Bayesian Pricing of News

Dmitry Livdan
Haas School of Business,
University of California, Berkeley
livdan@haas.berkeley.edu

Alexander Nezlobin*
Haas School of Business
University of California, Berkeley
nezlobin@haas.berkeley.edu

August 6, 2018

Abstract

This paper characterizes the equilibrium stock price reaction to arbitrarily distributed signals. This stock price reaction is shown to be proportional to the Fisher score of the news calculated under the risk-neutral probability measure. The expression for the Fisher score takes a particularly compact form for news whose distribution is obtained by conditioning a latent multivariate normal signal, a situation which often arises in partial-pooling equilibria. As an application of our analysis, we (i) characterize the stock price reaction to news whose arrival is content-dependent, and (ii) construct an equilibrium in a voluntary disclosure model with multidimensional information and risk averse investors.

Keywords: Persuasion games, voluntary disclosure, cost of capital, news pricing.

JEL Codes: D21, D82, D83, M41.

*Corresponding author.
1 Introduction

How should the stock market react to news about a firm? While a vast body of financial economics literature is devoted to pricing of financial instruments with non-Gaussian payoffs, much less is known about the pricing of non-Gaussian news. In this paper, we consider a model with fairly standard assumptions about the underlying security and market participants: the terminal payoff is normally distributed and investors have constant absolute risk aversion (CARA) preferences. Yet, in contrast to the earlier literature, we allow for news that are described by arbitrary distributions conditional on the firm’s true payoff.

We start by showing that the stock price reaction to a signal (piece of news) is proportional to the Fisher score of that signal under the risk-neutral probability measure. In our model, the relevant concept of the Fisher score captures how sensitive the marginal log-likelihood of the news is to the prior mean of the payoff. The stock price reaction to news is proportional to this quantity albeit calculated under the risk-neutral probability measure, which we show can be obtained by replacing the prior mean of the payoff with the pre-news price of the stock. This result holds for a large class of signal distributions, discrete and continuous, with essentially the only requirement being that the marginal log-likelihood of a given signal realization should be differentiable in the prior mean of the payoff.

Our result encompasses two news-pricing formulas obtained in the earlier literature. First, it applies in settings where signals and payoffs are jointly normally distributed, a modeling assumption ubiquitous in financial economics studies. Second, we consider a voluntary disclosure game based on Dye (1985), in which the manager may or may not be endowed with a disclosable signal. The equilibrium of such models is usually of the partial pooling type: an informed manager discloses the signal only if it exceeds a certain threshold; if the signal realization is low, the manager pools with the uninformed types. Consequently, the posterior distribution of the payoff conditional on non-disclosure is not normal even when the original signal and the prior are. Dye and Hughes (2018) provide a first characterization of the risk averse pricing of such voluntary disclosures, and we show that their news pricing equation is also nested in our general result.

Having established the link between the stock price reaction to a signal and its Fisher score, we proceed by deriving a simple formula for the Fisher score for a set of signal distributions that often arises in partial pooling equilibria. Specifically, we consider signals that

---


2 The resulting distribution, in fact, belongs to the skew normal (or power normal) family.
indicate that a latent multidimensional normal variable belongs to some subset of $\mathbb{R}^N$. Such events, for instance, occur when news arrival is correlated with its content and/or the firm’s true value. In practice, different information channels may have different propensity for disclosing the same information. For instance, the firm’s managers are more likely to release news that are favorable to the firm, and analysts might prefer to drop coverage instead of publishing a very negative report. In contrast, under the U.S. Generally Accepted Accounting Principles (GAAP), the firm’s financial statements are required, in many situations, to convey bad news in a more timely fashion than the good news. In such scenarios, investors face the problem of pricing the firm based on some reported signals and the understanding that the realizations of some other signals may not have cleared their respective sources’ disclosure screens. In our second result, we show that the Fisher score of a piece of news obtained by conditioning a latent multidimensional normal variable can be expressed as a linear function of the expected surprise in that variable, with the expectation taken over its possible locations consistent with the observed information.

