Optimal Disclosure Decisions When There are Penalties for Nondisclosure*

Ronald A. Dye

August 16, 2015

Abstract

We study a model of the seller of an asset who is liable for damages to buyers of the asset if, after the sale, the seller is discovered to have failed to disclose an estimate of the asset's value that the seller knew prior to sale. The model yields some surprising predictions concerning how the seller's disclosure decision changes with changes in the severity of this liability, and with other parameters of the model, including the precision of the estimate and whether the seller behaves myopically or nonmyopically.

1 Introduction

This paper contains an economic model of voluntary disclosures where the seller of an asset is presumed to have a duty to disclose whatever private information he has about the asset's value to potential buyers prior to the asset's sale, and where - if the seller violates that duty - the seller may be subject to damage payments. Thus, the model studies situations in which the seller's disclosure is mandatory, but the seller's compliance with the mandatory requirements is voluntary. There are a variety of settings in which the theory applies, from selling securities in IPO/investment settings to selling products in manufacturing/consumer settings.

The model is founded on previous models of a seller's optimal voluntary disclosure behavior, where the premise is that the seller sometimes privately receives information about an asset's value prior to the asset's sale which the seller

*The paper is currently undergoing revision. Comments welcome!

may disclose or withhold (as in Dye [1985], Farrell [1986], Jung and Kwon [1988], Shavell [1994], Hughes and Pae [2004], Hughes and Pae [forthcoming], among others). In these prior works, when the seller makes no disclosure, potential buyers of the asset are unsure whether the seller did not receive information about the asset prior to sale or whether he elected to withhold information that he received. This uncertainty regarding the reason for the seller's nondisclosure allows a seller who received unfavorable information about the asset's value to "pool" with sellers who received no pre-sale information about the asset's value by not making a disclosure. After the sale, if buyers discover that the seller withheld information, then buyers will know that they overpaid for the asset based on the information in the seller's possession. In the prior literature, buyers who made such a discovery were posited to have no recourse. The innovation in the present paper is that we presume that buyers do have recourse: sellers caught having previously withheld information are obliged to pay the buyers damages.

We study a base model along with two variants. In the base model, the asset being sold is indivisible and the seller cannot take any actions that influence the subsequent value of the asset to buyer. In one variant, the asset continues to be indivisible but the seller can enhance the asset's value to buyers by taking some action prior to sale; in the second variant, the seller can also enhance the asset's value prior to sale, but the asset is divisible and the seller can decide what fraction of the asset to retain for his own use. In both the base model and its variants, we assume that the seller's penalties for withholding take the natural form of being proportional to the buyer's overpayment for the asset, where "overpayment" is calculated based on the difference between what buyers paid for the asset and what they would have paid for the asset had the seller disclosed his information. The proportionality factor that determines what (possibly fractional) multiple of this overpayment constitutes the damages penalty is a parameter of the model that we refer to as the "damages multiplier."

In the base model and its variants, we obtain a variety of results concerning the seller's equilibrium behavior. Among these results, at least four stand out. To describe the first of these results, call the seller's disclosure decision *myopic* if he ignores the potential damage payments he is subject to if he is caught withholding information and instead bases his disclosure decision just on whether the selling price of the asset is highest with or without disclosure. Call the seller's disclosure decision *nonmyopic* if in deciding whether to disclose information he receives, he takes into account these damage payments. Our first result is that for a wide range of the model's parameter values, the seller's optimal myopic disclosure decision coincides with his optimal nonmyopic disclosure decision. As part of this first result, besides demonstrating the robustness of the coincidence of myopic and nonmyopic disclosure decisions for a wide range of parameter values, we also show that myopic and nonmyopic disclosure policies coincide both when the seller is subject to clawback provisions and when he is not subject to them. (A "clawback provision" is a requirement that the seller disgorge any overpayment made to him by buyers of the asset as a consequence of his withholding, and is a source of reduction in the seller's net receipts from the asset's sale distinct from the damage payments.)

In our second result, we show, also for a wide range of parameter values, that the seller optimally discloses the information he receives *less* often as the size of the damages multiplier *increases*. One might have expected that the seller would disclose his information more often as this penalty increases, but this turns out not to be the case. In our third result, we show that as the damages multiplier increases, the seller chooses to sell a larger fraction of the asset to buyers. One might have expected the seller to sell a smaller fraction of the asset as the damages multiplier gets larger, since the seller is not liable for failure to disclose information about portions of the asset that he retains for himself, and so the benefits to the seller of retaining a large stake in the asset would seem to increase with the seller's liability for nondisclosure, but this also turns out not to be correct. We defer presenting the intuition for the preceding results to the body of the paper.

In the last of the results we highlight here, we show that in the model variant where the seller can both sell just a fraction of the asset and undertake actions that enhance the value of the asset prior to sale, the seller optimally retains a smaller stake in the asset as the quality or precision of the estimate he receives, and sometimes discloses, increases. Among the four highlighted results, this last is perhaps the most intuitive, since as the precision of the estimate improves, the estimate provides a more accurate indicator of the seller's pre-sale activities to enhance the asset's value, and so it becomes less important to rely on the seller's retention of the asset to motivate the seller to undertake the pre-sale value-enhancing activities. When "the asset" is a firm that an entrepreneur starts up and an increase in the precision of the disclosed estimate is interpreted as an improvement in the quality of the entrepreneur's firm's financial accounting reports, this last result suggests the empirically testable implication that improvements in the quality of accounting information leads entrepreneurs to retain smaller ownership stakes in the firms they found.

We deduce many other results besides those highlighted here from the model in the paper, but we leave the presentation and discussion of those results to the main text below.

This paper is part of the growing literature that studies firms' voluntary disclosure decisions that was initiated by the fundamental contributions of Grossman [1981] and Milgrom [1981] in their "unravelling" result. This literature is too large to summarize here, but see for example, the surveys of Gertner [1988], Milgrom [2008], and Dranove and Jin [2010]. Pae [2005] proposed an analysis similar to the one contained here, but Pae did not carry out the analysis. There is also a large empirical literature on voluntary disclosures, surveyed by Healy and Palepu [2001]. The empirical work most closely tied to the present paper is that of Heitzman, Wasley, and Zimmerman [2010] who study empirically firms' disclosure decisions in the presence of mandatory disclosure requirements.

The paper proceeds as follows. The next section, Section 2, introduces the model. Section 3 describes the seller's preferences along with a formal specification of damages payments. Section 4 examines the seller's optimal disclosure policy, and it contains the general findings regarding the equilibrium disclosure probability mentioned above. Section 5 extends the model to a setting where the seller's investment choice is endogenous. Section 6 contains a summary of some of our main findings, and the appendix contains proofs of many of the results not proven in the text or accompanying footnotes.

2 Base model setup

In the base model, S (for "seller") has an asset he wants to sell. There are multiple homogenous potential buyers of the asset. The value of the asset is uncertain to all of these potential buyers at the time S sells the asset and is given by the realization z of the random variable $\tilde{z}$, which is taken to be normally distributed with mean m and variance $\frac{1}{\tau}$, henceforth written as $\tilde{z}\tilde{\ }N(m, \frac{1}{\tau})$. $\tilde{z}'s$ realization occurs after the sale. Before the sale takes place, with probability $p \in (0, 1)$ S privately receives an estimate $\tilde{v}$ of $\tilde{z}$. This estimate $\tilde{v}$ is taken to be unbiased and given by

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

Here, r denotes the precision of the estimate $\tilde{v}$. It is apparent that the prior distribution of $\tilde{v}$ is normal, with mean m and variance $\sigma^2 \equiv \frac{1}{\tau} + \frac{1}{r}$. In the following, the prior density and cdf of $\tilde{v}$ are denoted by $g(v)$ and $G(v)$ respectively. With probability $1 - p$, S receives no estimate before the sale.

Consistent with the now conventional assumptions of the voluntary disclosure literature (see, e.g., Dye [1985] and Jung and Kwon [1988]), we assume that: if S receives no information, he makes no disclosure; in particular, S is presumed incapable of credibly disclosing that he did not receive information;

and if S discloses information, the disclosure is confined to be truthful.

The new feature of the model is that, after the sale takes place, if S disclosed nothing prior to the asset's sale, then a fact finder (investor, reporter, auditor, etc.) undertakes an investigation and, if the reason S disclosed nothing turns out to be that S withheld information, then the fact finder with probability $q \in (0, 1)$ both detects and reports that S withheld information along with what the withheld information was. In the latter event, S is forced to pay damages to the buyer of the asset, as described below. Also, the fact finder fails to discover that S withheld information with probability $1 - q$, but the fact finder never wrongly asserts that S withheld information when that was not the case.

3 Preferences and damage payments

S's disclosure, or nondisclosure, is the only source of information about the asset's value to the asset's potential buyers. All the of buyers are risk neutral, and they all behave competitively. Consequently the equilibrium selling price of the asset is its expected value based on what potential buyers learn or can infer about the asset's value prior to sale.

If S learns v and discloses it, the asset's selling price is denoted $P(v)$. Since buyers are risk neutral, this price $P(v)$ is the asset's expected value conditional on $\tilde{v} = v$. By Bayes' rule applied to normal distributions (see, e.g., DeGroot [1970]), we know that this conditional expected value is given by:

$$P(v) \equiv E[\tilde{z}|v] = \frac{\tau m + rv}{\tau + r}. \tag{2}$$

This price is a weighted average of buyers' initial beliefs about the asset's value m and the disclosed estimate v, with the weights $\frac{\tau}{\tau+r}$ and $\frac{r}{\tau+r}$ determined by the relative precisions of the prior and the estimate.[1]

[1] Even though at this point in the expostion we have yet to describe what constitutes an equilibrium of the model, it is worth pointing out for future reference that the specification of the price $P(v)$ in (2) implicitly entails making off-equilibrium specifications, because this price is being specified for all v, and not just for those v that S is expected to disclose in

We let P^{nd} denote the selling price of the asset when the seller makes no disclosure. If S withheld v and the fact finder subsequently finds that out, then the difference $P^{nd} - P(v)$ constitutes the amount the buyer overpaid for the asset based on the seller's withheld information.[2] In this event, we suppose that S has to pay the damages payment

$$\beta \times (P^{nd} - P(v)) \tag{3}$$

to the buyer. β is formally what we referred to in the Introduction as the "damages multiplier."

