Why managers with low forecast precision select high disclosure intensity: an equilibrium analysis

Miles B. Gietzmann
Faculty of Finance
Cass Business School

Adam J. Ostaszewski
Department of Mathematics
London School of Economics

This version*January 2012

Abstract

Shin (2006) has argued that in order to understand the equilibrium patterns of corporate disclosure, it is necessary for researchers to work within an asset pricing framework in which corporate disclosures are endogenously determined. Echoing this sentiment, Larcker and Rusticus (2010) have argued that earlier empirical results claiming to find a *negative* relationship between disclosure and cost of capital may suffer fatally from endogeneity issues which, once addressed by a formal structural model, may reverse the sign of the relationship. The purpose of this paper is to introduce a general equilibrium model following the Black-Scholes paradigm with endogeneous disclosure in which firms select uniquely determined optimal probabilities of early equity-value discovery in a noisy environment. As firms may differ also in the uncertainty (precision) with which management can forecast the future, managers strategically increase the intensity of their (voluntary) disclosures to provide partial compensation for this perceived differential risk. A *positive* relationship then results between disclosure and the cost of capital.

*Preliminary version presented at Bachelier Society 5th World Conference, London (2008).

1 Introduction

An issue with studying research on corporate disclosure is that theory and empirical work are not closely related. Empirical work often proceeds under the a priori assumption that disclosure and cost of capital are *negatively* related. However historically theoretical models often focus upon 'hard to observe' (latent) variables that are difficult to implement empirically and so do not always give clear guidance to empirical work. Moreover, some most recent empirical and theoretical work finds the opposite to hold - a *positive* relationship.

In an empirical paper van Buskirk (2011) finds that "the practice of regularly providing monthly revenue disclosures is not associated with reduced information asymmetry. In contrast, I find that more detailed (greater quantity) disclosure is associated with reduced information asymmetry." He argues that this arises because differences in disclosure frequency between firms alter the incentives for sophisticated investors to collect more private information. In line with Shin's (2006) original comments (see abstract), Clinch and Verrecchia (2011) develop a theoretical structural model allowing an endogenized cost of capital calculation (taking as exogenous investor numbers, their common utility, their common belief about firm cash-flows, and the prospect of early managerial discovery of true cash-flow). Mirroring an early result with a different model setup, due to Gietzmann and Trombetta (2003), their theory predicts a *positive* relationship. We learn consistently from theory models that firms respond strategically to their environment. Thus firms subject to downgrade by investors, on account of their highly uncertain prospects, may find it rational to improve their disclosures to reduce uncertainty gaps, although this need not result in parity with other firms from the same sector that have higher precision (see Section 3.2).

An issue with the above model of Clinch and Verrecchia (2011) is that they do not explain how firms differ endogenously with respect to managerial discovery of true cash-flow. The principal contribution of this paper is to provide a model of disclosure in which management of firms endogenously choose the discovery probability, which influences the disclosure. We amend the Dye (1985) disclosure model; until now that framework has not been readily amenable to empirical study. One of the difficulties is that an underlying parameter (probability of information discovery, which is referred to here as the *information endowment parameter*, or briefly *endowment parameter*) is a latent variable that may vary between companies. Also, that framework has usually been applied to returns rather than equity value; working with arbitrage-free equity valuations (log-normal as in the Black-Scholes model[1]), rather than cash-flow valuations, we here address the influence of equity-volatility and find that firms with least precise forecasts of one-period-ahead equity value will adopt the highest disclosure intensity and face the highest cost of capital (see Section 2.4). Additional disclosure may be used strategically to compensate, but only *partially,* for this sort of risk (see the discussion in Section 3.2); this is consistent with the empirical findings of van Buskirk (2011). Indeed, Larcker and Rusticus (2010) have argued earlier, empirical results that claim to find a *negative* relationship between these two variables may suffer fatally from endogeneity issues which, once dealt with by a formal structural model, may reverse the sign of the relationship. That is, if one were formally to model the cost of capital with endogenized strategic disclosure, the relationship between cost of capital and disclosure might become *positive.*

The purpose of this paper is to introduce a structural model in which firms differ in equity volatility and in the uncertainty (precision) with which management can forecast the future, given their private opportunity of value discovery. After developing the model, we are able to point to several measures of disclosure and to trace their relation to the cost of capital. For example, we state formal conditions under which a strong monotonicity result holds in which disclosure when measured by intensity (frequency), see definition below in equation (11), is *positively* related to managerial forecast variance; furthermore, the relationship between cost-of-capital and an alternative measure of disclosure (range of undisclosed values) is found to be *positively* related to the cost of capital, in the sense that the cost of capital rises with the range of values disclosed.

Section 2 develops the equilibrium framework. Within this framework, it becomes possible (see Section 3) to show how the optimal disclosure strategy of a firm implies an observable disclosure intensity, and how that in turn can be used to form inferences about the underlying parameters of the firm which determine the actual (rational) equilibrium market valuation. These findings come from the assumption that investor risk-preferences may be analyzed by reference to a risk-measure called the *omega-ratio* (well-known in portfolio

[1] We find that passing back and forth, via logarithms, between the additive arithmetic averaging of classical linear regression in respect of normal returns and its log-normal counterpart – a multiplicative geometric averaging of asset values – is straightforward and intuitive. The non-linearity of the logarithm turns out to be highly tractable.

theory – see Section 3.2).

Section 4 provides the principal mathematical formulas for log-normal models which permit tractability of the proposed approach; Section 5 sketches possible empirical research design. Concluding commentary is in Section 6. Most of the mathematical analysis is split into appropriate appendices.

2 The Dye Disclosure Calculus with Endogenous Information Endowment

This section reviews and builds on the Dye disclosure calculus, as established in Dye (1985) and developed in Jung and Kwon (1988). In (2.1) the calculus is established and also its underlying no-arbitrage foundations are explained (in terms of the associated 'risk-neutral' valuation measure); then in (2.2) endogenous selection of information endowment by management is introduced. It is shown how the choice-model of information endowment is consistent with established utility theory approaches in portfolio theory, in particular those where downside and upside risk are not viewed as having symmetric influences on preferences. This is key to further developing the Dye calculus, which is essentially driven by a lower partial moment computed over the range up to the Dye cutoff[2], as defined by equation (1). Section (2.3) links all the preceding sections to show how the model of disclosure strategy leads to a well-defined and tractable definition of disclosure intensity, which is used in Section (2.4) to derive the cost-of-capital discount from the relationship between the risk-neutral measure and the associated 'physical' measure.

2.1 No-Arbitrage Risk-Neutral Valuation with Dye's disclosure calculus

In the Dye model there is a rational (equilibrium) reason why management might not disclose information voluntarily (a relaxation of the unraveling paradigm). This necessitates a procedure (due to Dye, further developed below) enabling investors to value the company at other than the Grossman-and-Hart (1980) unconditional minimum (in which 'bad news' is assumed), when investors observe non-disclosure. We point out that Dye's disclosure-cutoff should be viewed as yielding a valuation based on the methodology of

[2]Note that the equilibrium cutoff is below the opening expected value.

arbitrage-free pricing. That is, it is determined by a *no-arbitrage condition*, which has implications for the valuation of the embedded 'disclosure' put option, to be discussed below.

When analyzing information flows, the Dye disclosure model assumes three distinctive time $\theta = 0, 1, 2$: ex-ante, interim and terminal times. In the model a random variable X, relating to company valuation (see below for a comprehensive framework for this), has density $f(x)$, an associated distribution function $F_X(x)$ and an ex-ante (i.e. at time $\theta = 0$) expected value m_X. A realization of the random variable is observed by management at the interim time with a probability q (drawn independently of X). Management's decision whether or not to disclose an observed realization of company value x is a voluntary (strategic) decision. Dye (1985) establishes that under continuity and positivity of f there exists a unique threshold value $t = \gamma$ at which management will be indifferent between disclosure or non-disclosure. Here γ will be called the *Dye cutoff*, defined implicitly by equation (1).

It is important to note that the Dye model assumes that the fixed parameter q is known to market participants. We formulate a more general framework in which the value of q is both deduce-able and statistically inferable. See the discussion on 'first best' towards the end of Section 2.3.

The indifference point is characterized by equality between a credibly disclosed value γ and the valuation formed by investors when they face non-disclosure (ND); the latter is formally $\mathbb{E}[X|ND(\gamma)]$, the computed expected value of the company, conditioned on the absence of information (non-disclosure) of values observed below γ. The latter expression is a consequence of Dye's assumption that "investors cannot discern whether [the manager] has received information but chosen not to release it or whether the manager has not received information" (Dye 1985, §3). That is, the indifference is described by what we term the *Dye equation,* or the *equilibrium indifference* equation:

$$\gamma = \mathbb{E}[X|ND(\gamma)]. \tag{1}$$

Under the assumptions above, this implicit definition of a cutoff value γ in fact determines it uniquely; whenever context demands, we emphasize the cutoff's dependence on the underlying information endowment parameter q of the manager and the specific model X of firm value by using the notation $\gamma(q)$ or $\gamma_X(q)$. Later, when referring to a family of distributions $F(x, \sigma)$ parametrized by σ a more convenient variant of this notation will suggest itself.

The equation expresses the fact that in the absence of a disclosure the market downgrades the value of the firm from m_X to γ_X. That is, the initial market value, is reduced to reflect the perceived new risks, and so $m_X - \gamma_X$ should be interpreted as a risk-premium, in short it is the *ex-post cost of disclosure*.

Note that Dye's approach (see his §3) assumes that "the current shareholders prefer a disclosure policy which maximizes the first-period [$\theta = 1$] price of the firm" and that "this disclosure policy is adopted".

Based on the assumption of a rational expectations equilibrium (in respect of a conjectural threshold value for the manager's cutoff), Jung and Kwon (1988) derive (their equation (7)) the equation satisfied by γ to be

$$\frac{1-q}{q}(m_X - \gamma) = H_X(\gamma), \tag{2}$$

where

$$H_X(t) := \int_{x \le t} F_X(x)dx = \int_{x \le t} (t-x)dF_X(x) = \mathbb{E}[(t-X)^+]. \tag{3}$$

$H_X(t)$ is the lower first partial moment below t, well-known in risk management[3], and here, working in present value terms (equivalently, with riskless rate $r = 0$), is seen to be the value of a put option with strike t (to which the Black-Scholes formula can be applied in the log-normal context). As this function is central to the Dye calculus, in our analysis we explicitly name it the *hemi-mean function*[4].

The appeal of this form lies in the separation of the two independent factors of the model: the information *odds*, i.e. the ratio $(1-q)/q$ to be henceforth denoted by λ, which characterizes management information technology on one side, and on the other a convex function H_X containing all the

[3] See for example McNeil, Frey and Embrechts (2005), Section 2.2.4.

[4] $H_X(t)$ is strictly convex, positive and asymptotic to $t - m_X$ (by l'Hôpital's Rule); clearly $H_X(\underline{X}) = 0$ and $H'_X(\underline{X}) = F_X(\underline{X}) = 0$, where $\underline{X}$ is the lower boundary of the support of F_X (possibly $-\infty$, as for the normal, if that is admissible).

$$\lim_{t \to +\infty} \frac{\int_{x \le \gamma} (t-x)dF(x)}{t - m_X} = \lim_{t \to +\infty} \frac{\int_{x \le t} dF(x)}{1} = 1.$$

information[5] on the distribution of X. The equation (2), when rewritten as

$$(1-q)(m_X-\gamma)=qH_X(\gamma), \tag{4}$$

shows the expected downgrade $(1-q)(m_X-\gamma)$, i.e. downgrade conditional on the manager receiving no information and so impounding the ex-post cost of capital; this balances what must be interpreted as some form of 'expected upgrade', conditional on the manager discovering information. Indeed some integration by parts shows why $qH_X(\gamma)$ is the essential term in the expected upgrade given by the upper partial first moment above γ.[6]

Given that the Dye cutoff is about an optimal downgrade-threshold, it is not surprising that there is a direct link between the cutoff calculus and the portfolio-theory of target-returns thresholds (in respect of shortfalls below the mean). A key construct there is the Omega ratio defined by:

$$\Omega(t,\sigma)=\frac{\int_{x\geq t}(1-F(x,\sigma))dx}{\int_{x\leq t}F(x,\sigma)dx},$$

introduced by W. Shadwick and C. Keating (2002) as a performance measure for a family of assets parametrized by σ on the threshold-variance pairs (t,σ), as a refinement of the usual mean-variance pairs of the standard the Markowitz analysis; they view it as an aid to selecting σ and "the level of return[7] against which a given outcome will be viewed as a gain or loss". In view of the *reciprocal identity* linking Ω with the hemi-mean function:

$$\frac{m_X-\gamma}{H_X(\gamma)}=\Omega_X(\gamma)-1, \tag{5}$$

---

[5] See Ostaszewski and Gietzmann (2008). **Proposition.** *Let $H(t)$ be any twice differentiable, strictly convex function on $[\underline{X},\overline{X}]$ satisfying $H(\underline{X})=0$, $H'(\underline{X})=0$ and $H'(\underline{X})=m$. Then $H(t)$ is the hemi-mean function of a continuous distribution with mean given by*

$$m=\overline{X}-H(\overline{X}).$$

This need reinterpretation when either limit of the support interval is infinite. For instance, when $\overline{X}=+\infty$,

$$m=\lim_{t\to\infty}(t-H(t)).$$

[6] $\int_{u\geq\gamma}(u-m_X)dF(u)=\int_{u\leq\gamma}(m_X-u)dF(u)=(m_X-\gamma)F(\gamma)+\int_{u\leq\gamma}F(u)du$

[7] Level of equity, in our case.

when $F_X(x) = F(x, \sigma_X)$ and $\Omega_X(x) = \Omega(x, \sigma_X)$, the Dye equation may be reformulated as

$$\begin{aligned} (1-q)\Omega(\gamma) &= 1, \text{ or} \\ \Omega - 1 &= \frac{q}{1-q} \end{aligned} \tag{6}$$

We return to this in Section 3.2.