We further illustrate the application of our results with two novel examples. First, we consider a setting where the arrival of a piece of news is determined by a probit model and can be correlated with the signal and noise components of the news itself. This setting is reminiscent of Heckman’s (1977) model, and thus one might expect the price reaction to such news to incorporate a “news arrival” correction. There are, however, two important differences between our setting and that in Heckman (1977). First, the nature of the economic problem in our setting is Bayesian, whereas the statistical problem in Heckman (1977) is frequentist. Second, in characterizing the post-news price of the firm one needs to determine not only the posterior mean of the firm’s payoff but also the risk premium that risk averse investors will require for holding the stock. While the conditional distribution of the payoff based on such news is not normal, we provide a closed-form expression for the news arrival correction. We show that the price reaction to news with content-dependent arrival is convex (concave) in the news content if the likelihood of arrival depends more strongly on the signal (noise) component of the news. This result links the stock price reaction to public news with the source’s propensity to release such news and provides conditions under which the reaction is convex or concave.

As a second application of our theory, we characterize an equilibrium in a voluntary disclosure model based on Dye (1985) with multidimensional information and risk averse investors. To date, voluntary disclosure models with uncertain and multidimensional informational endowments have been solved under the assumption that at most two signals can
be disclosed by the manager (see, e.g., Fishman and Hagerty 1990, Pae 2005, Guttman et al., 2014). The first paper incorporating risk averse investors in this class of models is Dye and Hughes (2018), but their analysis focuses on the setting with one voluntarily disclosed signal. We construct an equilibrium with risk averse investors and any arbitrary number of signals. In this equilibrium, the pricing function incorporates the disclosed signals exactly as they would have been priced in the absence of reporting discretion and attaches a fixed penalty for the total number of undisclosed signals (either withheld or unobserved by the manager). We show that the manager’s incentives to disclose additional signals unequivocally decrease in the realizations of the top signals; i.e., firms with great news to share are less likely to disclose the more ordinary news. Disclosure incentives also generally strengthen in the degree of investors’ risk aversion.

The rest of the paper is organized as follows. In the remainder of this Introduction, we discuss the related literature. Section 2 presents the general model, demonstrates the link between the stock reaction to an event and its Fisher score, and provides two examples from the earlier studies that illustrate this relation. Section 3 characterizes the Fisher score for events obtained by conditioning a latent multivariate normal signal. Also in that section, we study the price reaction to news with content-dependent arrival. Section 4 presents an equilibrium in a voluntary disclosure game with multidimensional information. Section 5 concludes.

1.1 Literature Overview

Our paper contributes to the growing body of literature in information economics that employs Bayesian analysis with a large set of signal distributions (see, e.g., Kamenica and Gentzkow 2011, Kamenica 2017). For instance, in our first result, we solve for the price reaction to a piece of news whose distribution conditional on the true payoff is effectively arbitrary. One of the main applications of our results is cast in a voluntary disclosure setting based on Dye (1985) and Dye and Hughes (2018). Specifically, we are interested in a multi-signal extension of this model similar to, e.g., Fishman and Hagerty (1990), Pae (2005), Acharya et al. (2011), Guttman et al. (2014). One similarity between such extensions of the voluntary disclosure setting and the literature on Bayesian persuasion is that both

---

3 Models with voluntary disclosure of multiple pieces of news are also studied in Shin (2003) and Shin (2006), yet in those papers the firm’s payoff is modeled as an outcome of multiple independent projects. As such, the nature of the analysis in those papers is not Bayesian as the news carry information about unrelated parts of the firm’s value.
approaches enrich the set of choices available to the sender: in the former approach, the sender can choose which pieces of information to disclose, whereas, in the latter, the sender can choose the distribution of the signal. The two classes of models are, however, usually different in the assumed possibility of the sender’s commitment: in the setting studied in this paper, the sender cannot commit to any particular disclosure policy (and if she could, she would prefer that of full disclosure).

In the rational expectations setting, there is also a growing literature that aims to relax the standard CARA-Normal assumptions. For instance, Breon-Drish (2015) studies a class of models, in which the posterior distribution of the payoff given a signal realization is assumed to belong to the exponential family. There are multiple differences between his analysis and that in our paper. First, while we allow for information asymmetry between the sender and the market, all market participants in our model are symmetrically informed, whereas Breon-Drish (2015) considers heterogeneously informed investors. Second, while our assumption about the prior payoff distribution is more restrictive (we assume it is normal), we allow for a broader set of signal distributions. For instance, when we study the price reaction to news with content-dependent arrival, the posterior distribution of the payoff given the signal does not belong to the exponential- but rather to the curved exponential family. As a consequence, investors’ demand for the firm’s stock after such news is shown to be either convex or concave in the news content. In Breon-Drish (2015), informed investors’ demand functions are linear in their respective signal’s realizations. In fact, we show that, in voluntary disclosure settings, the posterior distribution of the payoff given non-disclosure of some information will generally not belong to the exponential family even if the distribution of the unreported information is jointly normal with the prior.