The specification of the damages payment does not indicate the full consequences to S of getting caught withholding information from buyers, because it does not address the "clawback" question. If there is a clawback, S has to return the overpayment $P^{nd} - P(v)$ to buyers, so when S is caught having withheld information, S will receive net (equivalently, the buyer will pay net[3]) $P(v) - \beta \times (P^{nd} - P(v))$. If $\beta > 0$, the buyer winds up owning an asset worth (in expectation) $P(v)$ but pays net only $P(v) - \beta \times (P^{nd} - P(v))$ for it. If $\beta = 0$, the buyer ends up paying $P(v)$ for the asset which is exactly what the asset is worth. If $\beta \in (-1, 0)$, the buyer ends up losing on the purchase, because he receives an asset worth $P(v)$ but pays $P(v) - \beta \times (P^{nd} - P(v)) > P(v)$ for it. Nevertheless, in this last event, conditional on having purchased the asset, the buyer is better off having the fact finder detect the seller's withholding (than not having the withholding detected), because the buyer receives a positive amount from the fact finder's efforts. If $\beta = -1$, the buyer is stuck with paying P^{nd} for the asset (since $P^{nd} = P(v) - \beta \times (P^{nd} - P(v))$ when $\beta = -1$) and hence

equilibrium. Clearly, the off-equilibrium specification of $P(v)$ in (2) is *the* uniquely natural way to specify $P(v)$ when v is disclosed, regardless of how other facets of the equilibrium are defined.

[2] We shall show below that when the seller withholds information, it is the case that $P^{nd} > P(v)$, so $P^{nd} - P(v)$ is indeed an *over*payment.

[3] The equivalence between amounts paid for by the seller and received by the buyer holds only when there is no "slippage" due to attorneys' fees or other legal costs. The analysis can be extended to cover cases of slippage.

is no better off than he would have been had the fact finder not detected S's withholding. In the following, when studying the clawback case, we confine attention to damage multipliers bounded below by $\beta \geq -1$ and bounded above so that the inequality

$$q(1+\beta) < 1 \tag{4}$$

holds. Additional discussion concerning this bound may be found in the accompanying pair of footnotes.[4] [5] Before S knows whether or not he will receive the estimate $\tilde{v}$ - S's expected proceeds from the buyer net of his cost of damage payments is given by:

$$E[(1-p)P^{nd} + p\max\{P(\tilde{v}), (1-q)P^{nd} + q(P(\tilde{v}) - \beta(P^{nd} - P(\tilde{v})))\}]. \tag{5}$$

We conclude this section with a short discussion of the "no-clawback" case. In this case, S keeps the original payment P^{nd} from buyers if he is caught withholding information, and pays the damages payment $\beta(P^{nd} - P(v))$, and so gets net $P^{nd} - \beta(P^{nd} - P(v))$. For there to be content for the damages payment (3) to be a penalty in this case, the damages multiplier β must be positive. As was true of the clawback case, our analysis of the no-clawback case also requires imposing an upper bound on β. In this case, we require β to be positive and that the following (weaker) upper bound hold:

$$q\beta < 1. \tag{6}$$

[4] In the context of securities litigation, buyers/investors are typically not made even close to being made whole by the damages payment when firms are found to have improperly withheld information, but investors usually receive some payment following a fact finder's discovery that a firm withheld information, so in the securities litigation context, $\beta \in (-1, 0)$. See, e.g., Ryan and Simmons [2009].

[5] The inequality (4) has an intuitive economic motivation. If one combines the clawback with the penality specified in (3), the total penalty (i.e., drop in price from P^{nd}) for S derived from getting caught withholding information is $P^{nd} - P(v) + \beta \times (P^{nd} - P(v)) = (1+\beta) \times (P^{nd} - P(v))$, and since the fact finder detects S's withholding with probability q, the expected total penalty is $q(1+\beta) \times (P^{nd} - P(v))$. The inequality (4) restricts S's expected total penalty for withholding to be no larger than the absolute amount by which the firm's market value is overstated as a consequence of S's withholding.

4 The equivalence of nonmyopic and myopic disclosure strategies

Suppose the seller receives information v. As we described in the Introduction, we call the seller's disclosure decision *myopic* if the seller decides whether to disclose v solely based on whether $P(v)$ is bigger or smaller than P^{nd}. That is, the seller's disclosure decision is myopic if the seller disregards the potential damages he is subject to were he to withhold v and the fact finder subsequently detects his withholding. In contrast, we call the seller's disclosure decision *nonmyopic* if he takes into account these potential damage payments when making his disclosure decision. Our first result below asserts that:

Lemma 1 *As long as the parameter restriction (4) holds in the clawback case, or as long as the parameter restriction (6) holds in the no-clawback case, the seller's optimal nonmyopic disclosure decision coincides with the seller's optimal myopic disclosure decision.*

The proof is simple for both cases. In the clawback case, the seller's optimal nonmyopic disclosure entails nondisclosure iff

$$(1-q)P^{nd} + q(P(v) - \beta(P^{nd} - P(v))) > P(v). \tag{7}$$

Collecting terms in LHS(7) that involve P^{nd} together, and also collecting terms in LHS(7) involving $P(v)$ together, nondisclosure is S's best nonmyopic choice iff:: $(1-q-q\beta)P^{nd} + (q+q\beta)P(v) > P(v)$,or equivalently, iff

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

When the bound (4) holds, this last inequality is obviously equivalent to:

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

This proves the lemma in the clawback case.

In the case of no-clawback, the seller's optimal nonmyopic disclosure entails nondisclosure iff

$$P^{nd} - q\beta(P^{nd} - P(v)) > P(v),$$

i.e., iff

$$P^{nd} - P(v) - q\beta(P^{nd} - P(v)) > 0,$$

i.e., iff

$$(P^{nd} - P(v))(1 - q\beta) > 0. \tag{9}$$

Clearly, when the bound $1 - q\beta > 0$ holds, this last inequality is equivalent to:

$$P^{nd} > P(v).$$

This proves the lemma in the no-clawback case.

The result in both cases obtains because when damage payments are proportional to buyers' overpayments, then the optimal nonmyopic disclosure decision is based on the difference between the "no disclosure" price of the asset and the price of the asset had S disclosed his information, and so results in the optimal nonmyopic disclosure decision coinciding with the optimal myopic disclosure decision.[6]

To avoid studying a proliferation of cases in the following, for the most part, we confine attention to the "non-clawback" case and, unless the contrary is explicitly stated, we assume the bound (6) holds. Given those maintained assumptions, it follows, in view of the preceding lemma, that we do not have to provide additional separate analyses depending on whether S is believed to behave myopically or nonmyopically in making his disclosure decision.

[6] It is perhaps useful to mention what happens if the multiplier β is so large that exceeds the upper bounds ((4) or (6)) referenced above. Consider the no-clawback case when $q\beta > 1$. Then according to (9), the optimal myopic and the optimal nonmyopic disclosure decisions are the opposite of each other: when S behaves in the optimal nonmyopic fashion, it is optimal for S to disclose v when $P(v)$ is smaller than P^{nd}, and it is optimal for S to withold v when $P(v)$ is larger than P^{nd}.

5 Determination of the equilbrium in the base model

We now turn to a discussion of how the equilibrium no disclosure price P^{nd} is set. Clearly, if S prefers to disclose rather than withhold v, then S will also prefer to disclose rather than withhold any $v' > v$, and so the nondisclosure set is given (for both myopic andor nonmyopic sellers) by a left-tailed interval of the form:

$$ND \equiv \{v|P^{nd} > \frac{\tau m + rv}{\tau + r}\}, \tag{10}$$

or equivalently by:

$$ND = \{v|v^c > v\}. \tag{11}$$

where v^c is defined implicitly by the solution to the equation

$$\frac{\tau m + rv^c}{\tau + r} = P^{nd}. \tag{12}$$

To specify P^{nd} more precisely, we next calculate buyers' perceptions of the expected value of the asset given that S makes no disclosure and also given that the buyer thinks S uses the cutoff v^c in deciding whether to disclose the private information he receives. There are three events related to the seller's nondisclosure that must be handled separately. (1). The seller did not receive any information. (2). The seller received and withheld information and the fact finder subsequently fails to detect that S withheld information. (3). The seller received and withheld information and the fact finder subsequently detects that S withheld information. Since buyers' ex ante perceptions of the probability S will make no disclosure is given by $1 - p + pG(v^c)$ (this is the probability S receives no information plus the probability that S receives information that is below the cutoff v^c), buyers can apply Bayes' Rule to conclude that the probability S received no information, conditional on S making no disclosure, i.e., buyers' perceptions of the probability of event (1) as described above given no disclosure is:

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

Likewise, the buyer's perceptions of the probability of events (2) and (3) above, given S makes no disclosure, are respectively:

$$\frac{p(1-q)G(v^c)}{1-p+pG(v^c)} \tag{14}$$

and

$$\frac{pqG(v^c)}{1-p+pG(v^c)}. \tag{15}$$

Next notice that, conditional on event (1), buyers expect to receive an asset whose value is given by

$$E[\tilde{z}] = m. \tag{16}$$

(Were buyers to know that the reason S made no disclosure was that S received no information, buyers would have no reason to make a "lemons" inference about the asset's value from S's nondisclosure, so the asset's expected value in that event is just the asset's unconditional expected value.) Buyers' perceptions of the asset's expected value conditional on no disclosure and event (2) is given by:

$$E[\tilde{z}|\tilde{v} < v^c] = E[E[\tilde{z}|\tilde{v}]|\tilde{v} < v^c] = E[\frac{\tau m + r\tilde{v}}{\tau + r}|\tilde{v} < v^c]. \tag{17}$$

(If buyers knew S withheld information and that withholding would not be subsequently detected by the fact finder, buyers would calculate the conditional expected value of the asset to be $E[\tilde{z}|\tilde{v} < v^c]$.) Finally, to calculate buyers' perceptions of the asset's expected value conditional on no disclosure and event (3), first fix a particular $v < v^c$ and suppose that buyers knew the fact finder was going to discover that S withheld this particular v. Then, the value the buyers would attach to purchasing the asset inclusive of the damage payments they expect to receive in this case is given by:

$$E[\tilde{z}|v] + \beta(P^{nd} - E[\tilde{z}|v]) = \frac{\tau m + rv}{\tau + r} + \beta(P^{nd} - \frac{\tau m + rv}{\tau + r}).$$

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

$$E[\frac{\tau m + r\tilde{v}}{\tau + r} + \beta(P^{nd} - \frac{\tau\hat{m} + r\tilde{v}}{\tau + r})|\tilde{v} < v^c]. \tag{18}$$

Thus, when buyers believes the seller uses the cutoff v^c in deciding whether to disclose the estimate he receives, the total value buyers expect to receive, conditional only on no disclosure by the seller, consists of the sum of the products of (13) and (16), (14) and (17), (15) and (18), i.e., consists of:

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

which may be rewritten as:

$$\frac{(1-p)m + pG(v^c)E[\frac{\tau\hat{m} + r\tilde{v}}{\tau + r}|\tilde{v} < v^c] + pqG(v^c)\beta E[P^{nd} - \frac{\tau\hat{m} + r\tilde{v}}{\tau + r}|\tilde{v} < v^c]}{1 - p + pG(v^c)}. \tag{19}$$

When S makes no disclosure, competition among the buyers will drive the price P^{nd} that S receives to this conditional expected value (19). Thus, the no disclosure price P^{nd} will satisfy:

$$P^{nd} = (19). \tag{20}$$

Now, recalling the link between v^c and P^{nd} described in (12) above, it follows from (19) and (20) that in equilibrium, the cutoff v^c must satisfy the equation:

$$\frac{\tau m + rv^c}{\tau + r} = \frac{(1-p)m + pG(v^c)E[\frac{\tau m + r\tilde{v}}{\tau + r}|\tilde{v} < v^c] + pqG(v^c)\beta E[\frac{\tau m + rv^c}{\tau + r} - \frac{\tau m + r\tilde{v}}{\tau + r}|\tilde{v} < v^c]}{1 - p + pG(v^c)}. \tag{21}$$

Rearranging this last equation (see the appendix for details) leads to the following theorem. In the statement of the theorem, $\phi(\cdot)$ and $\Phi(\cdot)$ refer to the density and cdf of a standard normal random variable $\tilde{x}$ respectively, and α is defined by $\alpha \equiv 1 - q\beta$.

Theorem 2 *The v^c that solves (21) is given by*

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

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, combined with the next three corollaries below, constitutes the main result of this section. It establishes the link between the equilibrium cutoff v^c and a "standardized" equilibrium cutoff x^c defined in terms of the density and cdf of a standard normal random variable associated with the solution to (23).

The theorem has several consequences. First, since the equilibrium probability that S will make no disclosure, given that it received information, is $\Pr(\tilde{v} \leq v^c)$, and the transformed random variable $\frac{\tilde{v}-m}{\sigma}$ has a standard normal distribution, it follows from (22) that

$$\Pr(\tilde{v} \leq v^c) = \Pr(\frac{\tilde{v} - m}{\sigma} \leq \frac{v^c - m}{\sigma}) = \Phi(x^c). \tag{24}$$

Since both the cdf $\Phi(x^c)$ and the density $\phi(x^c)$ are defined independently of each of: the mean m and variance σ^2 of $\tilde{v}$, and also the precision r of the estimate, we conclude:

Corollary 3 *The equilibrium probability of disclosure (and hence also the equilibrium probability of no disclosure) is independent of each of: the mean m and variance $\sigma^2 (= \frac{1}{\tau} + \frac{1}{r})$ of the estimate $\tilde{v}$, and also independent of $\tilde{v}'s$ precision r.*

In particular, this last corollary implies that were, say, the prior mean m of $\tilde{v}$ (equivalently, the prior mean of the asset's value $\tilde{z}$) to be made endogenous, then the equilibrium probability of disclosure is unaffected by whatever turns out to be the equilibrium value of that investment choice. This observation will be important in extensions of the model to be presented in later sections of the paper.

The independence of the equilibrium cutoff highlighted in this last corollary from the prior mean and variance of $\tilde{v}$ also implies that if the model were altered so that additional public information, say $\tilde{y}$, were disclosed before S had an opportunity to receive or disclose the private estimate $\tilde{v}$, and this public information $\tilde{y}$ was such that buyers' posterior beliefs about $\tilde{v}$ were still normally distributed (as would be true, for example, if $\tilde{y} = \tilde{z} + \tilde{\omega}$ for some normal random variable $\tilde{\omega}$ that is independent of all other variables in the model), then the release of this public information has no impact on S's subsequent voluntary disclosure decisions. That is to say, in this model:

Corollary 4 *The disclosure of public information before S has an opportunity to receive or disclose his private information has no effect on S's disclosure decision, and hence public disclosures are neither complementary to nor a substitute for voluntary disclosures.*

This corollary speaks to a common concern in the disclosure literature as to whether mandatory disclosures may "drive out" voluntary disclosures or more generally how the release of public information may influence private information production and/or disclosure decisions.[7]

Third, the theorem is the basis for the following additional comparative statics.

Corollary 5 *As long as the bound $q\beta < 1$ holds, the unique x^c that solves (23):*

(i) is negative;

(ii) implies that the ex ante probability S will make a disclosure is at least $p/2$;

(iii) is strictly decreasing in p;

(iv) is strictly increasing in β;

[7] In models related to, but distinct from, the present model, where the emphasis is not on voluntary disclosure but rather on the related phenomenon of traders' costly private information acquisition activities, sometimes very different conclusions emerge. E.g., Diamond [1985] shows that public disclosures can reduce investors' incentives to acquire information on private account if the public disclosures are sufficiently informative.

(v) is strictly increasing in q.

Before discussing these comparative statics individually, we make two observations that apply to all of them. First, in view of (24) and since $\Phi(\cdot)$ is an increasing function, each of these comparative statics is also a comparative static about the *probability* that S does not make a disclosure in equilibrium. Thus, for example, since part (iv) of the corollary asserts that x^c increases in β, it follows that the probability that S makes no disclosure in equilibrium also increases in β.[8] Second, each of these comparatives statics is also a comparative static about how the equilibrium cutoff v^c that solves (21) changes in any parameter, in view of the monotone relationship between x^c and v^c described in Theorem 2, namely that $v^c = \sigma x^c + m$. Thus, for example, since the corollary asserts that in equilibrium x^c increases in β, it follows that the equilibrium cutoff v^c 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 $\tilde{v}$). Part (ii) then follows immediately: since $x^c < 0$, the symmetry of the density of a standard normal random variable around $x = 0$ implies that $p(1 - \Phi(x^c)) > p/2$. In words, we obtain the robust conclusion that for any probability p that S receives information, the ex ante probability S discloses his information always exceeds $p/2$ regardless of the precision of the information S receives, regardless of the damages multiplier β applicable if S is caught withholding information, etc.

Part (iii) shows that the cutoff is decreasing in the prior probability S receives

[8] As applied to parameter p, the ex ante probability S receives information, this last statement in the text is also true (that is, since (iii) asserts that x^c is strictly decreasing in p, it follows that the ex ante probability that S will not disclose information is also decreasing in p), but the claim requires the following extra observation beyond what is stated in the corollary to confirm it: the ex ante probability that S does not make a disclosure is $1-p+pG(x^c(p))$. We contend that this probability declines in p. The derivative of this probability with respect to p is given by: $\frac{\partial}{\partial p}(1-p+pG(x^c(p))) = -1+G(x^c(p))+pg(x^c(p))\frac{\partial x^c(p)}{\partial p} = -(1-G(x^c(p)))+pg(x^c(p))\frac{\partial x^c(p)}{\partial p}$. Since the corollary asserts that $\frac{\partial x^c(p)}{\partial p} < 0$, we can further conclude that $\frac{\partial}{\partial p}(1-p+pG(x^c(p))) < 0$ too, which is the substance of the claim made in the text.

information. This result extends Jung and Kwon [1988] to situations where a value-maximizing firm (or S) is subject to potential damage payments when the (potentially disclosed) estimate of the firm's cash flows is normally distributed.

Part (iv) of the corollary asserts that x^c increases in β . At first blush, this result seems so counterintuitive as to be wrong, as it asserts that an increase in the damages multiplier β *reduces* the probability S will disclose the information he receives. But, the result has an easy explanation. Recall from Lemma 1 that, as long as the inequality in (6) is maintained, S's optimal nonmyopic disclosure policy is determined by the (myopic) comparison of P^{nd} and $P(v)$ in (8). While the parameter β does not appear explicitly in (8), β does appear implicitly in (8) through P^{nd}. P^{nd} increases in β (as long as the bound $q\beta < 1$ is preserved) because as β increases, buyers will receive a larger damages payment in the event the fact finder catches S withholding information from them. Since P^{nd} increases in β, it follows that inequality (10) will hold for more values of v as β increases, and hence S will withhold the information he receives more often.