A feature of the Dye equation, critical to later analysis, is its positive homogeneity[8], in the sense that

$$H_{\alpha X}(\alpha t) = \alpha H_X(t) \text{ for } \alpha > 0. \tag{7}$$

This simplifies intra-firm comparisons based on the Dye calculus; with m a given firm's mean $\mathbb{E}[X]$ and writing $\bar{\gamma}_j$ for γ_j/m_j, one may cancel by m on both sides of the Jung and Kwon equation, and so replace firms with different means by equivalent 're-normalized' firms with equal means.

The framework above can embrace an alternative interpretation: X may validly be replaced by a noisy signal of the true value X, say by $T = T(X, Y)$, where Y models noise. Then one may deduce (see Appendix 1) the existence of a cutoff γ_T above which the noisy signal T would in equilibrium be voluntarily disclosed. Given a disclosure, investors would then form expectations conditioning on the reported noisy signal, and the market values the firm at $\mathbb{E}[X|T]$ rather than at T. That is, referring to the *regression function* $\mu_X(t) := \mathbb{E}[X|T = t]$, the value is $\mu_X(T)$. If, however, no disclosure occurs, then the market valuation is $\mu_X(\gamma_T)$. The Dye disclosure calculus remains valid in a noisy setting, provided the X in Dye's model is interpreted, not as the true firm value, but as $\mathbb{E}[X|T]$, the estimated firm value given T. This is valid subject to $\mu_X(t)$ being an increasing function. All that needs doing in the Jung and Kwon equation is to replace H_X by another, related, hemi-mean function; the cutoff $\gamma_{\mathbb{E}[X|T]}$ defined implicitly in an amended Jung and Kwon equation then is $\mu_X(\gamma_T)$. (Here γ_T is, as above, the disclosure cutoff for the

[8] $\mathbb{E}[\alpha X|\alpha X < \alpha\gamma] = \alpha\mathbb{E}[X|X < \gamma]$ for $\alpha > 0$; or note that

$$H_{\alpha X}(\alpha\gamma) = \int_{\alpha x \le \alpha\gamma} \Pr(\alpha X \le \alpha x) d(\alpha x) = \alpha \int_{x \le \gamma} \Pr(X \le x) dx = \alpha H_X(t).$$

actual signal T.) For details see Appendix 1 (Isotonic transformations)[9].

A particular example we pursue is $T = XY$ with X, Y log-normal, in accord with the financial benchmark model. It is clear from this example that what may be termed the *sector variability,* as captured by variance in X and the *managerial variability* ,as captured by variance in Y, will both of them independently influence the cutoff, and so also the disclosure policy. We shall see at the end of Section 3.2 that they have opposite effects and we discuss the implications of these differing effects for empirical research design.

To clarify the no-arbitrage underpinnings of Dye's calculus, the market itself is modeled as *arbitrage-free* so equipped with a probability distribution function F_X which models events corresponding to the three times $\theta = 0$ and $\theta = 1$ and $\theta = 2$ in such a way that the distribution fully reflects the market price of risk at any of these points in time. That is, any contingent contract traded on the market is priced by computing an expectation of the claim under this distribution. This presumes the so-called complete market hypothesis and asserts that the distribution itself is an observable, i.e. there is a sufficient range of traded instruments to select a distribution from a proposed parametrized family, and so to identify the density of the risk-neutral measure and its distribution F. Thus at $\theta = 0$ the variance of X is known, and so too is $m_X := \mathbb{E}_{\theta=0}[X]$, the market share price of the firm at time $\theta = 0$. This is implicitly part of the Dye framework.

Dye assumes that q is known, and so any investor can compute the market value of the firm, denoted $V(t)$, when the manager is known to use an (arbitrary) cutoff t. That is, $V(t) = \mathbb{E}_{\theta=1}[X|ND(t)]$, where $ND(t)$ is the non-disclosure event at time $\theta = 1$ when the cutoff is t. Thus $m_X = \mathbb{E}_{\theta=0}[X \cdot 1_{D(t)} + V(t) \cdot 1_{ND(t)}]$. There is, however, a unique value $t = \gamma$ for which $V(t)$ assumes a minimum, and that value is characterized by the (unique) solution of the equation $t = V(t)$. That solution yields the minimum valuation consistent with the information available at $\theta = 1$, and coincides with the value at which the manager is indifferent between disclosing and not disclosing. It is this indifference-pricing approach that characterizes Dye's own justification for equilibrium. So this is also the unique value of t for which $m_X = \mathbb{E}_{\theta=0}[X \cdot 1_{D(t)} + t \cdot 1_{ND(t)}]$. That is, the time $\theta = 0$ valuation of the firm's

[9] Another way of extending the original framework: interpret X parsimoniously as incorporating real effects, e.g. observation of an input price in a managed production process with certain (short time-scale) implementation leading to a sure realizable (indirect) profit.

stock impounds the strategic disclosure option available to the manager.

It is well-known (see e.g. Jung and Kwon (1988)) that γ as a function of q is strictly monotone. So the equity-market model approach implies that, in the event that there is no disclosure at $\theta = 1$, the new market value, whatever it is, reveals the value of q through the equation $\gamma_{\text{revealed}} = \gamma(q)$. That is, there is a 'market implied q' in the non-disclosure scenario. Dye assumes that the market has a belief about q, but is mute in regard to how the market has discovered q. We provide a mechanism, whereby the market rationally infers q from its assumed knowledge of observable sector and managerial variability parameters. In the log-normal model above, for instance, the observables are the variances of X and Y.

2.2 Endogenizing the information endowment in the Dye model

The original Dye theory aimed to prove existence of an 'equilibrium non-disclosure' result and was not designed to explain differences in disclosure cutoff between firms. To apply the Dye paradigm within an empirical setting in which firm cutoffs vary, one needs to develop a rational equilibrium model[10], in which different managers choose different Dye cutoffs leading to different observable disclosure practices. To address this, here managers of firms are represented by a one-parameter family of distributions, rather than two parameters, since the ex-ante firm expected value may be omitted from consideration (via renormalization to unity), on the basis of the positive homogeneity (in respect to this parameter) of the Jung and Kwon equation, as discussed above in Section 2.1. Thus manager types are represented by a single parameter called σ; in the case of log-normal models, σ^2 is interpreted as the variance of the underlying normal variable responsible for adding noise to the true firm-value X in the received signal T. Thus, given differences between managers, there may exist incentives for them to act (disclose) differently in equilibrium.

Recalling the Jung and Kwon equation (2), and referring to:

$$z = z(\gamma) := m_X - \gamma, \text{ and } y = y(\gamma) := H_X(\gamma). \tag{8}$$

[10] Our approach starts with the risk-neutral (pricing) distribution of equity value; for an alternative approach to equilibrium pricing in which investor demand arises from mean-variance utility see Suijs (2011).

equation (2) becomes:

$$\frac{1-q}{q}z(\gamma) = y(\gamma),$$

which relate to the shielded-downgrade loss z and value-enhancement potential y. We interpret these two constructs as generalized measures of risk in the sense of Fishburn – see Appendix 2 for a discussion. So, following Fishburn, we model managerial preferences in terms of a utility function $U(z, y)$ (with domain the positive quadrant) over the two constructs z and y. The choice of q is now reduced to determining the solution of the optimization problem[11]:

$$\max_q U(m_X - \gamma(q), H_X(\gamma(q))) = \max_q U(z(\gamma(q)), y(\gamma(q))).$$

Equivalently, the optimal $q = \hat{q}$ may be determined via $\hat{z} = m_X - \gamma(\hat{q})$, where

$$(\hat{y}, \hat{z}) := \arg\max\{U(z, y) : \; y = H_X(m_X - z)\}.$$

This uses the fact that: (i) $\gamma(q)$ is strictly monotone in q, for which see Jung and Kwon (1988), and (ii) the function $H_X(t)$ is strictly convex for a positive density f_X, so the opportunity set $\{(y, z) : y = H_X(m_X - z)\}$ is strictly convex (compare footnote 5 in Section 2.1).

In Appendix 3, we show that: *for $U(y, z)$ concave and homothetic, there will be a unique pair $(\hat{y}, \hat{z})$ solving the maximization problem.*

When $U(y, z)$ is homothetic, or homogeneous of degree 0, that unique pair is characterized by the two optimality conditions[12]:

$$u(\hat{\lambda}) = F(\hat{\gamma}) \text{ and } \hat{\lambda} = \hat{y}/\hat{z}, \tag{9}$$

where $u(\lambda) := U_z(\lambda, 1)/U_y(\lambda, 1)$ will be referred to as the *marginal rate of substitution function*[13]. The second of these two equations is just the Dye equation in the variables y, z.

[11] So the optimization should be read as maximizing a ranking $U(\mu(F, q), \rho(F, q))$ over the parameters $\mu(F, q) = m_X - \gamma(q)$, i.e. an adjusted mean, and $\rho(F, q) = H_X(\gamma(q))$, a lower partial moment, as in the Fishburn analysis of Appendix 2. It is thus capable of being interpreted as an expected utility (rather than a utility of the two quantities μ, ρ which are expectations under the model (F, q)).

[12] The equilibrium condition places constraints on the size of λ, since $0 < F(\gamma) \leq F(m) < 1$. Note that $\lambda < 1$ iff $q > 1/2$.

[13] The marginal rate of substitution of a utility function is homogeneous of degree zero iff the utility function is homothetic or itself homogeneous of degree zero – see Forsund (1975).

For U any one of the text-book homothetic utility functions, the marginal rate of substitution function $u(\lambda)$ is increasing in the odds λ. As the cumulative distribution function $F(t)$ is also an increasing function, the *utility effect* of selecting higher odds in equilibrium is to increase the disclosure cutoff: thus decreasing the endowment parameter q leads to increasing the chances of non-disclosure occurring (when the manager is informed). This parallels and preserves a Dye version of the *gain-to-loss effect*: since here the gain-to-loss ratio, as represented by $\Lambda(\gamma) := H(\gamma)/(m-\gamma)$, is increasing in γ, the Dye equation $\lambda = \Lambda(\gamma)$ implies that increasing the odds λ (exogenously, as is the case in Dye's model), or equivalently decreasing q, leads to an increased cutoff.

The combined effect of the two equilibrium conditions in (9) can best understood in the special case of a Cobb-Douglas utility $U_{\text{C-D}}(y,z) = y^{\alpha}z^{\beta}$ for which one has $u(\lambda) = \beta\lambda/\alpha$ and so

$$\frac{\alpha}{\beta} = \frac{H(\gamma)}{(m-\gamma)F(\gamma)}. \tag{10}$$

Since $H'(\gamma) = F(\gamma)$, the right-hand side represents the 'growth rate' of $H(\gamma)$ (for background see Bingham et al. (1987), especially p. 44), which balances the substitution coefficient α/β. (For the general situation see Appendix 5).

These considerations are applicable to a range of explicit trading mechanisms in which it is possible to derive the implied preferences in the form of a utility function; we call such a derived function the manager's *implied equivalent utility.* In this sense the choice-model is reasonably robust.

2.3 Endogenous optimal disclosure intensity

Having established in the previous section that managers can be viewed as optimizing q, we return now to the fundamental question of what this means about the observed disclosure strategy. It turns out that this is now relatively easy to answer if attention is focused upon the disclosure intensity of such a strategy, shortly to be defined in (11).

Disclosure occurs when management are informed (which occurs with probability q) and furthermore the discovered value, be it the value X or $T(X,Y)$, is above the (respective) cutoff γ, which occurs with probability $1-F(\gamma)$. Here F denotes the probability distribution function F_X with $\gamma = \gamma_X$ if X is observed above γ_X, or F_S with $\gamma = \gamma_S$ for $S = \mathbb{E}[X|T]$ if T is observed above γ_T (cf. Appendix 1). Thus it is natural to define a firm's disclosure intensity τ as:

$$\tau = q(1 - F(\gamma)). \tag{11}$$

So at issue next is whether the firm's disclosure intensities τ varies in a systematic way (e.g. monotonically) with the underlying uncertainty (noise) faced by management, represented by the relevant model parameter σ, be it measured by sectorial or managerial variance (σ_X when X is observed, or by σ_Y when $T = T(X, Y)$ is observed with σ_X fixed). We note that Penno (1997) produced an existence result which suggested that one should not assume that the relationship between τ and σ may be a simple monotonic function, and hence that inference of the relative underlying uncertainty faced by different management from disclosure intensity would be problematic[14]. Actually, this suggestion turns out to have been over-pessimistic, as the following example shows.

The structurally minimal model. We illustrate in a tractable way the theory developed so far, by concentrating on details of the pay-offs that may arise, rather than on abstract utilities as above. We concentrate on the expression $(1-q)z$, i.e. $(1-q)(m_X - \gamma(q))$, which arises in (4), the notation here stressing the dependence of γ on q. With probability $1-q$, a payoff $m_X - \gamma(q)$ may arise to the manager in the following circumstance: the manager knows that no new information is available on the company's future value, but investors have nevertheless downgraded the value of the firm (because of non-disclosure). Here $\gamma(q)$ may be interpreted as γ_X (corresponding to the case when X is observed), or as $\mu_X(\gamma_T)$ (when T is observed). Conditional on this absence of information, the manager could, if permitted, buy the stock at the interim market price γ, and then liquidate the stock at the terminal time. The expected terminal value is m_X given the absence of information. Thus ex-ante the manager holds an option with expected value (under the measure F_X, which is here the risk-neutral valuation measure, or, in a noisy context, its 'noisy replacement' F_S – see Appendix 1) equal to

$$(1 - q)(m_X - \gamma). \tag{12}$$

[14] The work by Penno is a timely reminder of the care that needs to be taken when trying to extend theory to model real world practice. However, the following section shows that in fact the Penno result on non-consistency (monotonicity) arises because of the somewhat restrictive functional form he used. We generalize his result and show what class of probability functions admit consistency.