Finally, our paper is related to several studies in econometrics and statistics. The probit model of news arrival builds on Heckman (1977), but, similar to van Hasselt (2011), the news arrival correction arises in a Bayesian setting. The idea of considering distributions that are obtained by conditioning a latent multivariate normal variable has been also actively explored in the statistics literature; see, e.g., Azzalini and Dalla Valle (1996), Lee and McLachlan (2013).

2 General Model

Let $V$ be the main variable of interest, e.g., the firm’s terminal cash flow. Throughout the paper, we make the following two assumptions. First, the prior distribution of $V$ is normal
with mean $M$ and variance $\sigma^2_V$:

$$V \sim \mathcal{N}(M, \sigma^2_V).$$

(1)

Second, the preferences of the representative investor trading the firm’s stock are given by the constant absolute risk aversion (CARA) utility function:

$$U(\omega) = -\frac{1}{\rho} \exp\{-\rho \omega\},$$

(2)

where $\omega$ is the investor’s terminal wealth, and $\rho > 0$ is the coefficient of risk aversion. The case of risk-neutrality obtains when $\rho \to 0$. While assumptions (1) and (2) are not innocuous, they are ubiquitous in the earlier literature on information disclosure. We note that in our setting with symmetric information, there is no loss of generality in modeling the stock market as a single representative investor.\(^4\)

The firm’s stock is traded twice – before and after a signal about $V$ is released to the market. Our main goal is to describe the evolution of the firm’s stock price under sufficiently general assumptions about the structure of the signal, which will be denoted by $X$. Let $g(X|V)$ denote the conditional likelihood of $X$ given $V$. The function $g(X|V)$ can be a probability density function if $X$ is a continuous random variable or a probability mass function if $X$ takes on discrete outcomes. At this stage, we also allow for single- and multidimensional signals. Finally, $X$ can be any event whose probability (or its density) is measurable for all values of $V$.

A risk-free asset yielding a return of zero is in infinite supply. Without loss of generality, the initial wealth of the representative investor is normalized to zero, and the total supply of the firm’s stock is normalized to one share at each trading date. The sequence of events is as follows: at date 1, the firm’s stock is traded based on the prior information about the payoff; at date 2, signal $X$ is released; at date 3, the firm’s stock is traded by the same representative investor again; and, at date 4, payoff $V$ is revealed and consumed by the investor. At both trading dates, $t = 1$ and 3, the price of the firm, $P_t$, is set so that the representative investor is willing to hold exactly one share of the firm’s stock.\(^5\) Formally, $P_t$

\[^4\]With some additional algebra, our analysis can be extended to allow for multiple firms with correlated payoffs; we discuss such a setting later in this section.

\[^5\]It is not crucial for our analysis that the stock holdings at date 1 are at their equilibrium levels. However, introducing the equilibrium price $P_1$ is useful in demonstrating the intuition behind the pricing formula at date 3.
must satisfy the following equation:

$$1 = \arg \max_{\alpha} \mathbb{E}_t \left[ -\frac{1}{\rho} \exp \left\{ -\rho (\alpha V - \alpha P_t) \right\} \right],$$

(3)

where $\mathbb{E}_t [\cdot]$ is the expectation operator conditional on the information available at date $t$, and $\alpha$ is the share of the firm that the representative investor chooses to hold. It is well-known and straightforward to check that $P_1$ is given by

$$P_1 = M - \rho \sigma^2_V.$$  

(4)

We will refer to the expression above as the firm’s pre-news (or pre-disclosure) price.

In our first main result, we provide a simple expression for price $P_3$, which prevails in the market once signal $X$ is released. Let $f(t; m, \sigma^2)$ and $F(t; m, \sigma^2)$ denote, respectively, the normal probability density (pdf) and cumulative distribution (cdf) functions with mean $m$ and variance $\sigma^2$. If $m$ and $\sigma^2$ are omitted, they are understood to be zero and one, respectively. The posterior distribution of $V|X$, $\psi(V|X)$, can now be written as:

$$\psi(V|X) = \frac{\left( \mathbb{P}\{X\} \right)^{-1}}{\text{Normalizing constant}} f(V; M, \sigma^2_V) g(X|V),$$