The explanation for why $\frac{\partial x^c}{\partial q}$ is positive in part (v) is similar to the explanation for part (iv). When S makes no disclosure, the value buyers attaches to purchasing the asset is the sum of the buyers' perceptions of the value of the asset itself combined with the expected value of the damages claim. Since the expected value of the damages claim increases as the probability q the fact finder detects the withholding increases, it follows that P^{nd} increases with q, so (as in part (iv)) of the corollary) inequality (10) will hold for more values of v as q increases

5.1 Extension 1: When S can make a pre-sale investment that affects the expected value of the asset

In this section, we extend the base model to a setting where the prior mean m of the asset is endogenous and is affected by an investment I S selects prior to the asset's sale. Examples of settings where m is naturally endogenous include: the asset is a used car S drove prior to sale whose value is affected by how much care S took in maintaining the car; the asset is a business started by S which he wants to sell because he is retiring, and the value of that business is affected by marketing, production, and other actions that S took while the business was under his management, etc.

Formally, we now suppose that if S selects private investment $I \geq 0$ at personal cost $.5I^2$, this investment results in the distribution of the asset's value at the time of sale $\tilde{z}$ being distributed $\tilde{z}\tilde{\ }N(m(I), \frac{1}{\tau})$, for $m(I) = w \times I$. Here, w is the marginal productivity of investment. S's actual investment I is taken to be a private choice of S, so buyers must make a conjecture about I when trying to assess the asset's value. We let buyers' conjecture (assumed to be common to all buyers) be denoted by $\hat{I}$. In equilibrium we require that buyers' conjectures be correct: $I = \hat{I}$.

The rest of the model is the same as the base model described above: subsequent to selecting I but before the sale of the asset, with probability p S receives estimate v, the realization of the random variable $\tilde{v}$ distributed as in (1) above. The prior density and cdf of $\tilde{v}$ are now written as $g(v|I)$ and $G(v|I)$ to acknowledge their dependence on I. If S receieves v, S can elect whether to disclose or withhold v. If S withholds v, a fact finder detects the withholding with probability q, and if the fact finder does so, then S is liable for damages as described above.

Taking as given buyers' conjecture $\hat{I}$ about S's initial investment choice, then equation (21) above still uniquely defines the cutoff v^c – now denoted by $v^c(\hat{I})$ – that S will use in deciding whether to disclose the information he receives, provided the prior mean m is replaced by $m(\hat{I}) \equiv \hat{m}$ and $G(v^c)$ is replaced by

$G(v^c(\hat{I})|\hat{I})$. That is, the cutoff $v^c(\hat{I})$ is now defined by the equation:

$$\begin{aligned} & \frac{\tau\hat{m} + rv^c(\hat{I})}{\tau + r} \qquad (25) \\ = \quad & \frac{(1-p)\hat{m} + pG(v^c(\hat{I})|\hat{I})E[\frac{\tau\hat{m}+r\tilde{v}}{\tau+r}|\tilde{v} < v^c(\hat{I}), \hat{m}] + pqG(v^c(\hat{I})|\hat{I})\beta E[\frac{\tau\hat{m}+rv^c}{\tau+r} - \frac{\tau\hat{m}+r\tilde{v}}{\tau+r}|\tilde{v} < v^c(\hat{I}), \hat{m}]}{1 - p + pG(v^c(\hat{I})|\hat{I})}. \end{aligned}$$

The "no disclosure" price of the asset, now written as $P^{nd}(\hat{I})$, is connected to the cutoff $v^c(\hat{I})$ just as the no disclosure price P^{nd} was connected to the cutoff v^c in the base model above via (12), i.e.,

$$P^{nd}(\hat{I}) = \frac{\tau\hat{m} + rv^c(\hat{I})}{\tau + r}. \qquad (26)$$

At the time S initially chooses I, S takes buyers' conjecture $\hat{I}$, along with both the cutoff $v^c(\hat{I})$ and the no disclosure price $P^{nd}(\hat{I})$ implied by that conjecture, as given when deciding what investment I actually to adopt. Specifically, S will choose I so as to maximize:

$$\begin{aligned} & OBJ(I|\hat{I}) \\ \equiv \quad & (1 - p + pG(v^c(\hat{I})|I))P^{nd}(\hat{I}) + p\int_{v^c(\hat{I})}^{\infty} \frac{\tau\hat{m} + rv}{\tau + r} g(v|I)dv \qquad (27) \\ & -pq\beta \int_{-\infty}^{v^c(\hat{I})} (P^{nd}(\hat{I}) - \frac{\tau\hat{m} + rv}{\tau + r})g(v|I)dv - .5I^2. \end{aligned}$$

The first term in (27), $(1 - p + pG(v^c(\hat{I})|I)) \times P^{nd}(\hat{I})$, is the product of the ex ante probability S will make no disclosure (taking into account both buyers' conjectures and S's actual investment choice) and the price S will get if he makes no disclosure; the second term in (27), $p\int_{v^c(\hat{I})}^{\infty} \frac{\tau\hat{m}+rv}{\tau+r} g(v|I)dv$, is the ex ante probability S will make a disclosure, $p(1-G(v^c(\hat{I})|I))$, times the conditional expected value of the asset if he makes a disclosure, $E[\tilde{v}|\tilde{v} > v^c(\hat{I}), I]$.[9] The third term in (27), $pq\beta\int_{-\infty}^{v^c(\hat{I})}(P^{nd}(\hat{I}) - \frac{\tau\hat{m}+rv}{\tau+r})g(v|I)dv$, is the ex ante probability S will be subject to damages payments, $pqG(v^c(\hat{I})|I)$, times the conditional expected value of those damages payments $E[\beta \times (P^{nd}(\hat{I}) - \tilde{v})|\tilde{v} \leq v^c(\hat{I}), I]$.[10]

[9] Since, when multiplied out: $p(1 - G(v^c(\hat{I})|I)) \times E[\tilde{v}|\tilde{v} > v^c(\hat{I}), I] = p\int_{v^c(\hat{I})}^{\infty} \frac{\tau\hat{m}+rv}{\tau+r} g(v|I)dv)$

[10] Since, when multiplied out: $pqG(v^c(\hat{I})|I) \times E[\beta(P^{nd}(\hat{I}) - \tilde{v})|\tilde{v} \leq v^c(\hat{I}), I] = pq\beta\int_{-\infty}^{v^c(\hat{I})}(P^{nd}(\hat{I}) - v)g(v|I)dv$.

Given the preceding definition of $OBJ(I|\hat{I})$, an equilibrium is given as follows.

Definition 6 *An equilibrium investment level consists of an I^* that satisfies: $I^* = \arg\max_I \; OBJ(I|I^*)$.*

Identifying the equilbrium investment level I^* leads to a complete specification of the equilibrium since it determines both the equilibrium cutoff $v^c(I^*)$ and the equilibrium no disclosure price $P^{nd}(I^*)$.

Recalling that $\alpha \equiv 1 - q\beta$ and that x^c was specified in Theorem 2 above, and setting

$$X \equiv \frac{pr\left(1 - \alpha\Phi(x^c)\right)}{\tau + r}, \tag{28}$$

we prove in the appendix the following theorem, which fully characterizes the equilibrium value I^*.

Theorem 7 *The equilibrium value I^* is given by:*

$$I^* = wX. \tag{29}$$

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

Corollary 8 *I^* is:*

(a) increasing in r;

(b) increasing in p;

(c) decreasing in τ;

(d) increasing in β;

(e) increasing in q.

All the results in the corollary intuitive: consider part (a). It asserts that S's investment optimally increases in the precision r of the estimate $\tilde{v}$. As the precision of the estimate increases, $\tilde{v}$ reflects more accurately the action I S took when S makes a disclosure. Anticipating this, S is encouraged to choose a higher level of investment. Virtually the same explanation applies to explain part (b): as the probability S receives the estimate $\tilde{v}$ increases, S is more likely to disclose the estimate, and anticipating this, S is more inclined to pick a high value for I. Next, consider the result in part (c): as buyers' priors beliefs about the asset's value become tighter (τ goes up), then buyers will place relatively less weight on the estimate S sometimes discloses to them and more weight on their prior beliefs about the asset's value when valuing the asset after S makes a disclosure. Since S can do nothing to affect buyers' prior beliefs about the asset's value, increases in τ discourage S from choosing a high level of investment.

Consider part (d). Increases in β increase the liability S has to pay in the event his withholding gets caught. S can reduce the probability of having to pay any damages if he works harder, i.e., chooses a higher level of investment, because the realized value of $\tilde{v}$ is more likely to be above the cutoff, and hence S is less likely to have an incentive to withhold the information and be subject to the liability payment. Likewise, as part (e) reports, increases in q increase the likelihood that S will have to pay damages, holding I fixed. As was the case for the explanation of part (d), S can reduce the likelihood of having to pay such damages by increasing I.

