Economic Analysis of Information and Contracts: Essays in Honor of John E. Butterworth,* Gerald A. Feltham, Amin H. Amershi, and William T. Ziemba, eds., Springer 1988.

Jung and Kwon, "Disclosure when the Market is Uncertain of Information Endowment of Managers," Journal of Accounting Research Vol. 26 No. Spring 1988, pp. 146-153.

Leland, H. and D. Pyle, "Informational Asymmetries, Financial Structure, and Financial Intermediation," *Journal of Finance* Vol. 32, No. 2, May 1977.

Ramakrishnan, Ram T. S. and Anjan V. Thakor, "The Valuation of Assets under Moral Hazard," *The Journal of Finance* Vol. 39, No. 1 (Mar., 1984) , pp. 229-238.

Ryan, Ellen and Laura Simmons, "Securities Class Action Settlements 2009," *Cornerstone Research.*

Trueman, Brett, "The Relationship between the Level of Capital Expenditures and Firm Value," *Journal of Financial and Quantitative Analysis* Volume 21, Issue 02, June 1986, pp 115-129.