(5)

where $\mathbb{P}\{X\}$ is the marginal likelihood of $X$, given by:

$$\mathbb{P}\{X\} = \int_{-\infty}^{\infty} f(V; M, \sigma^2_V) g(X|V) \, dV.$$  

(6)

Note that $\mathbb{P}\{X\}$ depends on the prior mean, $M$, and we will write it as $\mathbb{P}\{X; M\}$ when we need to make this dependency explicit. It is important to mention that, in our notation, the probability operator $\mathbb{P}\{\cdot\}$ measures a priori probabilities, i.e., it treats $V$ as a random variable distributed according to its prior; conditioning on the true value of $V$ will be specified explicitly, where appropriate, as $\mathbb{P}\{\cdot|V\}$.

Using the posterior density in (5), observe that the expected utility of the representative investor who owns $\alpha$ shares of stock at date 3 is proportional, up to a factor independent of $\alpha$, to:

$$\mathbb{E}\left[ -\frac{1}{\rho} \exp \left\{ -\rho (\alpha V - \alpha P_3) \right\} | X \right] \sim -\frac{1}{\rho} \int_{-\infty}^{\infty} \exp \left\{ -\rho (\alpha V - \alpha P_3) \right\} f(V; M, \sigma^2_V) g(X|V) \, dV.$$  

6
Completing the square of $V$ in the exponent under the integral, the expression above simplifies to:

$$
E \left[ \frac{-1}{\rho} \exp \left\{ -\rho (\alpha V - \alpha P_3) \right\} \mid X \right] \sim \frac{-1}{\rho} \exp \left\{ -\rho \left( \alpha M - \frac{\rho \alpha^2}{2} \sigma_V^2 - \alpha P_3 \right) \right\} \int_{-\infty}^{\infty} f(V; M - \rho \alpha \sigma_V^2, \sigma_V^2) g(X \mid V) dV. \quad (7)
$$

Now note that the integral in the right-hand side above is of the exactly the same form as (6), and, therefore, we have:

$$
E \left[ \frac{-1}{\rho} \exp \left\{ -\rho (\alpha V - \alpha P_3) \right\} \mid X \right] \sim \frac{-1}{\rho} \exp \left\{ -\rho \left( \alpha M - \frac{\rho \alpha^2}{2} \sigma_V^2 - \alpha P_3 \right) \right\} \mathbb{P} \{ X; M - \rho \alpha \sigma_V^2 \}. \quad (8)
$$

We are now ready to calculate $P_3$. Maximizing (8) with respect to $\alpha$ yields the following first-order condition:

$$
\exp \left\{ -\rho \left( \alpha^* M - \frac{\rho (\alpha^*)^2}{2} \sigma_V^2 - \alpha^* P_3 \right) \right\} \left( (M - \rho \alpha^* \sigma_V^2 - P_3) \mathbb{P} \{ X; m \} + \sigma_V^2 \frac{\partial \mathbb{P} \{ X; m \}}{\partial m} \right) \bigg|_{m = M - \rho \alpha^* \sigma_V^2} = 0. \quad (9)
$$

The condition above is satisfied at $\alpha^* = 1$ if and only if\(^7\)

$$
P_3 = M - \rho \sigma_V^2 + \sigma_V^2 \frac{\partial \log \mathbb{P} \{ X; m \}}{\partial m} \bigg|_{m = M - \rho \sigma_V^2}. \quad (10)
$$

Recall that $M - \rho \sigma_V^2$ is the pre-disclosure price, $P_1$, and the derivative of the log-likelihood of some event with respect to a model parameter is known as its Fisher score.\(^8\) The characterization in (10) thus relates the stock price reaction to event $X$, measured as $\Delta P(X) \equiv P_3 - P_1$, to the a priori Fisher score of that event. We illustrate this result with several examples

\(^6\)Recall that $f(V; M, \sigma_V^2) = \frac{1}{\sqrt{2\pi \sigma_V^2}} \exp \left\{ -\frac{(V - M)^2}{2\sigma_V^2} \right\}$.

\(^7\)Note that if $\chi(\cdot)$ is a positive function, then $(\log \chi(t))' = \chi'(t) / \chi(t)$.