5.2 Extension 2: When S sells a divisible asset

In this section, we consider another variant of the first extension above, where now the asset S sells is divisible, and so S can choose, if he so desires, to sell only a fraction of the asset. We now also suppose that the realized value of the asset is determined far into the future after the sale takes place, so discounting becomes important to assess the expected present value of the asset. This extension is appropriate for considering, among other things, disclosures in an IPO setting, where S is viewed as an entrepreneur who sells some fraction of his asset/firm to outsider investors, and these outside investors share with him in the eventual cash proceeds generated by the firm. In this extension, unlike the first extension, no sale need take place: S can retain all of the asset for himself. Since we continue to assume that both S and the buyers of the asset are risk neutral, we need to introduce some additional feature to the model to motivate S to sell a nonzero fraction of the asset to the buyers. The feature we now add is that we suppose that S discounts the cash flows generated by the asset at a higher rate than do the buyers.[11] Specifically, we assume that the eventually realized cash flows z produced by the asset have present value δz to S, for some positive constant $\delta < 1$. Without loss of generality, we normalize the present value of those same cash flows to buyers of the asset to be z, and we further assume that the cash flows received or paid by S related to the (possibly fractional) sale of the asset and ensuring liability assessed on S for withholding information, if any, arrive or are paid sufficiently close in time to S's disclosure decision that S evaluates those receipts and payments at their undiscounted values. These latter conventions are irrelevant to the conclusions we reach below and are adopted solely to eliminate the notational clutter associated with the introduction of additional discount factors were these conventions not adopted.

With these assumptions and conventions in place, we see that if S decides to retain fraction f of the asset, chooses investment I, learns $\tilde{v} = v$ and withholds it, then his expected utility before knowing whether his withholding will be detected is given by:

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

Similar expressions hold for other possible scenarios, e.g., if S decides to retain fraction f of the asset, chooses investment I, learns $\tilde{v} = v$ and discloses it, then his expected utility at the time of disclosure is given by $\delta(1-f) \times E[\tilde{z}|I, v] +$

[11] The difference between S's and buyers' discount rates is natural, and can be motivated by S's life cycle or liquidity demands or S's relative lack of diversification.

Demarzo and Duffie [1999] use this same exogenous assumption in their liquidity based model of security design.

$f \times P(v) - .5I^2$, etc.

In ongoing work, we allow for S's decision about the fraction of the asset to sell to buyers to be delayed until S learns the realization of the estimate $\tilde{v}$ or learns that he is not going to learn the realization of the estimate. In the present analysis, we adopt the simplification that S chooses the fraction f of the asset to sell at the same time he makes his investment choice I. The tradeoff S faces in deciding what fraction of the asset to retain is the following: since buyers attach a higher present value to the cash flows eventually produced by the asset than S does (because of their lower discount rate), that encourages S to sell a large fraction of the asset so as to, in effect, arbitrage the difference between his and buyers' discount rates. But, offsetting that pressure to sell most or all of the asset is the pressure to retain the asset due to the positive incentive effects on S's pre-sale investment choice arising from retention (due to the fact that the cash flows eventually produced by the asset are more informative about S's initial investment choice than is the estimate S occasionally discloses, and the extra informativeness of those cash flows influences S's pre-sale investment only when S retains ownership of those cash flows). S's optimal fractional sales of the asset balances these two competing effects.

We start the analysis of this extension by observing that the ex ante value to S of choosing investment I when buyers conjecture he has chosen investment $\hat{I}$, and selling fraction f of the asset to the buyers is given by:

$$OBJ \equiv \delta(1-f)m(I) + f \times \Psi(I,\hat{I}) - .5I^2, \tag{30}$$

where

$$\begin{aligned}\Psi(I,\hat{I}) \equiv\ & (1-p+pG(v^c(\hat{I})|I)) \times \frac{\tau m(\hat{I}) + rv^c(\hat{I})}{\tau + r} + \\ & p\int_{v^c(\hat{I})}^{\infty} \frac{\tau m(\hat{I}) + rv}{\tau + r} g(v|I)dv - pq\beta \int_{-\infty}^{v^c(\hat{I})} \left(\frac{\tau m(\hat{I}) + rv^c(\hat{I})}{\tau + r} - \frac{\tau m(\hat{I}) + rv}{\tau + r}\right) g(v|I)dv.\end{aligned} \tag{31}$$

Here, $f \times \Psi(I,\hat{I})$ is the price S anticipates receiving from buyers when they purchase fraction f of the asset from him net of the expected damage payments he subsequently expects to pay buyers.[12]

The equilibrium value of OBJ in (30) can be obtained in two steps: first determine the equilibrium value of the investment $I^*(f)$ for any fixed f, and second, determine the optimal f. Using derivations similar to those which led to Theorem 7 of the previous section, one can show that the solution to this first step yields the following natural parallel to Theorem 7.

Theorem 9 *If S chooses to retain fraction f of the asset, then S's equilibrium investment level is unique and given by:*

$$I^*(f) = \delta w + fw \times (X - \delta). \tag{32}$$

Note that when $\delta = 1$ and $f = 1$, then $I^*(f) = wX$, which is the same as the equilibrium level of investment in the first extension above of the base model above. It is easy to show that all of the comparative statics associated with Theorem 7 hold here too, and in addition, in the present setting we get two additional comparative statics: $I^*(f)$ increases in δ, and $I^*(f)$ increases or decreases in f depending on whether X is bigger or smaller than δ. The former result is straightforward: as S's discount rate falls (or equivalently, as S's discount factor rises), the present value of the portion of the asset S retains increases in present value, which causes S to work harder. The latter result indicates that when S discounts the future more than buyers (i.e., when $\delta < 1$), then it is not always true that S has a greater incentive to work hard as his retained ownership stake in the asset increases. Whether that result obtains depends, as (32) shows, on how big S's discount factor δ is relative to X.

The next theorem completely characterizes S's optimal choice of what fraction f of the asset to sell to buyers. The statement of the theorem makes use

[12] Each of the components of $\Psi(I, \hat{I})$ has a natural counterpoint to each of the components in (27) above, when we substitute $\frac{\tau m(\hat{I}) + r v^c(\hat{I})}{\tau + r}$ for $P^{nd}(\hat{I})$, and so we forego explaining each of the components of $\Psi(I, \hat{I})$ here.

of the function:

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

Here, $X(r)$ is the same X as defined in (28) above, where the dependence of X on r is now emphasized.[13]

Theorem 10 *(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 equilibrium* f *is given by*

$$f^* = \frac{(1-\delta)\delta}{(1-X)^2 - (1-\delta)^2}; \tag{33}$$

(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 equilibrium* f *is given by (33).*

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

Corollary 11 *S's equilibrium retained ownership stake in the asset,* $1 - f^*$*, always (at least weakly):*

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

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

(iii) declines as the damages multiplier β *increases.*

The first two comparative statics (i) and (ii) are intuitive. Holding his investment choice fixed, as we already noted, S has an incentive to sell 100% of the

[13] The critical observation that underlies this next theorem is the calculation of the derivative of OBJ with respect to f, when evaluated at the equilibrium value $\hat{I}(f) = I^*(f)$. It can be shown to be given by:

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

asset to buyers, since they value the asset more than he does (because of their lower discount rates). But, since S's investment choice is endogenous, and depends in part on what fraction of the asset he retains, his ex ante investment choice may be inefficiently low unless he retains a substantial ownership stake in the asset. However, as the quality of the estimate $\tilde{v}$ (as measured by the precision of the estimate) S receives increases, or as the probability S receives the estimate increases, the estimate $\tilde{v}$ can be relied on more to ensure that S has good incentives to work hard to invest in the asset even if he sells some fraction of his original ownership stake. Hence, S can profitably sell a larger stake in the asset as either r or p increases.

As increases in the precision r of the estimate can be interpreted in practice as an improvement in the quality of the accounting information S provides to investors, the first result has the empirically testable implication that improvements in financial reporting result in lower retained equilibrium ownership stakes by entrepreneurs who found a company.

The explanation for the third comparative static (that S's optimal retained ownership stake declines as the damages multiplier β increases) is somewhat more complicated. One might think that an increase in β would increase S's optimal retained ownership stake in the asset, since an increased ownership stake, i.e., a reduced sale to buyers, economizes on S's liability risk. For example, S is obviously exposed to no liability for nondisclosure if he sells none of the asset to buyers. Further reinforcing this effect is that, as was noted above in Corollary 5 (iv), an increase in β leads to an increase in the threshold x^c,[14] and hence a reduction in S's propensity to disclose information, and in turn, more exposure to liability for having withheld information. Call these effects combined "the liability effect."

But, there is an offsetting benefit to selling a higher fraction of the asset to

[14] It is easy to check that that corollary remain valid if we replace the assumption there that S sells 100% of the asset by the assumption that S sells any *fixed* fraction of the asset.

buyers when β increases, because of the tight connections between x^c and v^c (recall (22)), and between v^c and P^{nd} (recall (12)). These connections imply that an increase in the damages multiplier β increases the no disclosure price P^{nd}. This latter effect, which we shall call "the price effect," by itself encourages S to increase the fraction of the asset he sells to buyers.[15] Thus whether increasing β will lead to an increase in the optimal fraction f of the asset S sells to outsiders depends on which of the "liability effect" or the "price effect" is larger. The proof of part (c) shows that the price effect is the larger of these two effects, and so S optimally increases the fraction of the asset he sells as β increases.

6 Summary

We have studied a disclosure problem where the seller of an asset sometimes receives private information regarding an estimate of the asset's value prior to sale, which he may decide to disclose to, or withhold from, potential buyers of the asset. If the seller elects to withhold the estimate from buyers, he is liable for damage payments in the event his withholding is subsequently detected after the sale. These damage payments are taken to be SEC 10b-5 like in that they are proportional to the amount the buyer(s) of the asset overpaid for the asset (calculated based on what the asset would have sold for were the seller's estimate made public). The analysis shows that for a broad range of parameter values, the seller's optimal disclosure policy is the same whether the seller is myopic and chooses to disclose his estimate just based on whether the disclosure increases or decreases the asset's market value at the time of the disclosure, or whether the seller is nonmyopic and also takes into account the possible damage payments he may be liable for if his withholding is subsequently detected. A collection of comparative statics were obtained, some of which are unexpected. For example, if the damages multiplier determining what fraction of the buyer's overpayment

[15] I wish to thank Xu Jiang of Duke University for providing me with intuition for this last result.

must be reimbursed by the seller goes up, the analysis shows that it is often the case that the seller's propensity to withhold the information he receives increases.