More generally, if the manager can receive an incentive share α of this value (with $0 < \alpha < 1$) in remuneration, then the expression above may be regarded as the manager's objective function. For instance, this does not necessarily have to involve explicit trading, instead simply assume the manager's share options are set 'at the money' immediately after investors, not seeing a disclosure, downgrade the firm, when the manager knows there is no new company information[15]. The expression $(1-q)(m_X - \gamma(q))$ is known to be strictly concave in q with a unique maximum, whose location we denote by $\hat{q}$; for this see Ostaszewski and Gietzmann (2008) and for an explanation that this expected value here (under F_X) is also a risk-neutral valuation, i.e. realizable through a trading strategy. The trading mechanism does assume that the manager's trade remains *unobserved* by the investors, as would be the case in the traditional Kyle (1985) one-shot market model. We refer to this as the *unobserved trading mechanism*. However, it is possible to relax this assumption (without changing the qualitative features of the results) in a setting with a sequence of observed trades[16], but at the price of less tractability.

The expression (12) above is easiest to interpret when the manager's choice of q becomes known through some mechanism (e.g. inferentially from earlier observations of disclosure intensity) to the market participants, so that γ takes the value $\gamma(q)$; then it becomes clear that the manager chooses $\hat{q}$. A difficulty arises if q is not known in this way, since γ in this expression is selected by the investor, and q by the manager. If $\gamma < m$, then the manager maximizes his objective by taking $q = 0$, i.e. selecting to be always uninformed. We view this difficulty as demanding some co-ordination mechanism for a 'co-operative game' approach, in which a 'first-best' outcome

[15] See Yermack (1997) , Aboody and Kaznik (2000), and Gao and Shrieves (2002) for a discussion of strategic granting of options.

[16] In such a sequential market model (allowing trades at dates in between the interim and terminal dates), the manager's trading, having become observable, is subject to inferential analysis. The revised managerial opportunity set necessitates that the optimal managerial behaviour (given the manager's incentive) employs a mixed strategy of buying and selling – to preserve optimally the manager's informational advantage. On game-theoretic grounds, one expects that the revised valuation of the manager's 'option to trade' is a convex function of q, say of the form $V(q)z(\gamma(q))$, with $V(q)$ taking zero value at the endpoints $q = 0$ and $q = 1$. (In this respect that is similar to the case with $(1-q)(m_X - \gamma(q))$). See De Meyer and Moussa Saley (2002) for a sequential auction model yielding just such a result. Our theory applies also to such general valuations – subject to permitting trades for then (non-verifiable) observations below $\gamma(q)$.

$(1-\hat{q})(m-\gamma(\hat{q}))$ is sought but is not achieveable. In fact, by offering an appropriate incentive scheme, the investor may achieve an outcome *arbitrarily close* to this 'first-best', setting the incentive share α small enough but positive.

We continue, parsimoniously, with this structurally minimalist assumption where a manger's trades or option grants are not observed in a timely fashion (before the final date) and with the manager's objective set at $(1-q)(m_X-\gamma(q))$. This turns out to be a very tractable model. Indeed, we find that under these current assumptions the manager behaves as though he was making a utility maximization in trading-off the shielded-downgrade loss z against the value enhancement potential y (that is, acting out the role defined in Section 2.3) and employing a uniquely determined CES utility function, namely $U(y,z)=(y^{-1}+z^{-1})^{-1}$. Recall that we refer to this as an *implied equivalent utility* (end of last section) in order to stress that the utility function is not imposed, but derived from the managerial payoff structure (12). See Appendix 4 for details.

The principal feature of the structurally minimal model is summarized in the following

Theorem 1. (Optimal Intensity in the structurally minimal model). *In the structurally minimal model the odds $\lambda=(1-q)/q$ and the intensity of disclosure τ sum to unity, i.e.*

$$\tau+\lambda=1,$$

iff the value of q is selected optimally as in ***Section 2.3 above****, i.e. $q=\hat{q}$, or, equivalently $\tau=\hat{\tau}$. In this case the corresponding Dye cutoff, denoted $\hat{\gamma}$, and the odds $\hat{\lambda}$ are related according to the rule*

$$\hat{\lambda}^2=F(\hat{\gamma}(\sigma),\sigma). \tag{13}$$

The result here is driven by the condition (13) which corresponds to the utility function $U(z,y)=(y^{-1}+z^{-1})^{-1}$ derived in Appendix 4. The definition of τ and some simple arithmetic yields[17]:

$$\tau=1-\lambda \text{ iff } F(\gamma,\sigma)=\lambda^2,$$

from which the sum-to-unity formula follows.

[17] $1-F(\gamma)=1-\lambda^2=(1-\lambda)(1+\lambda)$ iff $\tau=q(1-F(\gamma))=1-\lambda$, as $q=1/(1+\lambda)$.

According to this simple rule, the intensity of a firm selecting its optimal odds at λ is negatively linear in λ. As the optimal choice of q increases (across different firms) the intensity rises.

We stress that the simplicity of this formula is evidence of the tractability of the valuation (12).

The more general situation is given by the following result (see Appendix 5 for proof), which includes the structurally minimal case given by $u(\lambda) = \lambda^2$, where $u(\lambda)$ is the marginal rate of substitution of the utility function $U(z, y)$, as defined in Section 2.2 above.

Theorem 2 (First Monotonicity Theorem: Disclosure response to optimal odds). *The intensity of disclosure as a function of the optimal odds is decreasing in the following three circumstances:*

(i) *if* $u(\lambda)$ *is increasing,*

(ii) *if* $u(\lambda)$ *is convex and* $\tau'(0) < 0$, *and*

(iii) *if* $u(\lambda)$ *is concave and* $\tau'(\hat{\lambda}) < 0$, *then* $\tau'(\lambda) < 0$ *for all* $0 < \lambda < \bar{\lambda}$.

2.4 Disclosure effects on the ex-ante cost-of-capital formula

Up to this point the distribution F_X modelled the risk-neutral, or pricing, probability in describing the future equity X, which allowed us to identify γ_X, the no-arbitrage non-disclosure valuation. To avoid misunderstandings, we recall that we work with present-values (so that for our context the riskless rate is zero). We consider next the so-called 'physical' distribution, corresponding to a model of X on the basis of observed past prices, which we denote by $F_X^{\mathbb{P}}$. For clarity (but only in this Section), we will write $F_X^{\mathbb{Q}}$ for the risk-neutral distribution, whenever we want to draw attention to its nature. Thus under the latter we have $m_X = \mathbb{E}^{\mathbb{Q}}[X]$, i.e. the expected value of equity coincides with its market price; however, the same is not true under the physical expectation, where

$$\mathbb{E}^{\mathbb{Q}}[X] = (1 - \delta)\mathbb{E}^{\mathbb{P}}[X], \text{ with } 0 < \delta < 1, \tag{14}$$

as the physical-expectation of equity-value needs to be adjusted by a discount δ which we term the *cost of capital.*

The two measures are connected, just as in the Black-Scholes model, where the Girsanov transformation ('change of measure') shifts the anticipated instantaneous return to the riskless rate (without altering the volatility), thus applying an appropriate 'risk-premium' (or discount) – see Appendix 6.

However, here in the richer disclosure setting, when managerial information endowment is modelled by true (noiseless) equity-value discovery, there is an additional connection beyond the Girsanov transformation resulting from the presence of the interim date. As any disclosed value is the true equity-value, the disclosure-indifference cutoff level is also at 'true value' (or a risk-neutral expectation). In particular, given this information structure, the two measures now agree on the no-disclosure valuation of equity:

$$\gamma_X m_X = \mathbb{E}^{\mathbb{Q}}[X|ND] = \mathbb{E}^{\mathbb{P}}[X|ND].$$

Hence, we may compute δ by reference to the risk-neutral measure as follows. Again for clarity, we continue with the case of the noiseless signal; however, in the noisy signal context, X may be replaced throughout by $S = \mathbb{E}^{\mathbb{Q}}[X|T]$ and γ_X by γ_S. Noting that disclosure occurs with probability $\tau = q(1 - F_X^{\mathbb{Q}}(\gamma))$, one has:

$$\mathbb{E}^{\mathbb{P}}[X|ND] = \tau \int_{x \geq \gamma} x dF_X^{\mathbb{P}}(x|\text{Disclosure}) + (1-\tau)\gamma.$$

Using (14) and renormalizing equity-value to unity as usual, one has

$$\begin{aligned} 1 - \delta &= q(1 - F(\gamma)) \int_{x \geq \gamma} x dF_X^{\mathbb{P}}(x|\text{Disclosure}) + ((1-q) + qF(\gamma))\gamma \\ &= q(1 - F(\gamma)) \int_{x \geq \gamma} x dF(x) + ((1-q) + qF(\gamma))\gamma, \qquad (15) \end{aligned}$$

since disclosed values are true, and $F = F_X^{\mathbb{Q}}$ here is the distribution of true equity values. As before $\gamma = \gamma_X(q)$ for any q (that is, as always, investors believe at the interim date that the asset's expected value given non-disclosure is at the minimum possible, given knowledge of q). We check that $\delta > 0$, so that this does indeed represent a discount over the expected value (under the physical measure).

Substituting into (15) the identity

$$1 = \int_{x \leq \gamma} x dF(x) + \int_{x \geq \gamma} x dF(x) = \gamma F(\gamma) - H(\gamma) + \int_{x \geq \gamma} x dF(x),$$

yields, after some re-arrangement,

$$1-\delta = 1+\gamma F(\gamma)(1-q)+q\gamma F(\gamma)^2 - F(\gamma).$$

Hence, adding and subtracting two terms, one has

$$\delta = F(\gamma)(1-q-\gamma(1-q)+qF(\gamma)-q\gamma F(\gamma))+qF(\gamma)-qF(\gamma)^2$$

which leads to

$$\delta = \delta(q) = F(\gamma)[(1-\tau)(1-\gamma)+\tau],$$

which is non-negative, and in fact positive, for $0 < q < 1$. In a noisy signal context this formula for X remains valid with X replaced by S and γ_X by γ_S.

In particular, in the structurally minimal model of Section 2.3, for the optimized choice of $q = \hat{q}$ (see Theorem 1), one has

$$\hat{\delta} = \delta(\hat{q}) = \hat{\lambda}F(\hat{\gamma})(1-\hat{\gamma})+\hat{\tau}F(\hat{\gamma}). \tag{16}$$

Rewriting this formula as

$$\hat{\delta} = \hat{\lambda}F(\hat{\gamma})(1-\hat{\gamma})+\hat{q}\cdot F(\hat{\gamma})(1-F(\hat{\gamma})) = H(\hat{\gamma})F(\hat{\gamma})+\hat{\tau}F(\hat{\gamma}), \tag{17}$$

we see on the right-hand side two terms that are easily interpreted.

First, there is the ex-post cost-of-capital weighted by the odds against information discovery and the probability of a below-cutoff discovery. Equivalently, via the Dye equation, this balances the lost 'upward potential' in any discovery (refer to Section 2.1), a matter we return to below.

Secondly, $\hat{q}\cdot F(\hat{\gamma})(1-F(\hat{\gamma}))$ is the expected conditional variance of non-disclosure, conditioning on the early discovery of information, i.e. $\mathbb{E}[\text{var}(1_{ND}|1_E)]$. This measures the informativeness of non-disclosure (ND) about low value discovery. We term this the *ND-informativeness* to distinguish it from other definitions of informativeness.

In Statistics this term is the *unexplained variance* or more completely the 'unexplained component of total variance' – variance in 1_{ND} not accounted for by the knowledge of 1_E, see Bingham & Fry (2010), Th. 4.20, (see also Cox et al. (2003), Searle et al. (1992)).

In summary, in the structurally minimal model:

$$\hat{\delta} = \text{ lost upward potential } + \text{ ND-informativeness.}$$

More generally, using the identity from footnote 6 of Section 2.1, one obtains the decomposition of the equity value $(1-\gamma)(1-\tau)$, lost through non-disclosure, as

$$(1-\hat{\tau})(1-\hat{\gamma})-\hat{\delta}=\tau H(\hat{\gamma})+(1-\hat{\gamma})\cdot\hat{q}F(\hat{\gamma})(1-F(\hat{\gamma})), \tag{18}$$

i.e. with $\mathbb{E}$ signifying market-valuation (risk-neutral expectation):

$$\begin{aligned}&\mathbb{E}[\text{cost-of-capital}^{\text{ex-post}}|\text{ND}]-\text{cost-of-capital}^{\text{ex-ante}}\\ =\;&\mathbb{E}[\text{upward potential}|\text{D}].\\ &+\text{ND-informativeness}\times\text{cost-of-capital}^{\text{ex-post}}.\end{aligned}$$

It is no surprise that the Dye cutoff determines the cost of capital; the surprise is how transparently it does so. On the left-hand side is:

the *incremental cost of capital* arising through non-disclosure.

This balances on the right-hand side:

the inevitable loss of any valuation up-grade associated with a disclosure (above cutoff) together with a further penalty for low-value discovery, moderated by the informativeness of non-disclosure,

Numeric investigation of $\delta(q)$ in the structurally minimal case for the log-normal models of Section 4 finds it to be increasing in q. We consider in Section 3 the monotonic dependence of $\hat{q}$ on sources of uncertainty about the firm (sector/managerial variability) and report on numerical investigation of the corresponding dependence of $\hat{\delta}$ on these uncertainties.