\(^8\)Here, we abuse the standard terminology but only slightly. Usually, the Fisher score is defined as the derivative of the log-likelihood function of the data with respect to the unobserved parameter, which, in our case, would be given by be the derivative of the log of density in (5) with respect to $V$. In equation (10), we take the derivative of the logarithm of the marginal likelihood function, $\mathbb{P} \{ X; m \}$, with respect to one of the parameters of the prior distribution. Nonetheless, as we will show later, our notion of Fisher score has properties very similar to the traditionally defined one, so we will use this term throughout the paper, sometimes qualifying it as the “a priori” Fisher score.
below; at this stage, it is useful to note that our derivation of equation (10) has only relied on the existence of several integrals involving signal \( X \). Specifically, we assume that the following expressions, viewed as functions of \( M \),

\[
P\{X\}, \mathbb{E}[V|X], \mathbb{E}[V^2|X]
\]

exist are everywhere continuous in \( M \). (11)

To confirm that the stock market at date 3 clears at price \( P_3 \), it remains to check the second-order condition for the problem of maximizing (8) with respect to \( \alpha \). We do so in the Appendix, thus proving the following Proposition.

**Proposition 1** Assume that event \( X \) satisfies conditions (11). Then, the stock price reaction to event \( X \) is given by the product of the prior variance of \( V \) and the a priori Fisher score of \( X \) evaluated with the prior mean set equal to the pre-disclosure price:

\[
\Delta P(X) = \sigma^2_V \frac{\partial \log \mathbb{P}\{X; m\}}{\partial m} \bigg|_{m = P_1}.
\]

(12)

To shed further light on the expression in (12), consider the case of risk-neutrality, \( \rho \to 0 \). Then, the stock price reaction is simply given by:

\[
\Delta P(X) = \sigma^2_V \frac{\partial \log \mathbb{P}\{X\}}{\partial M}.
\]

The equation above can be used to derive the expected price at date 3:

\[
\mathbb{E}_1[P_3] = P_1 + \sigma^2_V \mathbb{E}_1\left[\frac{\partial \log \mathbb{P}\{X\}}{\partial M}\right] = P_1 + \sigma^2_V \int \frac{\partial \mathbb{P}\{X\}}{\partial M} \mathbb{P}\{X\} dX
\]

\[
= P_1 + \sigma^2_V \frac{\partial}{\partial M} \int \mathbb{P}\{X\} dX = P_1,
\]

(13)

where the last equality obtains since the probability density function integrates to one, regardless of \( M \). Therefore, when investors are risk-neutral, the expected price at date 3 is exactly equal to the price at date 1, and this observation is, of course, nothing more than the law of iterated expectations.

\(^9\)In the expression below, we assume that \( X \) is a continuous random variable and, therefore, \( \mathbb{P}\{X\} \) is its marginal density. The expression we derive is a well-known property of the Fisher score.
Note that the argument above does not apply when $\rho \neq 0$ and $P_1 \neq M$. One can still write down an analogue of (13), but this sequence of equalities will only hold when the future states of the world are weighted by probabilities $\mathbb{P}\{X; P_1\}$ rather than $\mathbb{P}\{X; M\}$ in the calculation of the expectation operator $\mathbb{E}_1[\cdot]$. Hence, the substitution $m = P_1$ in equation (12) effectively converts the expression for the Fisher score to the risk-neutral probability measure.

In fact, it follows from the proof of Proposition 1 that, in the general case, the stock price reaction in (12) can also be written as:\footnote{To see this, recall that
\[
\frac{\partial \log \mathbb{P}\{X;m\}}{\partial m} = \frac{\partial \mathbb{P}\{X;m\}}{\partial m} / \mathbb{P}\{X;m\}
\]
and apply equation (45). While the expression in (14) is more compact than (12), as it will become clear later, it is less useful for practical purposes. The reason for this is that it is often easier to work with the direct density of $X|V$ rather than with the density of $V|X$, which needs to be calculated using the Bayes rule.}

\[
\mathbb{E}_1[ V - m | X; m ] \bigg|_{m = P_1}, \quad (14)
\]

where $\mathbb{E}[V|X;m]$ denotes the expectation of $V$ given $X$, calculated under the assumption that the unconditional mean of $V$ is equal to $m$. If $\rho > 0$, then $P_1 \neq M$, and the physical measure is not risk-neutral; then, for instance,

\[
\mathbb{E}_1 \left[ \mathbb{E}[ V - m | X; m ] \bigg|_{m = P_1} \right] \neq \mathbb{E}_1 \left[ \mathbb{E}[ V - m | X; m ] \bigg|_{m = P_1} \right] = 0,
\]

and the expected price at date 3 will depend on the probabilistic structure of signal $X$.