In an extension of the base model where the seller of the asset can influence the distribution of the asset's value by making an investment, we show that the seller's optimal investment increases in each of: the precision of the estimate; the probability the seller receives the estimate, the damages multiplier, and - conditional on having withheld information - the probability that fact finder will detect that the seller withheld information.

When the asset is divisible, we showed that the fraction of the asset the seller optimally retains decreases in each of: the precision of the estimate the seller sometimes discloses, the probability the seller receives the information that he sometimes discloses, and the damages multiplier. When interpreted in the context of an entrepreneur who starts up a firm and then subsequently engages in an IPO, these last results lead to the testable predictions that an entrepreneur will retain a smaller fraction of the IPO for himself as the quality of the firm's accounting system increases, as the probability the entrepreneur receives the estimate increases, and as the damages multiplier increases.

7 Appendix: Proofs

The following lemma is instrumental in simplifying some of the expressions that arise in various proofs.

Lemma 12 *If $\tilde{u}$ is normally distributed with mean $m(I)$ and variance σ^2, density $g(u|I)$ and cdf $G(u|I)$, then:*

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

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

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

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

(v) $G_I(u^c|I) = -m'(I)g(u^c|I)$;

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

Proof of Lemma 12 Start by recalling the definitions: $g(u|I) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(u-m(I))^2}{2\sigma^2}}$ and $G(u^c|I) = \int_{-\infty}^{u^c} g(u|I)du$ and so

$$\frac{dg(u|I)}{du} = -\frac{u - m(I)}{\sigma^2} \times g(u|I).$$

Thus, for any u^c :

$$\int_{-\infty}^{u^c} (u - m(I))g(u|I)du = -\sigma^2 \times \int_{-\infty}^{u^c} (-\frac{u - m(I)}{\sigma^2})g(u|I)du = -\sigma^2 g(u^c|I). \tag{34}$$

so

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

This proves (iii). Next observe

$$g_I(u|I) = \frac{u - m(I)}{\sigma^2} \times m'(I) \times g(u|I),$$

and so,

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

(The second to last equality follows from (34).) This proves (v).

We now recall the notation as in the text: $\phi(x)$ and $\Phi(x)$ denote the density and cdf of a standard normal random variable. It is a standard observation concerning expectations of 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 $\phi(\frac{u^c - m(I)}{\sigma}) = \sigma g(u^c|I)$ and $1 - \Phi(\frac{u^c - m(I)}{\sigma}) = 1 - G(u^c|I)$:

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

This proves (iv). This also shows that

$$
\begin{aligned}
\int_{-\infty}^{\infty} \max\{u^c, u\} g(u|I)du &= \Pr(\tilde{u} \leq u^c)u^c + \int_{u^c} ug(u|I)du \\
&= G(u^c|I)u^c + m(I)(1 - G(u^c|I)) + \sigma^2 g(u^c|I) \\
&= G(u^c|I)(u^c - m(I)) + m(I) + \sigma^2 g(u^c|I).
\end{aligned}
$$

This proves (ii). This further implies:

$$
\begin{aligned}
\int^{u^c} (u^c - u)g(u|I)du &= \int_{-\infty}^{\infty} \max\{u^c - u, 0\} g(u|I)du \\
& \quad \int_{-\infty}^{\infty} \max\{u^c - u, u - u\} g(u|I)du \\
&= \int_{-\infty}^{\infty} (\max\{u^c, u\} - u) g(u|I)du \\
&= \int_{-\infty}^{\infty} \max\{u^c, u\} g(u|I)du - \int_{-\infty}^{\infty} ug(u|I)du \\
&= \int_{-\infty}^{\infty} \max\{u^c, u\} g(u|I)du - m(I) \\
&= G(u^c|I)(u^c - m(I)) + \sigma^2 g(u^c|I).
\end{aligned}
$$

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

$$\begin{aligned}
\int_{-\infty}^{\infty} \max\{u^c, u\} g_I(u|I) du &= \frac{\partial}{\partial I} \int_{-\infty}^{\infty} \max\{u^c, u\} g(u|I) du \\
&= \frac{\partial}{\partial I} \left(G(u^c|I)(u^c - m(I)) + m(I) + \sigma^2 g(u^c|I) \right) \text{ (from (ii))} \\
&= \frac{\partial}{\partial I} \left(G(u^c|I)(u^c - m(I)) + m(I) + \sigma \phi(\frac{u^c - m(I)}{\sigma}) \right) \\
&= G_I(u^c|I)(u^c - m(I)) - m'(I) G(u^c|I) + m'(I) + \sigma \frac{\partial}{\partial I} \phi(\frac{u^c - m(I)}{\sigma}) \\
&= G_I(u^c|I)(u^c - m(I)) - m'(I) G(u^c|I) \\
&\quad + m'(I) + m'(I) \frac{u^c - m(I)}{\sigma^2} \sigma \phi(\frac{u^c - m(I)}{\sigma}) \\
&= G_I(u^c|I)(u^c - m(I)) - m'(I) G(u^c|I) + m'(I) \\
&\quad + m'(I)(u^c - m(I)) g(u^c|I) \\
&= -m'(I) g(u^c|I)(u^c - m(I)) - m'(I) G(u^c|I) + m'(I) \\
&\quad + m'(I)(u^c - m(I)) g(u^c|I) \text{ (from (v))} \\
&= m'(I)(1 - G(u^c|I)).
\end{aligned}$$

This proves (vi).■

Proof of Theorem 2 Start by observing that (21) can be rewritten using the definition of conditional expectation and the obvious fact that $m = \frac{\tau m + r m}{\tau + r}$ as:

$$\frac{\tau m + r v^c}{\tau + r} = \frac{(1-p)\frac{\tau m + r m}{\tau + r} + p \int^{v^c} \frac{\tau m + r v}{\tau + r} g(v) dv + p \frac{r q \beta}{\tau + r} \int^{v^c} (v^c - v) g(v) dv}{1 - p + p G(v^c)},$$

or alternatively as:

$$\frac{\tau m + r v^c}{\tau + r} = \frac{(1 - p + pG(v^c)) \frac{\tau m}{\tau + r} + \frac{r}{\tau + r} \left((1-p)m + p \int^{v^c} v g(v) dv + p q \beta \int^{v^c} (v^c - v) g(v) dv \right)}{1 - p + p G(v^c)}. \tag{35}$$

Now, multiply both sides of equation (35) by the denominator of RHS(35) to get:

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

or equivalently,

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

or equivalently

$$(v^c - m)(1-p) + pv^c G(v^c) = p\int^{v^c} vg(v)dv + pq\beta \int^{v^c} (v^c - v)g(v)dv. \quad (36)$$

From Lemma 12 parts (iii) and (i) we know:

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

and

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

so (36) can be written as:

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

i.e., as:

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

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

$$(\frac{v^c - m}{\sigma})(1 - p + \alpha p G(v^c)) + \alpha p \sigma g(v^c) = 0. \quad (37)$$

Now define x^c by

$$x^c \equiv \frac{v^c - m}{\sigma}, \quad (38)$$

and substitute this x^c into (37), after observing that $G(v^c) = \Phi(x^c)$ and $\sigma g(v^c) = \phi(x^c)$, to conclude that (37) can be rewritten as:

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

This last equation is what was labeled (23) in the statement of the theorem.

Next notice that since - throughout - we have restricted attention to those β for which both $\beta \geq -1$ and $q(1+\beta) < 1$, we know that $\alpha = 1 - q\beta$ is positive. (This follows since α can be written as $\alpha = (1 - q(1+\beta)) + q$, and so α is positive as the sum of two positive terms.) We also note that α is bigger than 1 when β is negative, and α is smaller than 1 when β is positive (the preceding is clear since $\alpha = 1 - q\beta$).

Since α is always positive, we know 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 so

$$\frac{\partial LHS(23)}{\partial x} = 1 - p + \alpha p \Phi(x) + x\alpha p\phi(x) - \alpha p\phi(x)x = 1 - p + \alpha p\Phi(x) > 0. \quad (39)$$

Thus, from (39) we know that if equation (23) has a solution, that solution is unique. 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, by the intermediate value theorem, (23) has a negative solution. This proves the theorem.

■

Proof of Corollary 5

Part (i) follows from the observations made in the course of proving that (23) has a unique negative solution.

Part (ii) follows part (i) since if $x^c < 0$ implies $\Phi(x^c) < .5$, and so $p(1 - \Phi(x^c)) > \frac{p}{2}$.

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

$$\frac{\partial LHS(23)}{\partial x}\frac{\partial x^c}{\partial p} + \frac{\partial LHS(23)}{\partial p} = 0, \quad (40)$$

or equivalently, using (39):

$$[1 - p + \alpha p\Phi(x^c)]\frac{\partial x^c}{\partial p} = x^c(1 - \alpha\Phi(x^c)) - \alpha\phi(x^c). \quad (41)$$

Now, notice that when evaluated at its solution x^c, (23) can be written as

$$x^c - p\{x^c(1 - \alpha\Phi(x^c)) - \alpha\phi(x^c)\} = 0,$$

so

$$x^c(1 - \alpha\Phi(x^c)) - \alpha\phi(x^c) = \frac{x^c}{p} < 0.$$

Hence, RHS(41) is negative. Thus, $\frac{\partial x^c}{\partial p} < 0$.