3 Monotonic linking of disclosure intensity τ to signal variance

In Theorem 2 we have just traced the dependence of τ on $\hat{\lambda}$. In the noisy signal model $T=T(X,Y)$ there are two sources of uncertainty:
(i) in X, i.e. in the sector return variability, as captured by the variance σ_X^2, and
(ii) in Y, i.e. in the variability of the noise which models the managerial 'vision', as captured by the variance σ_Y^2 (or the precision $1/\sigma_Y^2$).
Equivalently, in the case of the log-normal model $T=XY$ of Section 2.1 and Section 4 these variabilities are captured respectively by σ_u^2 and σ_v^2, the variances of the corresponding underlying normals.

This formalization focusses on managerial variability as a way to model differences between the firms of a single sector regarded as having a single parameter of riskiness. Since we do not wish to preclude cross-sector comparisons, we will work with a general family of distributions parametrized by $\sigma = \sigma_{\text{agg}}$, (to signify a general aggregation of sectorial and managerial variances); from there we trace dependencies on the two separate sources. Thus, we identify henceforth the corresponding Dye cutoff as $\gamma(\lambda, \sigma)$.

We continue to use hatted notation to refer to situations where the uncertainty parameter q, or λ, takes its optimal value, thus we write $\hat{\lambda}(\sigma)$ and $\hat{\gamma}(\sigma) := \gamma(\hat{\lambda}(\sigma), \sigma)$. That is, as in (9), $\hat{\lambda}$ and $\hat{\gamma}$ solve the two equations:

$$\lambda = H_X(\gamma, \sigma)/(m_X - \gamma) \text{ and } u(\lambda) = F_X(\gamma, \sigma),$$

or the amended version for a noisy observation T (cf. Section 2.1 or Appendix 1).

We now consider the dependence of $\hat{\lambda}$ on the aggregate variability σ^2_{agg}. For this work we will need to exploit stochastic dominance (and a refinement of the first-order notion – see Appendix 7). This naturally complements the Dye model (in which risk preferences were embodied in the one risk-neutral distribution F_X), as now investor risk-preferences as between different distributions, i.e. as between different σ^2_{agg}, need to be introduced. This should comes as no surprise in view of the seminal work by Vijay Bawa (1975) on lower partial moments, where he mapped out the relationship between lower partial moment and stochastic dominance. Indeed Bawa (1975) was the first to consider the lower partial moment (LPM) as creating a general family of *below-target* risk measures, among them the below-target semivariance, and studied them in regard to risk tolerance. See Levy (1992) for a survey, or Nawrocki (1999) for a more recent review of the issues.

We will see first that $\hat{\gamma}$ decreases with σ (subject to dominance assumptions). We then formulate an assumption about the preferences of investors facing increased risk; from this and the comparative statics of $\hat{\gamma}$ follows the sensitivity to changes in σ of $\hat{\lambda}$, or equivalently of $\hat{q}$. This may be reduced to a modeling condition on the distribution F_X to be used in relation to the log-normal models in the next section (Section 4). This section ends with the corollary that under the circumstances $\hat{\tau}$, the equilibrium value of the disclosure intensity, decreases with σ_{agg}.

The entire analysis of these comparative statics is necessitated by the fact that the statics conducted by Jung and Kwon (1988) are inappropriate here, because theirs is a 'partial statics analysis' with one parameter held fixed namely q; we need to relax their assumption and allow variation in q as well as in the distributions (which in our case are parameterized by σ). As commented before, in Section 2.1, we hold the expected firm value m_X fixed.

3.1 Stochastic dominance

Recall that F_1 dominates F_2 in the sense of *first degree stochastic dominance* (FDSD) if $F_1 \neq F_2$ and $F_1(t) \leq F_2(t)$ for all t, i.e. $\bar{F}_1(t) \geq \bar{F}_2(t)$ for all t, so that the event $X_1 \geq t$ (of interest in regard to disclosure) is more likely than $X_2 \geq t$, so that the former is preferred over the latter.

Thus a family of distributions $F(t, \sigma)$ parametrized by variance σ exhibits FDSD if for all t

$$F(t, \sigma_1) \leq F(t, \sigma_2) \text{ provided } 0 < \sigma_1 < \sigma_2.$$

Likewise F_1 dominates F_2 in the sense of *second degree stochastic dominance* (SDSD) if $F_1 \neq F_2$ and $H_1(t) \leq H_2(t)$ for all t.

Recall two relevant results from Jung and Kwon (1988). By their Prop. 2 one has[18] for fixed σ

$$\gamma(\lambda_1, \sigma) < \gamma(\lambda_2, \sigma) \text{ provided } 0 < \lambda_1 < \lambda_2. \tag{JK1}$$

Also their Prop. 3, for fixed q, implies[19]

$$\gamma(\lambda, \sigma_2) \leq \gamma(\lambda, \sigma_1) \text{ provided } 0 < \sigma_1 < \sigma_2, \tag{JK2}$$

when the family $F(t, \sigma)$ parametrized by variance σ exhibits FDSD or SDSD.

We prove the complementary result for $\hat{\gamma}$ below in Theorem 2 that, subject to stochastic dominance assumptions,

$$\hat{\gamma}(\sigma_2) := \gamma(\hat{\lambda}(\sigma_2), \sigma_2) < \gamma(\hat{\lambda}(\sigma_1), \sigma_1) := \hat{\gamma}(\sigma_1). \tag{19}$$

This does not follow from their results, because here in fact $\hat{\lambda}(\sigma_2) > \hat{\lambda}(\sigma_1)$, or equivalently, $\hat{q}(\sigma_2) < \hat{q}(\sigma_1)$. For the latter result see the Section 3.2. Our first statics result is as follows: for a proof and technical terms here, see Appendix 7.

[18] In their notation this result would read $\gamma(q_1, \sigma) < \gamma(q_2, \sigma)$ provided $0 < q_2 < q_1$, since λ is decreasing in q.

[19] As before, in their notation, one has $\gamma(q, \sigma_2) \leq \gamma(q, \sigma_1)$ provided $0 < \sigma_1 < \sigma_2$.

Theorem 3 (Cutoff statics). *Given two distributions with log-concave hemi-means, with $F_1 = F(.,\sigma_1)$ increasingly dominating $F_2 = F(.,\sigma_2)$ for $\sigma_1 < \sigma_2$, the corresponding optimized Dye cutoffs satisfy*

$$\hat{\gamma}(\sigma_2) < \hat{\gamma}(\sigma_1), \text{ i.e.,} \gamma(\hat{\lambda}(\sigma_2), \sigma_2) < \gamma(\hat{\lambda}(\sigma_1), \sigma_1) \text{ provided } 0 < \sigma_1 < \sigma_2.$$

Thus the cutoff falls if the aggregate variability σ^2_{agg} increases.

3.2 Monotonicity and Investor Preferences

Here we argue that in modelling a firm by a family of distributions $F(x,\sigma)$ corresponding either to a perfectly observed value X or its estimate $\mu_X(T)$, i.e. $S := \mathbb{E}[X|T]$ (based on the observed signal $T = T(X,Y)$), the family should reflect the fact that investors require to be rewarded for accepting increased risk, when appropriately measured. Our approach is *axiomatic* – rather than introduce a further modelling micro-structure involving the investors selecting which manager to employ by some formal mechanism. We take the view that investors employ an overall risk-measure of below-target-shortfall relative to a threshold t and variance σ given by the gain-to-loss ratio Omega, introduced in Section 2.1 and recalled here as

$$\Omega(t,\sigma) = \frac{\int_{x \geq t}(1 - F(x,\sigma))dx}{\int_{x \leq t} F(x,\sigma)dx}.$$

There is further support in the literature for the view taken by Shadwick and Keating (2002) on this performance ratio – see C. S. Pedersen and S. E. Satchell, (2002), and A. E. Bernardo and O. Ledoit (2000): the Omega ratio, which is intimately connected to the *Sortino* ratio, should be regarded as an adjusted type of *Sharpe* ratio. Our interest focusses on the fact that it is usual in the Omega approach to assume that investors' risk-preference is to rank one pair (t,σ) as preferred over another, when $\Omega(t,\sigma)$ is larger. Indeed $\Omega(t,\sigma)$ is decreasing in t and in σ (see Cascon, Keating and Shadwick (2003)). We make the following monotonicity assumption.

Assumption MIEP (Monotonic Investor Equilibrium Preferences). *We assume that, following an increase in σ, the equilibrium 'market-demand for Ω', by way of compensation, is strictly positive, i.e. as a function of σ the equilibrium Ω-value decreases.*

The axiom thus captures the effect that, even when the disclosure cut-off is lowered to share more information with investors, the riskiness of the firm is only *partially* offset; the Ω value can only be compensated by reducing the σ. Our attention on the Omega function is motivated by its natural connection to the Dye equation (see 6) and the very simple form that it gives that equation.

In view of the reciprocal identity (compare (5)) linking $\Omega(t,\sigma)$ with $H(t,\sigma)$, the hemi-mean function corresponding to the distribution $F(t,\sigma)$, we will usually expect that the hemi-mean function $H(t,\sigma)$ (i.e H_X or H_S, as the case may be) is increasing in σ, so that decreases in γ are counter-balanced by increases in σ. This is the case for the log-normal hemi-mean function $H_{\mathrm{LN}}(\gamma,\sigma)$, i.e. it is increasing in σ, as its 'vega' is $D_\sigma H_{\mathrm{LN}}(\gamma,\sigma) = \varphi(d_1) = \varphi((\log\gamma - \frac{1}{2}\sigma^2)/\sigma)$ – for which see J.C. Hull (2011) (3rd ed. or later), and so is positive. (Likewise, $D_\sigma H_{\mathrm{N}}(\gamma,\sigma) = \varphi(\gamma/\sigma)$.)

In order to study changes in $\hat{\lambda}$ in response to σ, we first rewrite the equilibrium condition of Section 2.3 for utility optimization, namely

$$u(\lambda) = F(\gamma,\sigma),$$

in the format $\lambda = \Pi(\gamma,\sigma)$ similar to

$$\lambda = \frac{H(\gamma,\sigma)}{1-\gamma}$$

(working as usual with a re-normalized risk-neutral mean, i.e. $m_X = 1$) by referring to the inverse function of $u(\lambda)$ and taking

$$\Pi(\gamma,\sigma) := u^{-1}(F(\gamma,\sigma)).$$

Given this, rather than eliminate γ between the two equations $\lambda = H(\gamma,\sigma)/(1-\gamma)$ and $\lambda = \Pi(\gamma,\sigma)$, we relate the MIEP to the Jacobian of the bi-variate system of equations.

Theorem 4 (Jacobian condition). *Suppose $H(\gamma,\sigma)$ is log-concave as a function of γ, that $F(x,\sigma)$ exhibits first order stochastic dominance with respect to (increasing) $\sigma = \sigma_{\mathrm{agg}}$ and that the transformation $(\gamma,\sigma) \to (\pi,\lambda)$ given by:*

$$\begin{aligned}\lambda &= \frac{H(\gamma,\sigma)}{1-\gamma} = \frac{1}{\Omega(\gamma,\sigma)-1},\\ \pi &= u^{-1}(F(\gamma,\sigma)),\end{aligned}$$

is invertible. The Assumption MIEP is equivalent to the Jacobian (determinant) $\partial(\lambda, \pi)/\partial(\gamma, \sigma)$ being positive for $0 < \gamma < 1$ and $\sigma > 0$, i.e. to the condition

$$\frac{\partial(\lambda, \pi)}{\partial(\gamma, \sigma)} = \frac{1}{u'(\hat{\lambda})(1-\gamma)\,(\Omega-1)^2} \begin{vmatrix} -\Omega_\gamma & -\Omega_\sigma \\ F_\gamma & F_\sigma \end{vmatrix} > 0,$$

or in terms of distributional data:

$$\frac{\partial(\lambda, \pi)}{\partial(\gamma, \sigma)} = \frac{1}{u'(\hat{\lambda})(1-\gamma)} \begin{vmatrix} F(\gamma,\sigma) + H(\gamma,\sigma)/(1-\gamma) & H_\sigma(\gamma,\sigma) \\ F_\gamma(\gamma,\sigma) & F_\sigma(\gamma,\sigma) \end{vmatrix} > 0.$$

This is just the two-variable analogue of the well-known positive slope condition used to characterize monotonic increasing functions of one variable. The condition asserts that the transformation $(\gamma, \sigma) \to (\lambda, \pi)$ from $[0, 1) \times [0, 1)$ to $[0, 1) \times [0, 1)$ defined above preserves orientation, a feature that is equivalent to the Assumption MIEP.

For any of the standard text-book homothetic utility functions, as remarked in Section 2.2, one has $u'(\lambda) > 0$ and so since $\gamma < 1$ the Theorem identifies the simple condition

$$J(\gamma, \sigma) := \begin{vmatrix} F + H/(1-\gamma) & H_\sigma \\ F_\gamma & F_\sigma \end{vmatrix} > 0,$$

which condition holds for F the log-normal distribution, but fails for the normal (as there the determinant is zero).

As an immediate corollary, we obtain the following monotonicity result. For technical terms here see Appendix 7.

Theorem 5 (Second Monotonicity Theorem: Disclosure response to Aggregate variability). *Assume MIEP and that H is log-concave. For $u(\lambda)$ strictly and regularly increasing, $\hat{\tau}$ is decreasing in $\sigma = \sigma_{\text{agg}}$.*

Proof. By MIEP $\Omega(\hat{\gamma}(\sigma), \sigma)$ is decreasing in σ. Since $1 = (1-\hat{q}(\sigma))\Omega(\hat{\gamma}(\sigma), \sigma)$, the endowment parameter is decreasing and so $\hat{\lambda}(\sigma)$ is increasing in σ. Hence, $\hat{\tau}(\sigma)$ is decreasing (by Theorem 2), as required.