We now prove $\frac{\partial x^c}{\partial \alpha} < 0$. (Once we do this, then parts (iv) and (v) will follow directly, since $\alpha = 1 - q\beta$, and hence $sgn\frac{\partial x^c}{\partial \beta} = sgn\frac{\partial x^c}{\partial \alpha}\frac{\partial \alpha}{\partial \beta} = -sgn\frac{\partial x^c}{\partial \alpha}q = -sgn\frac{\partial x^c}{\partial \alpha}$; similarly $sgn\frac{\partial x^c}{\partial q} = sgn\frac{\partial x^c}{\partial \alpha}\frac{\partial \alpha}{\partial q} = -sgn\frac{\partial x^c}{\partial \alpha}\beta$, and so $sgn\frac{\partial x^c}{\partial q}$ is positive if $\beta > 0$ and $sgn\frac{\partial x^c}{\partial q}$ is negative if $\beta < 0$.) Differentiate (23) totally with respect to α, in a fashion analogous to (40) 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:

$$\frac{\partial x^c}{\partial \alpha} = -p\frac{x^c\Phi(x^c) + \phi(x^c)}{1 - p + \alpha p\Phi(x^c)}. \tag{42}$$

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

$$f(x) \equiv x\Phi(x) + \phi(x) \tag{43}$$

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

$$\lim_{x\to-\infty} x\Phi(x) = \lim_{x\to-\infty}\frac{\Phi(x)}{\frac{1}{x}} = \lim_{x\to-\infty}\frac{\phi(x)}{-\frac{1}{x^2}} = -\frac{1}{\sqrt{2\pi}}\lim_{x\to-\infty} x^2 e^{-x^2/2} = 0. \tag{44}$$

Since $\lim_{x\to-\infty}\phi(x) = 0$ too, it follows from (44) 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 (43) that we can write $\frac{\partial x^c}{\partial \alpha}$ in (42) 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 $\frac{\partial x^c}{\partial \alpha} < 0$. ■

Proof of Theorem 7

Initially, we take $\hat{I}$ as fixed, and so, to save space, we write v^c in place of $v^c(\hat{I})$.

Rewrite $OBJ(I|\hat{I})$ as:

$$\begin{aligned} OBJ(I|\hat{I}) &= (1-p+pG(v^c|I))\frac{\tau\hat{m}+rv^c}{\tau+r} + p\int_{v^c}^{\infty}\frac{\tau\hat{m}+rv}{\tau+r}g(v|I)dv \\ &\quad -pq\beta\int_{-\infty}^{v^c}(\frac{\tau\hat{m}+rv^c}{\tau+r} - \frac{\tau\hat{m}+rv}{\tau+r})g(v|I)dv - .5I^2 \\ &= (1-p)\frac{\tau\hat{m}+rv^c}{\tau+r} + pG(v^c|I)\frac{\tau\hat{m}+rv^c}{\tau+r} + p\int_{v^c}^{\infty}\frac{\tau\hat{m}+rv}{\tau+r}g(v|I)dv \\ &\quad -pq\beta\int_{-\infty}^{v^c}(\frac{\tau\hat{m}+rv^c}{\tau+r} - \frac{\tau\hat{m}+rv}{\tau+r})g(v|I)dv - .5I^2. \end{aligned} \tag{45}$$

Now observe that

$$\begin{aligned} \int_{-\infty}^{v^c}(\frac{\tau\hat{m}+rv^c}{\tau+r} - \frac{\tau\hat{m}+rv}{\tau+r})g(v|I)dv &= \int_{-\infty}^{\infty}\max\{\frac{\tau\hat{m}+rv^c}{\tau+r} - \frac{\tau\hat{m}+rv}{\tau+r}, 0\}g(v|I)dv \\ &= \int_{-\infty}^{\infty}\max\{\frac{\tau\hat{m}+rv^c}{\tau+r} - \frac{\tau\hat{m}+rv}{\tau+r}, \frac{\tau\hat{m}+rv}{\tau+r} - \frac{\tau\hat{m}+rv}{\tau+r}\}g(v|I)dv \\ &= \int_{-\infty}^{\infty}\left(\max\{\frac{\tau\hat{m}+rv^c}{\tau+r}, \frac{\tau\hat{m}+rv}{\tau+r}\} - \frac{\tau\hat{m}+rv}{\tau+r}\right)g(v|I)dv \\ &= \int_{-\infty}^{\infty}\max\{\frac{\tau\hat{m}+rv^c}{\tau+r}, \frac{\tau\hat{m}+rv}{\tau+r}\}g(v|I)dv - \int_{-\infty}^{\infty}\frac{\tau\hat{m}+rv}{\tau+r}g(v|I)dv \\ &= \int_{-\infty}^{\infty}\max\{\frac{\tau\hat{m}+rv^c}{\tau+r}, \frac{\tau\hat{m}+rv}{\tau+r}\}g(v|I)dv - \frac{\tau\hat{m}+rm}{\tau+r}. \end{aligned} \tag{46}$$

(Notice that the last equality invokes $\int_{-\infty}^{\infty}\frac{\tau\hat{m}+rv}{\tau+r}g(v|I)dv = \frac{\tau\hat{m}+rm}{\tau+r}$, where $m = m(I)$.) Also notice that

$$\begin{aligned} \frac{\tau\hat{m}+rv^c}{\tau+r}G(v^c|I) + \int_{v^c}^{\infty}\frac{\tau\hat{m}+rv}{\tau+r}g(v|I)dv &= \int_{-\infty}^{v^c}\frac{\tau\hat{m}+rv^c}{\tau+r}g(v|I)dv + \int_{v^c}^{\infty}\frac{\tau\hat{m}+rv}{\tau+r}g(v|I)dv \\ &= \int_{-\infty}^{\infty}\max\{\frac{\tau\hat{m}+rv^c}{\tau+r}, \frac{\tau\hat{m}+rv}{\tau+r}\}g(v|I)dv. \end{aligned} \tag{47}$$

Substitute (46) and (47) into (45) to get

$$\begin{aligned} OBJ(I|\hat{I}) &= (1-p)\frac{\tau\hat{m}+rv^c(\hat{I})}{\tau+r} + p\int_{-\infty}^{\infty}\max\{\frac{\tau\hat{m}+rv^c}{\tau+r},\frac{\tau\hat{m}+rv}{\tau+r}\}g(v|I)dv \\ &\quad -pq\beta\left(\int_{-\infty}^{\infty}\max\{\frac{\tau\hat{m}+rv^c}{\tau+r},\frac{\tau\hat{m}+rv}{\tau+r}\}g(v|I)dv - \frac{\tau\hat{m}+rm}{\tau+r}\right) - .5I^2 \\ &= (1-p)\frac{\tau\hat{m}+rv^c(\hat{I})}{\tau+r} + p(1-q\beta)\int_{-\infty}^{\infty}\max\{\frac{\tau\hat{m}+rv^c}{\tau+r},\frac{\tau\hat{m}+rv}{\tau+r}\}g(v|I)dv \\ &\quad +pq\beta\frac{\tau\hat{m}+rm}{\tau+r} - .5I^2. \end{aligned} \tag{48}$$

Differentiate this last expression with respect to I to get:

$$\frac{\partial OBJ(I|\hat{I})}{\partial I} = p(1-q\beta)\int_{-\infty}^{\infty}\max\{\frac{\tau\hat{m}+rv^c(\hat{I})}{\tau+r},\frac{\tau\hat{m}+rv}{\tau+r}\}g_I(v|I)dv + pq\beta\frac{rm'(I)}{\tau+r} - I.$$

Recall from Lemma 12(vi) that: $\int_{-\infty}^{\infty}\max\{v^c,v\}g_I(v|I)dv = m'(I)(1-G(v^c|I))$. Combined with $\int_{-\infty}^{\infty} g_I(v|I)dv = 0$, it follows that:

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

and hence recalling $\alpha = 1-q\beta$ it further follows that

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

The second-to-last equality follows from Theorem 2. The equilibrium I is obtained by setting this last expression equal to zero, which proves the theorem.

■

Proof of Theorem 10

The first part of the proof entails establishing when OBJ, as expressed in (**??**), is evaluated at the equilibrium value $I^*(f)$ specified in Theorem 9, that

$$\frac{d}{df}OBJ = (1-\delta)\delta w^2 + (2-\delta-X)\, w^2 f(X-\delta). \tag{49}$$

This will take several steps to oprove. First step. Compute this derivative by applying the envelope theorem to it. Note that while the envelope theorem allows us to ignore the effect of changes in f onS's equilibrium choice $I^*(f)$, the envelope theorem does not allow us to disregard the effect of changes in f on buyers' conjecture$\hat{I}(f)$, since S controls (and can adjust) $I^*(f)$ as he f changes, but he does not control how investors' conjectures $\hat{I}(f)$ change with f. When we evaluate this derivative at the equilibrium investment level $I^*(f) = \hat{I}(f)$, we get:

$$\frac{\partial}{\partial f}OBJ = -\delta m(I^*(f)) + \Psi + f \times \frac{\partial \Psi}{\partial \hat{I}}|_{I=\hat{I}=I^*(f)} \times \frac{\partial \hat{I}(f)}{\partial f}. \tag{50}$$

(Each instance of Ψ in this derivative is evaluated at $I = \hat{I} = I^*(f)$.)

$$OBJ \equiv \delta(1-f)m(I) + f \times \Psi(I,\hat{I}) - .5I^2, \tag{51}$$

where

$$\begin{aligned}\Psi(I,\hat{I}) \quad &\equiv \quad (1-p+pG(v^c(\hat{I})|I)) \times \frac{\tau m(\hat{I}) + rv^c(\hat{I})}{\tau + r} + \\ &p\int_{v^c(\hat{I})}^{\infty} \frac{\tau m(\hat{I}) + rv}{\tau + r} g(v|I)dv - pq\beta \int_{-\infty}^{v^c(\hat{I})} \left(\frac{\tau m(\hat{I}) + rv^c(\hat{I})}{\tau + r} - \frac{\tau m(\hat{I}) + rv}{\tau + r}\right) g(v|I)dv.\end{aligned} \tag{52}$$

Second step. Simplify this derivative. To that end, notice that

$$f \times \Psi(I^*(f), I^*(f)) = f \times m(I^*(f)) = fwI^*(f). \tag{53}$$

That is, the net proceeds S expects to receive from buyers for selling them fraction f of the asset, $f \times \Psi(I,I)$, equals fraction f of the asset's total value $m(I)$. This result follows because in expectation the damage payments constitute a net wash to S in equilibrium: competition among buyers drives the price they pay for the fraction of the asset they buy up to the total value they anticipate receiving from their purchase. So, in expectation, whatever amount buyers expect S to pay them in damages after the asset's sale equals the extra amount the buyers offer S "up front" at the time of the asset's sale. (This claim can be derived formally following (35), but we omit its simple proof.)

Next, calculate $\frac{\partial}{\partial \hat{I}}\Psi(I,\hat{I})|_{I=\hat{I}=I^*(f)}$. Following the same steps that allowed us to write OBJ in (27) as (48), we can write $\Psi(I,\hat{I})$ as

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

Hence:

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

In this last line, we utilized (22) (with $\hat{m}$ replacing m) to conclude $\frac{\partial v^c}{\partial \hat{m}}=1$, and hence $\frac{\partial v^c}{\partial \hat{I}}=\frac{\partial v^c}{\partial \hat{m}}\frac{\partial \hat{m}}{\partial \hat{I}}=m'(\hat{I})=w$. Hence, when $\hat{I}=I:$

$$\begin{aligned}
\frac{\partial}{\partial \hat{I}}\Psi(I,I) &= (\{1-p+p(1-q\beta)G(v^c(I)|I)\}\frac{r}{\tau+r}+\frac{\tau}{\tau+r})w\\
&= (\frac{r}{\tau+r}-\frac{pr}{\tau+r}(1-(1-q\beta)G(v^c(I)|I))+\frac{\tau}{\tau+r})w\\
&= (1-\frac{pr}{\tau+r}(1-\alpha G(v^c(I)|I)))w\\
&= (1-\frac{pr}{\tau+r}(1-\alpha\Phi(x^c)))w\\
&= (1-X)w. \qquad (54)
\end{aligned}$$

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

Next, we observe that, expressed in the notation (28), the result in (32) can be stated as

$$I^*(f) = \delta w + fw \times (X - \delta), \tag{55}$$

and so

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

and we can write $I^*(f)$ alternatively as:

$$I^*(f) = \delta w + f \times \frac{\partial I^*(f)}{\partial f}. \tag{56}$$

Since f is public information, $I^*(f) \equiv \hat{I}(f)$ is an identity in f, so:

$$\frac{\partial I^*(f)}{\partial f} = \frac{\partial \hat{I}(f)}{\partial f}. \tag{57}$$

Thus, putting (54) to (57) together, we can express the derivative (50) as:

$$\begin{aligned}
\frac{\partial}{\partial f} OBJ &= -\delta m(I^*(f)) + \Psi + f \times \frac{\partial \Psi}{\partial \hat{I}}|_{I=\hat{I}=I^*(f)} \times \frac{\partial \hat{I}(f)}{\partial f} \\
&= (1-\delta) w I^*(f) + f \times \frac{\partial \Psi}{\partial \hat{I}}|_{I=\hat{I}=I^*(f)} \times \frac{\partial \hat{I}(f)}{\partial f} \\
&= (1-\delta) w (\delta w + f \times \frac{\partial I^*(f)}{\partial f}) + f \times (1-X) w \times \frac{\partial I^*(f)}{\partial f} \\
&= (1-\delta)\delta w^2 + (1 - \delta + 1 - X)\, w f \frac{\partial I^*(f)}{\partial f} \\
&= (1-\delta)\delta w^2 + (2 - \delta - X)\, w^2 f (X - \delta).
\end{aligned} \tag{58}$$

Note that this last line is (49), so the first part of proof of Theorem 10 is complete.

The next part of the proof entails establishing the following lemma.

Lemma 13 *Case 1 If $\delta \leq X$, then the optimal f is $f = 1$*

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

$$f^* = \frac{(1-\delta)\delta}{(1-X)^2 - (1-\delta)^2}. \tag{59}$$

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

Proof of Lemma 13

First, we note that since

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

(58) 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 (59) is positive. In this case, this ratio is less than 1 iff

$$\begin{aligned}(1-\delta)\delta &< (1-X)^2-(1-\delta)^2 \text{ iff} \\ (1-\delta)^2+\delta-\delta^2 &< (1-X)^2 \\ 1-\delta &< (1-X)^2. \qquad (60)\end{aligned}$$

When inequality (60) holds - Case 2(a) - $\frac{\partial}{\partial f}OBJ$ is positive for $f < f^*$, where f^* is as given in (59), and $\frac{\partial}{\partial f}OBJ$ is negative for $f > f^*$. Hence, in Case 2(a), the f^* as defined in (59) is the global maximum of OBJ defined in (**??**). When inequality (60) is reversed - Case 2(b) - $\frac{\partial}{\partial f}OBJ$ is positive for all $f < f^*$ where f^* is as defined in (59). In particular, since in Case 2(b) the f^* defined in (59) 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.■

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

We now prove Case B (in the statement of Theorem 10). 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 (59) for all $r > 0$ in Case B.

In Case A (in the statement of Theorem 10), 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 (60) 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 10.■

Proof of Corollary 11

(i) First, suppose Case A characterizes f^*, i.e., $(1 - p(1 - \alpha\Phi(x^c)))^2 < 1 - \delta$. Further suppose $r < r^\delta$, i.e., Case Ai applies. That is, suppose f^* is defined by

(33). For such r, observe that as r increases:

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

This shows that an increase in r leads to an increase in f^* in Case Ai. Also notice that if, by increasing r, the relevant case goes from Case Ai to Case Aii, it remains true that f^* increases, since as r increases from being below r^δ to being above r^δ, f^* goes from some amount less than 1 to 1, So the corollary remains true when the case that characterizes f^* goes from Case Ai to Case Aii. If Case B applies, then f^* is also defined by (33), and so, by what has been shown previously, we are done.

(ii) The essential step of this demonstration is to show that (33) 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 10 decreases in p. Then, similar to the steps taken in the proof of part (i) above, it is easy to show that (33) is strictly increasing in r. The remainder of the proof (where we consider whether Case Ai or Case Aii or Case B applies) is straightforward and omitted.

(iii) The essential step is to show that as β increases, f^* as defined in (33) 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 Theorem 10. 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 (42) 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 (43) 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.

Finally, 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 (33) 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 (33)), so it is always true that, as β increases, the optimal f always weakly increases.■

8 References

Demarzo, P. and D. Duffie, "A Liquidity-based Model of Security Design, *Econometrica* Volume 67, Issue 1, pages 65–99, January 1999.

Dranove, David and Ginger Jin, "Quality disclosure and certification: Theory and Practice," *Journal of Economic Literature*. 2010; 48(4):935-963.

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

Farrell, Joseph, "Voluntary Disclosure: Robustness of the Unraveling Result," *in Antitrust and Regulation*, R. Grieson, ed., Lexington Books (1986), 91-103

Gertner, R., 1998. Disclosures and Unraveling. In: Peter, N. (Ed.), The New Palgrave Dictionary of Economics and the Law. Macmillan Press, London.

Grossman, S. "The Informational Role of Warranties and Private Information about Product Quality," Journal of Law and Economics Vol. 24, No. 3,1981 pp. 461-483.

Healy, Paul and Krishna Palepu, "Information Asymmetry, Corporate Disclosure and the Capital Markets: A Review of the Empirical Disclosure Literature," Journal of Accounting and Economics Vol. 31, Nos. 1-3, September 2001

Heitzman, Shane, Charles Wasley, and Jerold Zimmerman, "The Joint Effects of Materiality Thresholds and Voluntary Disclosure Incentives on Firms' Disclosure Decisions," *Journal of Accounting and Economics* Volume 49, Issues 1–2, February 2010, Pages 109–132.

Hughes, John and Suil Pae, "Voluntary Disclosure of Precision Information," Journal of Accounting and Economics 37 pp. 261-289, 2004

Hughes, John and Suil Pae, "Discretionary Disclosure and Efficiency of Entrepreneurial Investment," *Contemporary Accounting Research* forthcoming.

Jin, G. and P. Leslie (2003): "The Effectof Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," Quarterly Journal of Economics, 118(2), 409—51.

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.

Milgrom, P., "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics* Vol. 12, No. 2. (Autumn, 1981), pp. 380-391.

______ "What the Seller Won't Tell You: Persuasion and Disclosure in Markets." Journal of Economic Perspectives 2008, 22(2): 115-131.

Pae, S. "Selective Disclosures in the Presence of Uncertainty about Information Endowment," *Journal of Accounting and Economics*, Vol. 39, No. 3 (September 2005): 383-409.

Pae, Suil, "Discretionary Disclosure, Efficiency, and Signal Informativeness, " *Journal of Accounting and Economics* 33 pp. 279-311, 2002.

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

Shavell, Steven, "Acquisition and Disclosure of Information Prior to Sale," *RAND Journal of Economics*, Vol. 25, No. 1 (Spring, 1994), pp. 20-36