It is significant that Theorem 5 and the condition of Theorem 4 depend on the choice of $u(\cdot)$ only in regard to the sign of u', that is – the monotonicity is *robust* as to the details of the utility expressing managerial preferences (or the implied equivalent utility of a trading mechanism under which the manager may act) – see Section 2.2.

Our final aim is to apply Theorem 5 to the log-normal model of a manager receiving a noisy signal of firm value to show that $\hat{\tau}$ increases as the managerial vision deteriorates.

4 Log-normal models

In this section we note the details of the log-normal model. The explicit formulas quoted below show its tractability, all the way down to an application of Theorem 4 which yields the desired monotonicity result.

Recall that we are concerned with the possibility that at time $\theta = 2$ the manager observes either true value X, or a transform $T = T(X, Y)$ of the random variable X with Y a source of noise. Here we take $T = XY$ with $X = m_X e^{u - \frac{1}{2}\sigma_u^2}$ and $Y = e^{v - \frac{1}{2}\sigma_v^2}$ with u, v the underlying independent, normal zero-mean random variables (in the sense of the risk-neutral probability) with variances σ_u^2 and σ_v^2. Thus X is log-normally distributed, in accord with the financial benchmark model, as is the signal T. Here $T = m_X e^{w - \frac{1}{2}\sigma_w^2}$ with $w = u + v$ a mean-zero normal with variance

$$\sigma_w^2 = \sigma_u^2 + \sigma_v^2. \tag{20}$$

In this context F_{LN}, the cumulative distribution and H_{LN} the hemi-mean function for the log-normal model, are given by

$$\begin{aligned} F_{\mathrm{LN}}(\gamma, \sigma) &= \Phi_{\mathrm{N}} \left(\frac{\log(\gamma) + \frac{1}{2}\sigma^2}{\sigma} \right), \\ H_{\mathrm{LN}}(\gamma, \sigma) &= \gamma \cdot \Phi_{\mathrm{N}} \left(\frac{\log(\gamma) + \frac{1}{2}\sigma^2}{\sigma} \right) - \Phi_{\mathrm{N}} \left(\frac{\log(\gamma) - \frac{1}{2}\sigma^2}{\sigma} \right), \end{aligned}$$

where Φ_{N} denotes the standard normal probability distribution function. By (3) this is just the Black-Scholes put-formula for zero riskless rate, unit time to expiry, unit initial asset value, and strike γ. Hence one readily deduces that the relevant Jacobian for Theorem 4 is (up to a positive factor of proportionality) given by:

$$\begin{aligned} J(\gamma, \sigma) &= \gamma d_1 (\Phi_{\mathrm{N}}(d_1) - \Phi_{\mathrm{N}}(d_2)) - (1 - \gamma)\varphi_{\mathrm{N}}(d_1), \text{ where} \\ d_1 &= (\log 1/\gamma) + \frac{1}{2}\sigma^2)/\sigma, \text{and } d_2 = (\log 1/\gamma) - \frac{1}{2}\sigma^2)/\sigma. \end{aligned}$$

We have verified numerically that for $0 < \gamma < 1$ and a range of σ's this is positive, hence fulfilling the conditions of Theorem 4.

We denote by $\gamma_{\rm LN}(\lambda, \sigma)$ the unique solution of the Jung and Kwon equation for a log-normal random variable with unit mean and volatility σ

$$\lambda(1-\gamma) = H_{\rm LN}(\gamma, \sigma). \tag{21}$$

Then the clean (Dye) signal cutoff for an observation of true value X is given by

$$x = m_X \cdot \hat{\gamma}, \tag{22}$$

where $\hat{\gamma} = \gamma_{\rm LN}(\lambda, \sigma_u)$.

Since T is log-normal, it is straightforward[20] from the put-formula above to compute the regression function $\mu_X(t)$, which is

$$\mu_X(t) = \mathbb{E}[X|T=t] = m_X e^{\frac{1}{2}\kappa(1-\kappa)\sigma_w^2} \left(t/m_X\right)^{\kappa},$$

with κ the usual (normal) regression coefficient, and this is strictly increasing in t (as required by the Isotonic Theorem of Appendix 1); the conditional expectation estimator of Section 2.1 is then given via substitution using $T = m_X e^{w-\frac{1}{2}\sigma_w^2}$, so that in terms of $w = u + v$ one has:

$$X^{\rm est} = \mathbb{E}[X|T] = m_X \exp\left(\kappa w - \frac{1}{2}\kappa^2\sigma_w^2\right) = m_X \exp\left(\kappa w - \frac{1}{2}\kappa\sigma_u^2\right).$$

Here, the regression coefficient is

$$\kappa := \frac{\sigma_u^2}{\sigma_u^2+\sigma_v^2} = \frac{p_v}{p_u+p_v} = \frac{p_v/p_u}{1+p_v/p_u}, \text{ employing the precision } p_u = 1/\sigma_u^2, \text{etc.} \tag{23}$$

The aggregate variability is thus given by:

$$\kappa^2\sigma_w^2 := \frac{\sigma_u^4}{(\sigma_u^2+\sigma_v^2)^2}(\sigma_u^2+\sigma_v^2) = \frac{\sigma_u^4}{\sigma_u^2+\sigma_v^2} = \kappa\sigma_u^2.$$

As a result the cutoff for the estimator $X^{\rm est}$ is given by

$$\hat{x}^{\rm est} = m_X \cdot \hat{\gamma}^{\rm est}, \text{ where } \hat{\gamma}^{\rm est} = \gamma_{\rm LN}(\hat{\lambda}(\kappa\sigma_w), \kappa\sigma_w). \tag{24}$$

From here one readily deduces the following:

[20] Details available from the authors

Theorem 6 (Log-normal disclosure intensity). *With the notation of (20), (23) and (24), the disclosure intensity in the noisy log-normal model is given by*

$$\hat{\tau} = \hat{q}(\kappa\sigma_w)\left(1 - F_{\mathrm{LN}}(\hat{\gamma}^{\mathrm{est}}, \kappa\sigma_w)\right),$$

and $\hat{\gamma}^{\mathrm{est}} = \gamma_{\mathrm{LN}}(\hat{\lambda}(\kappa\sigma_w), \kappa\sigma_w)$.

The above formulation in general needs $u(\lambda)$ to be given in order to determine $\hat{q}$ (or $\hat{\lambda}$). The next result completes our analysis and in the case of σ_v justifies the title of the paper.

Theorem 7 (Counter-effects of sector and managerial variability). *For the log-normal model with $u(\lambda)$ strictly increasing, the optimal disclosure intensity $\hat{\tau}$ is decreasing in σ_u^2 (for σ_v^2 constant) and increasing in σ_v^2 (for σ_u^2 constant).*

Proof. Noting that the Jacobian condition depends on $u'(\lambda)$ only in relation to its sign, we may take $u(\lambda) = \lambda^2$ and check that the Jacobian condition is fulfilled. So $\hat{\lambda}$ is increasing in σ_{agg}^2 (by Theorem 3). Now $\hat{\tau}$ is decreasing as a function of $\hat{\lambda}$ (Theorem 1). The conclusion is now clear, because

$$\sigma_{\mathrm{agg}}^2 = \frac{\sigma_u^4}{\sigma_u^2 + \sigma_v^2} = \frac{\rho}{1+\rho}\sigma_u^2, \text{ with } \rho = \sigma_u^2/\sigma_v^2, \tag{25}$$

as we have seen above, so σ_{agg} is increasing in σ_u^2 (for σ_v^2 constant) and decreasing in σ_v^2 (for σ_u^2 constant) – equivalently, increasing in the 'relative precision' ρ.

Returning to the cost of capital formula of Section 2.4, numerical investigation of the above log-normal models implemented in the structurally minimal model, finds $\hat{\delta}(\sigma)$, to be increasing in $\sigma = \sigma_{\mathrm{agg}}$. From equation (25), we see that the cost of capital is increasing in the sectorial variance (when the managerial variance is held fixed), but is decreasing in the managerial forecast variance σ_v (when the sectorial variance is held fixed). Larger managerial precision (smaller managerial forecast variance σ_v) leads to a higher probability of discovery, a larger likelihood of undisclosed bad news, and a larger cost of capital.

The seemingly opposing effects above are easy to understand: the probabilistic discovery of new information by the manager reduces the aggregate variability σ_{agg} of the future equity value.

By Theorem 3, quite generally, $\hat{\gamma}(\sigma)$ is decreasing in $\sigma = \sigma_{\text{agg}}$, whilst $\hat{\delta}(\sigma)$ is increasing in $\sigma = \sigma_{\text{agg}}$. Consequently the relationship between cost-of-capital and disclosure is found to be positive in the sense that as the cutoff decreases, the range of values disclosed increases. By monotonicity, the cost of capital increases with the range of values that firms endogenoeusly choose to disclose.

5 Empirical evidence for disclosure intensity effects

A recent paper in this area is Cousin and de Launois (2006). In their work they consider traditional competing models of conditional volatility: the GARCH specification and a Markov two-state volatility switching model. They argue that changes in the rate of information arrival may cause a switch between high or low (stock return) volatility. In their GARCH framework the specification of conditional variance is given by

$$\sigma_{i,t}^2 = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2 + \lambda_i N_{i,t}, \tag{26}$$

where the new term $N_{i,t}$ is a proxy[21] for the number of news events specific to company i announced to the stock market per interval t. Their main objective is to compare and contrast the performance of this adjusted GARCH model to a two-state Markov Switching Regression (MSR) model, where now the disclosure intensity determines the probability that a company under consideration is either[22] in a low or high volatility regime.

What is of particular interest for us is that, on the basis of an empirical analysis, they conclude that disclosure intensity is an important explanatory variable for conditional volatility. In the GARCH framework their empirical findings are consistent with our theoretical predictions in that the conditional volatility is increasing in disclosure intensity, and in the MSR framework the probability of being in the high volatility state is increasing in disclosure intensity. Thus their empirical tests appear to be broadly in line with our

[21] They measure the variable by identifying the frequency of a subset of firm news releases on Factiva.

[22] To be more precise, the disclosure intensity in part determines whether the state regime dummy variable $D_{i,t}$ is above or below a threshold, qualifying whether the firm is in the high volatility regime.

theoretical predictions. However, before coming to this conclusion we believe it is important to raise a note of caution. What is critical is how Cousin and de Launois measure disclosure intensity. As their Table 1 makes clear, they simply record the frequencies of Factiva disclosures by category. However, if one just records all the raw empirical disclosure intensities for companies, this does not capture the essential features of our generalized Dye model, for the following reason. The theoretical model is of *voluntary* disclosures, that is the Dye model concerns itself only with those news-wires which correspond to management discovering information about future events that affect their *voluntary* ability to issue the news-wires, and thus may indicate value *above* the Dye cutoff. In addition companies are required under regulatory provisions to make *mandatory* disclosures. Thus the raw data on disclosure intensities is a mix of disclosure 'types', whereas the theory only speaks to voluntary disclosures. Thus, when working with raw disclosure-intensity data, an essential step is to implement an estimation procedure for separating out the voluntary Dye-type disclosures.

With this empirical issue in mind, one procedure could be to exploit the distributional assumptions of the model. The Dye cutoff can be shown to be close to the mean (just below), and one can use this to validate an empirical approach which measures *dimensionless relative intensity,* i.e. excess relative to the mean in proportion to standard deviation. Looking at disclosure intensities above the mean rate ('high rates') abstracts away from mandatory good news disclosures that occur on a regular basis. Restricting attention only to high-intensity disclosure periods, one needs to distinguish between those that approximately correspond to good news (voluntary disclosures) and those that approximately correspond to bad news (mandatory disclosures); the latter are typically driven by regulations put in place to protect investors from delay of bad news disclosure. In order to identify which are good news and which are bad news disclosures, when there is no established standard "message space" for voluntary disclosures, it is suggested here that one could identify good news disclosures as those that give rise to an increase in analysts' consensus forecasts (and so exclude those that give rise to a decline in analysts' consensus forecasts for the company).

In contrast recent research by Rogers, Schrand and Verrecchia (2008) (RSV) use an EGARCH model which allows them to estimate the conditional variance when modeled as being given by one of two functions, the choice depending on the sign of the return shock. The intuition behind this asymmetric modeling assumption is that "bad news" seems to have a more pronounced effect on conditional volatility than has "good news". For many companies there is a strong negative correlation between the current stock returns and future volatility. The tendency for conditional volatility to decline when returns rise (following good news) and rise when returns fall (following bad news) is typically referred to in behavioural finance as the *leverage effect*. RSV propose that, when companies follow a strategy of reporting good news and withholding bad news, this can be described as 'strategic disclosure'. In a setting where good news is taken at face value, bad news below the cutoff threshold has to be inferred by investors; it is this difference (i.e. observed versus inferred) in the formation of expectations that leads to the asymmetric responses in the market. To see this in the limiting case of full disclosure, remove the leverage (asymmetric) effect, whereupon current changes in valuation (impounded in returns) would always be associated with recent news arrival rather than the need for investors to make inferences following non-disclosure. Rather than look at actual disclosures, RSV instead develop two hypotheses about the leverage effect. The first is that the leverage effect is stronger for companies about which there is less private information; that feature is assumed to increase the threshold level of disclosure (implying a lower disclosure intensity). The second is that the leverage effect will be weaker when increased litigation risk affects a company's propensity to adopt a 'strategic disclosure' strategy. RSV report interesting results; however, our research on disclosure intensity suggests an alternative empirical implementation. Specifically, they use the variable PUBINFO as a measure of private information. That measure captures the extent to which information is likely not to be private, because in their analysis, if company returns move together then, ceteris paribus, homogeneity subsists in that sector of industry; so there is less private information when results of company operations are similar. Thus, they do not actually measure disclosure intensities. Accordingly, on the view that our model may have wider empirical applicability than the special two-case scenario investigated by RSV, we suggest that an EGARCH model variant of the standard GARCH model, redesigned so as to refer to disclosure intensities in (26), bears investigation.

6 Conclusion

By adapting the Dye (1985) model, this paper addresses Shin's (2006) challenge to understand corporate disclosure by studying a model of asset pricing that endogenously includes disclosure. Our model sheds light on three issues raised in the literature: (i) Larcker and Rusticus (2010) suggested that the relation between disclosure and cost of capital could perhaps be positive despite earlier claims to the contrary; (ii) van Buskirk's (2011) counter-intuitive empirical findings that higher disclosure is associated with reduced information asymmetry; (iii) Penno (1997) suggests that there might not be any monotonicity at all on the relation between informational endowment and disclosure intensity.

The Dye (1985) model of partial disclosure contains a latent variable: the chance discovery of new information; if the chance of discovery is turned into a choice variable and disclosure is voluntary, then it is possible to price the 'option to disclose' and hence to establish a cost-of-capital formula involving explicit disclosure components derived from the disclosure cutoff for discovered information.

In our model the *cutoff* is determined by the managerial variability and the sector variability; it is increasing in the managerial variance (when the sectorial variance is held fixed), and decreasing in the sectorial variance (when the managerial variance is held fixed) – i.e. implementing a larger range of disclosures.

Likewise, the *disclosure intensity* is decreasing in the sectorial variance (when the managerial variance is held fixed), but is increasing in the managerial variance (when the sectorial variance is held fixed) – that is, in a market with arbitrage-free pricing, managers with relatively low forecast precision (of the equity value of their firm) select a relatively high disclosure intensity in order to reduce the 'uncertainty gap'.

Contrarily, the *cost of capital* is increasing in the sectorial variance (when the managerial variance is held fixed), but is decreasing in the managerial variance (when the sectorial variance is held fixed). A reduction in managerial forecast variance leads to smaller probability of discovery and so smaller likelihood of undisclosed news; for the same reason, larger forecast variance leads to larger cutoff (larger range of undisclosed news).

Consequently the relationship between cost-of-capital and disclosure is found to be *positive,* in the sense that the cutoff decreases, i.e. the range of values disclosed increases, as does the cost-of-capital.

Shin warned that in an asset pricing model with endogenized disclosure relations between quantities of interest could turn out to be counter-intuitive. Here, the seemingly opposing effects noted above arise, because the probabilistic discovery of new information by the manager reduces the aggregate variability σ_{agg} of the future equity value.

The theory developed here suggests new empirical testing procedures and also, critically, a need to control for different directions in hypothesized causation. The theory shows why one should not base empirical hypotheses on an a priori assumption that 'better' companies make more voluntary disclosures. In fact, as we have shown, it is companies with the most poorly informed management (facing highest noise) that will – in equilibrium – disclose with the greatest intensity, albeit this may be offset by a cross-sector variability effect.

Furthermore, our strong monotonicity results provide an *empirical proxy* for the hitherto latent variable of information discovery namely the disclosure intensity which is statistically inferrable.

Whilst we focus on log-normal models of equity (following Black and Scholes), the framework is robust to a range of changes in its its microstructure, for which the findings remain unchanged, as the various theorems testify (in particular this justifies the above mentioned empirical proxy). The research is subject to a number of caveats. We abide by the assumptions of the Dye model in that, when disclosing, the manager does so truthfully and, when uformed, the manager is unable credibly to *announce* the absence of information. Furthermore, the manager's interests and those of the investors are aligned. The model is essentially a single-period project model, in which success in one period does not influence successes in later periods. That is, multi-period project dependence (and related disclosure) is not modeled. This is clearly a topic for future research.

7 References

Aboody, D. and Kaznik, R., (2000), "*CEO stock option awards and the timing of corporate voluntary disclosures*", Journal of Accounting and Economics 29, 73-1000.

An, M.Y., (1998), "*Log-concavity versus Log-convexity: a complete characterization*", Journal of Economic Theory, 80, 350-369.

A. E. Bernardo and O. Ledoit (2000), "*Gain, loss and asset pricing*", J. Political Economy, vol 108 no.1, 144-172.

Bergstrom, T.C. and M. Bagnoli, (2005), "*Log-concave Probability and its Applications*", Economic Theory, 26, 445-469.

Bingham, N. H. , Goldie, C. M. , Teugels, J. L. , *Regular variation,* 2nd edition, Encycl. Math. Appl. 27, Cambridge University Press, Cambridge, 1989 (1st edition 1987).

Bingham, N. H. and Fry, J.M., *Regression: Linear Models in Statistics,* SUMS, Springer, 2010.

Caplin, A. and B. Nalebuff, (1991), "*Aggregation and social choice: a mean voter theorem",* Econometrica 59 (1) 1-24.

Cascon, A., Keating C. and Shadwick, W., (2002/3), "*The Omega Function*", working paper.

Clinch, G., and Verrecchia, R.E., (2011) "*Endogeneous disclosure choice",* working paper.

Cox, D. R., Solomon, P. J., (2003), "*Components of variance*", Chapman & Hall.

Cousin, J.-G. and T. de Launois, (2006), "*News Intensity and Conditional Volatility on the French Stock Market", Finance,* Presses Universitaires de France, Vol. 27.1, pages 7-60.

De Meyer, B. and Moussa Saley, H., (2002), "*On the strategic origin of Brownian motion in finance*", International Journal of Game Theory, 31 (2002), 285-319.

Dye, R.A., (1985), "*Disclosure of Nonproprietary Information*", Journal of Accounting Research, 23, 123-145.

Easley, D., Hvidkjaer, S., and O'Hara, M., (2002), "*Is Information Risk a Determinant of Asset Returns*", Journal of Finance, Vol. 57 Issue 5, 2185-2221.

Easley, D., and O'Hara, M., (2004), "*Information and the cost of capital*", Journal of Finance, Vol. 59 Issue 4, 1553-1583.

Easley, D.; O'Hara, M.; Paperman, (1998), "*Financial Analysts and Information-Based Trade*", Journal of Financial Markets; Vol. 1, 175-201.

Forsund, F. R., (1975), "*The homothetic production function*", Swedish J. Economics 77, 234-244.

Fishburn, P. C., (1977), "*Mean-risk analysis with risk associated with below-target returns*", American Economic Review, 67.2, 116-126.

Gao, P., and Shrieves, R. E., (2002), "*Earnings Management and Executive Compensation: A Case of Overdose of Option and Underdose of Salary?"*, working paper SSRN.

Grossman, S., and O. Hart, (1980) "*Disclosure Laws and Take-over bids",* Journal of Finance, 35, 323-34.

Hull, J.C., (2011), "*Options, Futures and Other Derivatives*", Prentice-Hall.

Jung, W. and Y. Kwon, (1988), "*Disclosures when the market is unsure of information endowment of managers",* Journal of Accounting Research, 26, 146 - 153.

Keating C. and Shadwick W., (2002), "*A universal performance measure*", Journal of Performance Measurement, 59-84.

Kyle, A. (1985), "*Continuous auctions and insider trading*", Econometrica 53, 1315 - 1335.

Larcker, D. F., and Rusticus, T. O., (2010), "*On the Use of Instrumental Variables in Accounting Research*", Journal of Accounting and Economics. 49(3): 186–205.

Levy, H. (1992), "*Stochastic dominance and expected utility: survey and analysis*", Management Science, Volume 38.4, 555 - 593.

Markowitz, H. M. (1952), "*Portfolio Selection*", Journal of Finance 7.1, 77–91.

McNeil, A.J., Frey, R., and Embrechts, P., "*Quantitative Risk Management*", Princeton University Press, (2005)

Nawrocki, D. (1999), "*A brief history of downside risk*", The Journal of Investing, Vol 8.3, 9-25.

Ostaszewski, A., and M. Gietzmann, (2008), "*Value Creation with Dye's Disclosure Option: Optimal Risk-Shielding with an Upper Tailed Disclosure Strategy*", Review of Quantitative Finance and Accounting, Vol 31, Issue 1, 1- 27

C. S. Pedersen and S.E. Satchell, (2002), "*On the foundations of performance measures under asymmetric returns*", Quantitative Finance, Vol2 Issue 3, 217-223.

Penno, M.C., (1997), "*Information Quality and Voluntary Disclosure*", The Accounting Review, 72, 275-284.

Rogers, J.L., C.M. Schrand and R.E. Verrecchia, (2008) "*The impact of strategic disclosure on returns and return volatility: Evidence from the 'leverage effect'",* working paper Wharton School.

Roman, S. "*Introduction to Mathematical Finance: From Risk Management to Options*", Springer (2004).

Searle, S. R., Casella, G., McCulloch, C. E., (1992), "*Variance components",* Wiley.

Shin, H.S., (2006), "*Disclosure Risk and Price Drift",* Journal of Accounting Research, Vol. 44 Issue 2, 351-379.

Suijs, J., (2011), "*Asymmetric news and asset pricing*", working paper.

Van Buskirk, A., (2012), "*Disclosure frequency and information asymmetry*", Review of Quantitative Finance and Accounting, to appear.
Varian, H.R., (1992), *Microeconomic Analysis*, Norton, 3rd ed.
Verrecchia, R.E., (1990), "*Information Quality and Discretionary Disclosure*", Journal of Accounting & Economics, Vol. 12, Issue 4, 365-380.
Yermack, D., (1997) "*Good timing: CEO stock option awards and company news announcements*", Journal of Finance, Vol. 52.2, 449-476.

8 Appendix 1: Isotonic transformations

We are concerned here with a general noisy observation T of the signal X in the form $T = T(X, Y)$, where X and Y are independent random variables. Our interest below focuses on the regression function $\mu_X(t) = \mathbb{E}[X|T = t]$ and on the random variable, which we call the estimator of X

$$X^{\text{est}} = \mathbb{E}[X|T] = \mu_X(T),$$

which it is occasionally more convenient to abbreviate to S.

We now re-derive the Dye equation basing the analysis on the estimated value of X (given T) in place of X itself[23]. We find that the Dye equation holds for the estimator X^{est} of the true value. Let $L(x)$ be the inverse function of $\mu_X(t)$. The entire analysis rests on the following simple observation:

$$\Pr[X^{\text{est}} \leq x] = \Pr[\mu_X(T) \leq x] = \Pr[T \leq L(x)] = F_T(L(x)),$$

where F_T is the probability distribution function of T. The argument identifies that $F_S(x) := F_T(L(x))$ is the probability distribution function of X^{est}.

Theorem 8. (Isotonic Reduction Theorem)
Let $\mu_X(t)$ be the conditional expected value of of a random variable X given an observation t of $T(X, Y)$. Suppose that $\mu_X(t)$ is strictly increasing in t. Then the noisy signal cutoff $\hat{t}$ (cutoff for the disclosure of a noisy signal) reduces to the clean cutoff for the estimator distribution F_S with $S := \mathbb{E}[X|T]$, where $\hat{s} = \mu_X(\hat{t})$ solves the Dye equation

$$\frac{1-q}{q}(m_X - s) = H_S(s).$$

[23] This general argument includes the Penno (1997) analysis for the cutoff level of a normally distributed noisy signal as a special case.

and where $\mathbb{E}[S] = \mathbb{E}[X] = m_X$ *is also the mean of the estimator distribution* F_S *.*

Comment. The Theorem assumes a minimum specification regarding how the random variable T is related to X, namely that $\mu_X(t)$ is increasing in t. It may thus be applied both to an additive noise structure $X + Y$ like that of Penno, as well as our preferred multiplicative structure XY.

Proof. We begin by observing that the mean of the F_S distribution is m_X; this is immediate from the 'conditional mean formula' that

$$\mathbb{E}[S] = \mathbb{E}[\mathbb{E}[X|T]] = \mathbb{E}[X] = m_X,$$

so that the estimator $S = \mathbb{E}[X|T]$ is unbiased. To set this in context, start from the distribution F_T and put $L(s) := \mu_X^{-1}(s)$ so, since $\Pr[S \leq s] = \Pr[T \leq L(s)]$, one has that

$$\begin{aligned} \int_0^\infty s dF_S(s) &= \int_0^\infty s dF_T(L(s)) = \int_0^\infty \mu_X(t) dF_T(t) \\ &= \mathbb{E}[\mathbb{E}[X|T = t]] = \mathbb{E}[X], \end{aligned}$$

and this may, but need not in general, agree with $\mathbb{E}[T]$. For the noisy signal T the indifference equation for its non-disclosure cutoff $\gamma := \gamma_T$ triggering reads:

$$\mu_X(\gamma) = V(\gamma) := \mathbb{E}[X|ND(\gamma)] = \frac{(1-q)m_X + q\int_0^\gamma \mu_X(t)dF_T(t)}{(1-q) + qF_T(\gamma)}, \tag{27}$$

i.e. the signal disclosure γ yields the valuation $\mu_X(\gamma)$ whereas non-disclosure leads to $V(\gamma)$. The substitution

$$z = \mu_X(t)$$

gives

$$t = L(z).$$

From (27), integrating by parts, the indifference equation is equivalent to

$$
\begin{aligned}
\frac{1-q}{q}(\mu_X(\gamma)-m_X) &= \int_{t\leq\gamma}\mu_X(t)dF_T(t)-\mu_X(\gamma)F_T(\gamma) \\
&= [\mu_X(t)F_T(t)]^{\gamma}-\int_{t\leq\gamma}F_T(t)d\mu_X(t)-\mu_X(\gamma)F_T(\gamma) \\
&= -\int_{s\leq\mu_X(\gamma)}F_T(L(s))ds \\
&= -\int_{s\leq\mu_X(\gamma)}F_S(s)ds \\
&= -H_S(\mu_X(\gamma)),
\end{aligned}
$$

that is, the transformed trigger $s=\mu_X(\gamma)$ now solves the equilibrium indifference equation but using the hemi-mean of the F_S-distribution namely:

$$\frac{1-q}{q}(m_X-s)=H_S(s). \tag{28}$$

We will refer to this as the Dye equation adjusted for noise. We regard F_S and its hemi-mean as a replacement for F_X resulting from noise in the observed signal.

Appendix 2: Fishburn preferences

Explicit modeling assumptions need to be made concerning management preferences; in contrast to standard utility theory concerned with expected utility of outcomes, we need to consider preferences expressed by utility functions over pairs of parameters associated with distributions, rather like but more general than the Markowitz mean-variance pair.

Our concern with the hemi-mean functions may be connected to earlier literature in which modeling preferences reflecting differential concern with lower (versus upper) tail events also recognized the need to steer away from the simple mean-variance paradigm. This was indeed acknowledged by Markowitz (1952) explicitly in his seminal work, as he also proposed that semivariance be used to measure the risk of a portfolio, but did not exploit this[24]. Subsequently, a more general risk-measure (for below-target t risk)

[24] Possibly the reason he did not develop this is that the analysis of semivariances was not known to be tractable at that time.

was studied by Fishburn (1977), which he called the (α, t)-model, namely

$$F_\alpha(t) := \int_{x \leq t} (t-x)^\alpha dF(x), \qquad (\alpha > 0).$$

(with t an exogenous target), and showed it to be tractable. One sees immediately that the hemi-mean function, $H_X(t)$, appearing in the Dye equation is just F_1 for $F = F_X$, i.e. a lower partial moment of order $\alpha = 1$, cf. McNeil-Frey-Embrechts (2005); indeed, an integration by parts yields

$$H_X(t) = \int_{x \leq t} (t-x) dF_X(x).$$

For *arbitrary* distributions F, Fishburn studies preferences over F representable by a utility $U(\mu(F), \rho(F))$ over two parameters associated with F: the mean $\mu(F)$ and a risk-measure $\rho(F)$ of the general form $\int_{\leq t} \varphi(t-x)dF(x)$ for φ non-negative, non-decreasing with $\varphi(0) = 0$. The latter captures notions of 'riskiness' for outcomes x below the target t.

As both parameters are expectations under F, Fishburn's preference is a 'utility of expectations' rather than an expected utility in the von Neumann-Morgenstern sense.

In this connection we recall Fishburn's result, when specialized to the case $\varphi(t-x) = t-x$, that for such a dominance to be consistent with an expected utility, specifically taking the form $\mathbb{E}_F[v(X,t)]$ for some $v(x,t)$ increasing in y, with $v(t,t) = t$ and $v(t+1,t) = t+1$, it is necessary and sufficient for the existence of a constant $k = k(t) > 0$ such that

$$v(x,t;k) := \begin{cases} x, & x \geq t, \\ x - [k(t)(t-x)], & x \leq t. \end{cases}$$

Moreover,

$$\mathbb{E}_F[v(X,t)] = \mu(F) - k(t)\rho(F).$$

Call the utility v here the Fishburn *kinked* utility, to distinguish it from the utility $U(\mu(F), \rho(F))$ above. The kinked utility has left-sided slope[25] at t

[25] Since $v(t-1,t) = t-1-k$, one has

$$k(t) + 1 := \frac{v(t,t) - v(t-1,t)}{v(t+1,t) - v(t,t)},$$

and so $k(t)$ is independent of any scaling or shifting in the utility space.

greater than the right-sided slope, which recognizes the greater aversion to performance below target.

To summarize: in general, one can still analyze preferences with traditional expected utility analysis when risk is measured by a more general Fishburn risk measure $\rho(F)$, rather than 'symmetric' variance σ^2, provided one identifies the appropriate kink location t in the utility function.

Appendix 3: Optimized odds

To begin with we let a general concave utility function $U(y,z)$ describe managerial preference over the risk-shielding loss, assessed as in (8) by $z = m - \gamma$ when setting the cutoff at γ, and value enhancement, assessed by $y = H(\gamma)$. The context includes, as one interpretation, the case $m = m_X$ and $H = H_X$, where the random variable X models the terminal value of the firm, and, as another, the case $m = m_S$ and $H = H_S$, where S is a filtered signal, i.e. a random variable which is some well-defined transform of a noisy observation of X (for which see Section 2.1). Eventually we consider specific examples of U which include the utility of *Constant Elasticity of Substitution* (CES utility) and also its limiting version the *Cobb-Douglas* utility.

We continue the general analysis by supposing that the manager chooses among the points in the *opportunity set* defined by:

$$\Omega := \{(y,z) : 0 \leq z \leq m, \qquad y = H(m-z)\},$$

employing a general differentiable concave utility function $U(y,z)$, that is one with smooth convex contours. The set Ω is in general a convex curve, since $H''(\gamma) = f(\gamma) > 0$.

For our general discussion we assume there is a unique point $(\hat{y}, \hat{z})$ of Ω which corresponds to utility optimization under U. Subject to differentiability assumptions on U, the *optimal utility contour* $U(y,z) = U(\hat{y}, \hat{z})$ is tangential to Ω at $(\hat{y}, \hat{z})$. We analyze the significance of this observation.

In (y,z) co-ordinates the Dye cutoff condition (4) for expected indifference between non-disclosure and disclosure is:

$$(1-q)z = qy.$$

Corresponding to the optimal pair $(\hat{y}, \hat{z})$, from the preceding equation we define $\hat{\gamma}$ and $\hat{q}$ (or $\hat{\lambda}$) by

$$\hat{\lambda} = \frac{1-\hat{q}}{\hat{q}} := \frac{\hat{y}}{\hat{z}}, \text{ and } \hat{\gamma} := m - \hat{z}.$$

Thus we find that the common tangency of the optimal utility contour and the opportunity set (illustrated in Figure 2) implies a corresponding choice $\hat{\lambda}$ of *optimal odds*.

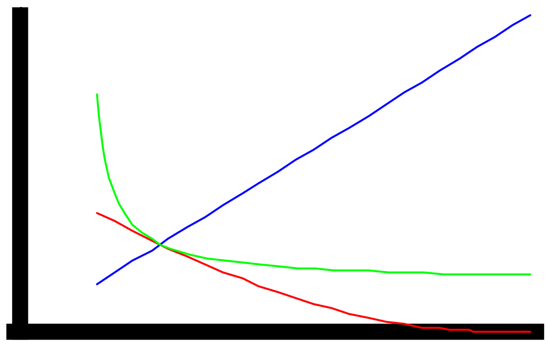

Figure 1. The arbitrage line (blue), the opportunity curve (red), and the tangential utility contour (green).

Our next aim is to characterize the optimal odds. We derive a condition, which we shall describe henceforth as *the optimal odds equation,* as follows. Note that the tangent slope at $(\hat{y}, \hat{z})$ along the optimal utility curve is given by the marginal rate of substitution $-U_z/U_y$. Put

$$\mathcal{M}_U(y, z) := \frac{U_z(y, z)}{U_y(y, z)}.$$

But, since $y = H(m - z)$, the slope of the opportunity curve is given by dy/dz, i.e. by $y' = -F(m - z)$. Common tangency thus requires that

$$\frac{d}{dz}H(m - z)\bigg|_{z=\hat{z}} = -\frac{U_z(\hat{y}, \hat{z})}{U_y(\hat{y}, \hat{z})}.$$

Since $F(\hat{\gamma}) = F(m - \hat{z})$, we obtain $F(\hat{\gamma}) = \mathcal{M}(\hat{\lambda}(m - \hat{\gamma}), m - \hat{\gamma})$. In summary, the model is fully endogenized by the choice of q via the pair of equations

$$\begin{aligned} \hat{\lambda}(m - \hat{\gamma}) &= H(\hat{\gamma}), \\ \mathcal{M}_U(\hat{\lambda}(m - \hat{\gamma}), m - \hat{\gamma}) &= F(\hat{\gamma}). \end{aligned}$$

Illustrative Examples. In the *Cobb-Douglas* case $U_{\text{C-D}}(y, z) = y^{\alpha} z^{\beta}$ and, setting $\kappa = \beta/\alpha$, the marginal rate of substitution of the utility is

$$\mathcal{M}_{\text{C-D}}(\hat{y}, \hat{z}) := \frac{\beta}{\alpha}\frac{\hat{y}}{\hat{z}} = \kappa\hat{\lambda}.$$

Similarly in the *Constant Elasticity of Substitution* case (CES case) $U_{\text{CES}}(y, z) = (\alpha y^{-\delta} + \beta z^{-\delta})^{-1/\delta}$, but now with $\kappa = (\beta/\alpha)^{1/(1+\delta)}$, we have

$$\mathcal{M}_{\text{CES}}(\hat{y}, \hat{z}) = \left(\kappa\frac{\hat{y}}{\hat{z}}\right)^{1+\delta} = \left(\kappa\hat{\lambda}\right)^{1+\delta}.$$

Thus $\delta = 0$ corresponds to the Cobb-Douglas case. In all these cases the optimal odds equation takes the elegant form:

$$F(\gamma) = (\kappa\lambda)^{1+\delta},$$

which in particular implies that

$$0 < \kappa\hat{\lambda} < 1, \text{ i.e. } 0 < \hat{\lambda} < \bar{\lambda} := 1/\kappa.$$

Consider the case when $\alpha = \beta$ (where equal weights are given to the two values being traded-off); then we have $0 < \hat{\lambda} < 1$, i.e. $1 - \hat{q} < \hat{q}$ and the manager prefers to be more informed than uninformed (i.e., $\hat{q} > 1/2$).

We now specialize our discussion to utilities $U(y, z)$ for which, as in the examples above, the marginal rate of substitution is *homogeneous of degree* 0 in y, z and so is a function of the ratio y/z, i.e. where $\mathcal{M}_U(y, z) = \mathcal{M}_U(y/z, 1)$. The class of such utilities is natural to use and quite wide as it comprises the homothetic utilities and those that are themselves homogeneous of degree zero – see Forsund (1975), and for further examples Appendix 5.

We thus write

$$u(\lambda) := \mathcal{M}_U(\lambda, 1). \tag{29}$$

If the U-contours are strictly concave functions of the form $y = y(z)$, then $u(\lambda)$ is increasing in λ. Similarly if the U-contours are strictly convex functions, then $u(\lambda)$ is decreasing in λ. If u is strictly increasing/decreasing (as in the illustrative examples), then the fully endogenized model permits a separation of the variables and leads to the simplified equations:

$$\begin{aligned} \hat{\lambda} &= H(\hat{\gamma})/(m - \hat{\gamma}), \\ \hat{\lambda} &= u^{-1}(F(\hat{\gamma})), \end{aligned}$$

since, the common tangency condition now reads $u(\hat{\lambda}) = F(\hat{\gamma})$.

In Section 3 we consider a family of distributions $F_X(.;\sigma)$ dependent on a parameter σ. Under such circumstances one writes $\gamma = \gamma(\lambda)$ or $\gamma = \gamma(\lambda, \sigma)$ in place of $\gamma(q)$ for the Dye cut-off defined (implicitly) by (4). For given σ the optimal odds in equilibrium, as above, are now defined implicitly as $\lambda = \hat{\lambda}(\sigma)$ via

$$u(\lambda) = F(\gamma(\lambda, \sigma), \sigma). \tag{30}$$

We have in mind that the manager's type is summarized by σ, and so one studies how the intensity of disclosure varies with the type. One may do so by first studying the dependence of the intensity of disclosure as a function of exogenously given odds λ.

By definition, the intensity of disclosure is given by

$$\tau(\lambda) := q(1 - F(\gamma(\lambda))).$$

Since $q = 1/(1 + \lambda)$, at equilibrium, one has

$$\tau(\lambda) = \frac{1 - u(\lambda)}{1 + \lambda}. \tag{31}$$

It is evident that if $u(\lambda)$ is an increasing function, then $\tau(\lambda)$ is decreasing (as a product of decreasing functions).

As a simple illustration, note the special case of the CES with $\delta = \kappa = 1$ above and $u(\lambda) = \lambda^2$, when we have $\tau = 1 - \lambda$, so that τ is linearly decreasing in λ.

The monotonicity here is in fact a quite general phenomenon, as was identified in Theorem 2 (Section 2.3 and Appendix 5 below).

Appendix 4: The structurally minimal model

We show that the implied utility of Section 2.3 for the structurally minimal model is $U(z, y) = (z^{-1} + y^{-1})^{-1}$.

To prove this, suppose as in Appendix 3 that the manager chooses among the points in the opportunity set defined by:

$$\{(y, z) : 0 \leq z \leq m, \qquad y = H(m - z)\},$$

employing a general differentiable utility function $U(y, z)$. In reduced form the Dye cutoff condition (4) for expected indifference between non-disclosure and disclosure is, as before:

$$(1-q)z = qy \qquad \text{or} \qquad (1-q)(y+z) = y, \tag{32}$$

which yields

$$q = q(y,z) = \frac{z}{y+z}. \tag{33}$$

Note that $q(y,z)$ is homogeneous of degree zero. So the manager's maximization objective is now $(1-q)\cdot z$, hence eliminating q via (32)

$$U(y,z) = (1-q)\cdot z = \frac{yz}{y+z} = (y^{-1}+z^{-1})^{-1}.$$

Compare Caplin and Nalebuff (1991), who consider generalized averages such as this harmonic one. Note also that the contour $U = c$ is a pair of rectangular hyperbolae with centre of symmetry at $y = z = c$, and is expressible as

$$(y-c)(z-c) = c^2.$$

Appendix 5: First Monotonicity Theorem

In this section we prove the Theorem 2 (First Monotonicity Theorem) of Section 2.3.

Assertion (i) is immediate from equation (31). (Compare the closing text of Appendix 3.)

As to assertion (ii), we note from equation (31) that

$$\begin{aligned}\tau' &= \frac{-u'(\lambda)(1+\lambda)-(1-u(\lambda))}{(1+\lambda)^2} \\ &= \frac{[u(\lambda)-\lambda u'(\lambda)]-u'(\lambda)-1}{(1+\lambda)^2} = \frac{\varphi(\lambda)}{(1+\lambda)^2}, \text{ say.}\end{aligned}$$

By assumption $\varphi(0) < 0$. Now

$$\varphi'(\lambda) = -(\lambda+1)u''(\lambda) < 0 \text{ for } u(\lambda) \text{ convex.}$$

So $\varphi'(\lambda) < 0$ i.e. $\varphi(\lambda)$ is decreasing for $\lambda > 0$. So $\varphi(\lambda) \le \varphi(0) < 0$ and hence $\tau'(\lambda) < 0$ for all $\lambda > 0$.

(iii) Here $\varphi(\bar{\lambda}) < 0$ and this time $\varphi(\lambda)$ is increasing so $\varphi(\lambda) \le \varphi(\bar{\lambda}) < 0$ for all $0 < \lambda < \bar{\lambda}$. Here again $\tau'(\lambda) < 0$ for all $0 < \lambda < \bar{\lambda}$.

Examples. (1) In the Cobb-Douglas case $u(\lambda) = \kappa\lambda$ (with $\kappa = \beta/\alpha$) and we have from (31) explicitly that

$$\tau = \frac{1 - \kappa\lambda}{1 + \lambda} = -\kappa + \frac{\kappa + 1}{1 + \lambda},$$

and so τ here is again decreasing with λ. (In fact $\varphi(\lambda) = -(1 + \kappa)$.)

2) In the CES case, taken in the form $U(y, z) := (\alpha y^{-\delta} + \beta z^{-\delta})^{-1/\delta}$ with $\delta \neq 0$, write $u(\kappa\lambda)$ in place of $u(\lambda)$ and $\varphi(\kappa\lambda)$ in place of $\varphi(\lambda)$ above. Then $\tau(\lambda) = (1 - u(\kappa\lambda))/(1 + \lambda)$ and so here from (31) we have

$$\begin{aligned} \tau' &= \frac{-\kappa u'(\kappa\lambda)(1 + \lambda) - (1 - u(\kappa\lambda))}{(1 + \lambda)^2} \\ &= \frac{[u(\kappa\lambda) - \kappa\lambda u'(\lambda)] - \kappa u'(\kappa\lambda) - 1}{(1 + \lambda)^2} = \frac{\varphi(\kappa\lambda)}{(1 + \lambda)^2}. \end{aligned}$$

We have

$$\begin{aligned} \varphi(\kappa\lambda) &= -\delta(\kappa\lambda)^{1+\delta} - (1 + \delta)\kappa(\kappa\lambda)^{\delta} - 1 \\ &= \kappa^{1+\delta}[-\delta\lambda^{1+\delta} - (1 + \delta)\lambda^{\delta}] - 1. \end{aligned}$$

Case a: with $\delta > 0$. (Here $u = \lambda^{1+\delta}$ which is convex.) We have $\varphi(0) < 0$. Furthermore, since $\delta > 0$ we have

$$\begin{aligned} \varphi'(\kappa\lambda)/\kappa^{\delta} &= -\delta(1 + \delta)\lambda^{\delta} - \delta(1 + \delta)\lambda^{\delta-1} \\ &= -\delta(1 + \delta)\lambda^{\delta-1}(1 + \lambda) < 0, \end{aligned}$$

thus $\varphi(\kappa\lambda)$ is decreasing for $\lambda > 0$ and so $\varphi(\kappa\lambda) < 0$ for $\lambda > 0$. Hence $\tau' < 0$, and so τ is decreasing in λ.

Case b: with $\delta < -1$. This is again the convex case. The argument above continues to hold.

Case c: $-1 < \delta < 0$. Here $\varphi(\kappa\lambda)$ is instead increasing for $\lambda > 0$; however, we also have $\varphi(1) < 0$, so $\varphi(\kappa\lambda) < 0$ in the range $0 < \kappa\lambda < 1$. Thus $\tau'(\lambda) < 0$ for $0 < \kappa\lambda < 1$. But the equilibrium condition $F(\gamma) = u(\kappa\lambda) = (\kappa\lambda)^{1+\delta}$ implies that $0 < \kappa\lambda < 1$ at equilibrium. Here again we have τ decreasing in λ.

9 Appendix 6: Physical vs. risk-neutral

Standardly, in the Black-Scholes model under the physical measure

$$S_t = S_0 e^{\mu t} e^{\sigma w_t^{\mathbb{P}} - \frac{1}{2}\sigma^2 t},$$

with μ the anticipated instantaneous return, σ the volatility and $w_t^{\mathbb{P}}$ the underlying $\mathbb{P}$-Wiener process, whereas under the equivalent risk-neutral measure $\mathbb{Q}$ with r the riskless rate

$$S_t = S_0 e^{rt} e^{\sigma w_t^{\mathbb{Q}} - \frac{1}{2}\sigma^2 t} \text{ with } w_t^{\mathbb{Q}} = \gamma t + w_t^{\mathbb{P}} \text{ and } \gamma = (\mu - r)/\sigma.$$

Here γ is the (Girsanov) risk-premium[26] and $w_t^{\mathbb{Q}}$ is a Wiener process under the risk-neutral measure $\mathbb{Q}$. Specializing to our context, setting $t = 1$, $r = 0$ (we use present-values in our analysis) and $S_0 = 1$ (in view of our renormalization of the initial equity to unity), given the ex-ante information one has under the physical measure that

$$X = S_1 = e^{\mu} e^{\sigma w_1^{\mathbb{P}} - \frac{1}{2}\sigma^2},$$

whereas under the risk-neutral measure that

$$X = S_1 = e^{\sigma w_1^{\mathbb{Q}} - \frac{1}{2}\sigma^2} \text{ with } w_1^{\mathbb{Q}} = \gamma t + w_1^{\mathbb{P}} \text{ and } \gamma = (\mu - r)/\sigma.$$

So the cost-of-capital discount is $e^{-\mu}$, since

$$\mathbb{E}^{\mathbb{Q}}[S_1] = 1 \text{ and } \mathbb{E}^{\mathbb{P}}[S_1] = e^{\mu}.$$

Appendix 7: Technicalities on stochastic dominance

Recall from Appendices 3 and 4 (cf. equation (29)) that $u(\lambda) := \mathcal{M}_U(\lambda, 1) = U_z(\lambda, 1)/U_y(\lambda, 1)$ is the marginal rate of substitution for the utility $U(z, y)$, assumed to be either homothetic or homogeneous of degree 0 in (y, z). We

[26] This is often derived by writing:

$$\frac{dS}{S} = \mu dt + \sigma dw_t = r dt + \sigma dw_t', \text{ where } w_t' = t(\mu - r)/\sigma + \sigma w_t.$$

assume that $u(\lambda)$ is increasing in λ and so has an inverse. Throughout we regard u as fixed. We will need to assume below that $u(\lambda)$ varies regularly enough.

Definition. Put

$$G^u(t), \text{ or } G(t) := \frac{u^{-1}(F(t))}{H(t)} \tag{34}$$

and say that the distribution F_1 *increasingly dominates* F_2 if $m_2 \leq m_1$ and

$$G_1(t) \geq G_2(t), \text{ for all } t \text{ with } 0 < t \leq m_2,$$

and also $H_1(m_2) < H_2(m_2)$. Thus we demand that

$$\frac{u^{-1}(F_1(t))}{H_1(t)} \geq \frac{u^{-1}(F_2(t))}{H_2(t)}.$$

Note the example $u(\lambda) = \lambda^2$, for which $u^{-1}(t) = t^{1/2}$, yields

$$G(t)^2 = \frac{F(t)}{H(t)^2} = -\frac{d}{dt}(H(t))^{-1}.$$

Here G is decreasing iff G^2 is decreasing iff $H(t)^{-1}$ is convex and so $-H(t)^{-1}$ is concave.

It will be helpful to recall the definition of ρ-concavity of a function g, due to Caplin and Nalebuff (1991), which includes log-concavity when $\rho = 0$ (i.e. $\log g$ is concave), and signifies that for $\rho > 0$ the function g^ρ is concave, whereas for $\rho < 0$ the function $(-g^\rho)$ is concave. We note that if g is ρ-concave then it is also ρ'-concave for all $\rho' < \rho$. In general if H is log-concave, then it is (-1)-concave.

Note that increasing dominance is a restricted form of dominance. Taking now $u(t) = t$, the mean-zero normal family directed by σ exhibits first order dominance to the left of the mean, as $\sigma_1 < \sigma_2$ implies that for $x < 0$

$$\frac{x}{\sigma_1} < \frac{x}{\sigma_2} \text{ and so } F_{\mathrm{N}}(x; 0, \sigma_1) = F_{\mathrm{N}}(\frac{x}{\sigma_1}; 0, 1) < F_{\mathrm{N}}(\frac{x}{\sigma_2}; 0, 1) = F_{\mathrm{N}}(x; 0, \sigma_2),$$

where $F_{\mathrm{N}}(t; \mu, \sigma)$ denotes the normal distribution function (with mean and standard deviation μ, σ). We recall that a twice differentiable function $g(t)$ is log-concave if

$$g''g - (g')^2 < 0. \tag{35}$$

(See e.g. Bergstrom and Bagnoli (2005), An (1998), or Caplin and Nalebuff (1991).)

Our first result shows that in cases of interest $G(t) = G^u(t)$, as defined above, is a decreasing function of t. The regularity condition quoted below is connected to the theory of regular variation, for which see Bingham et al. (1987), Section 1.8.

Proposition 1. *For $H(t)$ log-concave, $u(\lambda)$ strictly increasing and satisfying the regularity condition*

$$\frac{u(\lambda)}{\lambda u'(\lambda)} \leq 1, \tag{36}$$

in particular for $u(\lambda) = \lambda^\delta$ with $\delta \geq 1$, the function G^u is decreasing.

Proof. For $t = u(\lambda)$ with u strictly increasing put $v = u^{-1}$ so that $\lambda = v(t)$ and $u'(\lambda) = 1/v'(t)$ and note that

$$\frac{v(t)}{tv'(t)} = \frac{\lambda u'(\lambda)}{u(\lambda)} \geq 1. \tag{37}$$

But

$$\frac{d}{dt}\left(\frac{v(F(t))}{H(t)}\right) = \frac{v'(F)fH - v(F)F}{H^2} = \frac{v'(F)}{H^2}\left(Hf - F^2\frac{v(F)}{Fv'(F)}\right),$$

so in view of (35) and the growth condition (36) via (37)

$$\frac{d}{dt}\left(G^u(t)\right) \leq \frac{v'(F)}{H^2}\left(Hf - F^2\right) < 0.$$

Remark. When $u(\lambda) = \lambda^2$ the above condition reduces to the convexity of $H(t)^{-1}$ – i.e. the 'reciprocal convexity' of $H(t)$.

We have just seen that on $(0, m)$ if $H(t)$ is log-concave, then $G^u(t)$ is decreasing, so the condition that $G^u(t)$ be decreasing is a weakening of log-concavity.

Our next result shows that the assumption of a decreasing $G^u(t)$ represents only a mild form of regularity and does not need even as much as log-concavity.

Proposition 2. *Suppose the density function is such that the limit* $f(0+) := \lim_{t \searrow 0} f(t)$ *exists, then assuming the following limit exists we have:*

$$\lim_{t \searrow 0} \frac{f/F}{F/H} \leq 1.$$

If further $f'(0+)$ *exists, then we have in fact*

$$\lim_{t \searrow 0} \frac{f/F}{F/H} = 1.$$

Consequently, under the regularity condition (36), *the function*

$$G(t) = \frac{u^{-1}(F(t))}{H(t)}$$

is decreasing in an interval to the right of the origin.

We recall and prove Theorem 3 from Section 3.2.

Theorem 3. *Given two distributions, with log-concave hemi-means, with* F_1 *increasingly dominating* F_2, *the corresponding optimized Dye triggers satisfy*

$$\hat{\gamma}_1 > \hat{\gamma}_2.$$

Proof. The optimality condition requires $\gamma_F = \gamma(\hat{q})$ to satisfy

$$F_i(\gamma) = u(\hat{\lambda}) = u\left(\frac{H_i(\gamma)}{m_i - \gamma}\right),$$

or, with $v = u^{-1}$,

$$\frac{1}{m_i - \gamma} = v(F_i(\gamma))/H_i(\gamma).$$

or simply

$$G_i^v(\gamma) = \frac{1}{m_i - \gamma}.$$

Recalling that $m_2 < m_1$, we conclude, for $0 < t < m_1$, that

$$\frac{1}{m_2 - t} > \frac{1}{m_1 - t}.$$

The function $1/(m_1 - t)$ is strictly increasing, so intersects the graph of $G_2(t)$ earlier than it intersects the graph of $G_1(t)$. The intersection of G_1 with $1/(m_1 - t)$ is in turn later than its intersection with $1/(m_2 - t)$. The result is now clear. See the illustration below.

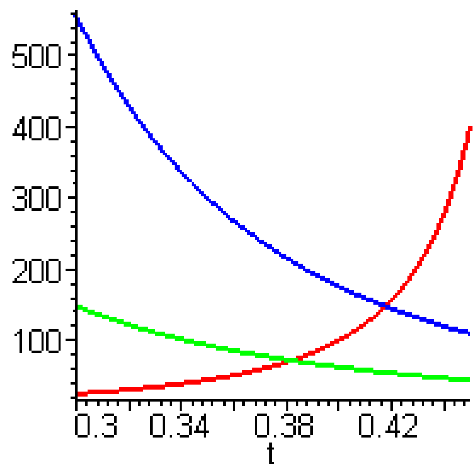

Figure 2. The blue and green curves exhibit the strict dominance. The red curve represents $(m - t)^2$ with $m = 1$.