We conclude our discussion of the risk-neutral case by calculating the variance of $P_3$ from the vantage point of date 1. Letting $\mathbb{V}_t[\cdot]$ denote the variance operator conditional on date $t$ information, we have:

\[
\mathbb{V}_1[P_3] = \mathbb{V}_1 \left[ \mathbb{E}_1 \left[ \frac{\partial \log \mathbb{P}\{X\}}{\partial M} \right] \right] = \mathbb{V}_1 \left[ \mathbb{E}_1 \left[ - \frac{\partial^2 \log \mathbb{P}\{X\}}{\partial M^2} \right] \right], \quad (15)
\]

where the variance operator in the second term is the a priori Fisher information of event $X$ with respect to $M$, and the last equation is a well-known property of Fisher information. We will return to this expression when we discuss the special case of normally distributed signals below; note, however, that our derivation of (15) does not rely on the joint normality
of \( X \) and \( M \).

Our result in Proposition 1 can be generalized to a setting with multiple firms with correlated payoffs. Let \( V \) denote the vector of normally distributed payoffs with mean \( m \) and covariance matrix \( \Sigma_V \). Furthermore, let \( P_1 \) and \( P_3 \) denote the vectors of prices at dates 1 and 3. Lastly, let \( \nabla_m \) denote the gradient operator with respect to \( m \).

**Observation 1** In the setting with multiple firms, the price reaction to event \( X \) is given by:

\[
P_3 - P_1 = \Sigma_V \nabla_m \log \mathbb{P}\{X; m\} \mid m = P_1.
\]

The proof of Observation 1 follows essentially the same steps as that of Proposition 1 above. Let \( \alpha \) denote the vector of shareholdings of the representative investor at date 3, and let \( T \) be the transpose operator. Then, completing the square of \( V \) in the calculation of the investor’s expected utility yields:

\[
\mathbb{E}\left[-\frac{1}{\rho} \exp \left\{ -\rho \left( \alpha^T V - \alpha^T P_3 \right) \right\} \mid X \right] \sim -\frac{1}{\rho} \exp \left\{ -\rho \left( \alpha^T M - \alpha^T P_3 - \frac{\rho}{2} \alpha^T \Sigma_V \alpha \right) \right\} \int_{-\infty}^{\infty} f_N (V; M - \rho \Sigma_V \alpha, \Sigma_V) g(X \mid V) \, dV,
\]

which is analogous to equation (7). The market clears at date 3 if the optimal \( \alpha \) is equal to \( 1_{N \times 1} \), the vector of ones. It is straightforward to check that the first-order condition for market clearing is:

\[
(M - P_3 - \rho \Sigma_V 1_{N \times 1}) \mathbb{P}\{X; M - \rho \Sigma_V 1_{N \times 1}\} + \Sigma_V (\nabla_m \mathbb{P}\{X; m\}) \mid m = M - \rho \Sigma_V 1_{N \times 1} = 0.
\]

Observation 1 then follows by noting that

\[
P_1 = M - \rho \Sigma_V 1_{N \times 1}.
\]

We have thus shown that even in the setting with multiple risky securities, the stock price reaction event \( X \) is given by its Fisher score calculated under the risk-neutral measure. For notational simplicity, we revert to the single firm scenario for the remainder of this paper.
2.1 Examples

In this section, we further illustrate the formula in (12) with two examples.

Example 1. Normally distributed signals. Assume that the market receives one signal about \( V, Y \), which is distributed normally with mean \( V \) and variance \( \sigma^2 \):

\[
Y \propto N(V, \sigma^2).
\]

Then, the a priori distribution of \( Y \) (its marginal likelihood) is normal with mean \( M \) and variance \( \sigma^2 + \sigma_e^2 \). It follows that

\[
P \{ Y; m \} = f(Y; m, \sigma^2 + \sigma_e^2),
\]

and

\[
\frac{\partial \log P \{ Y; m \}}{\partial m} = \frac{Y - m}{\sigma^2 + \sigma_e^2}.
\]  

(16)

Equation (12) now becomes:

\[
P_3 - P_1 = \sigma^2 \frac{Y - P_1}{\sigma^2 + \sigma_e^2},
\]

or, equivalently,

\[
P_3 = \frac{\sigma^2}{\sigma^2 + \sigma_e^2} P_1 + \frac{\sigma^2}{\sigma^2 + \sigma_e^2} Y
\]

\[
= \frac{\sigma^2}{\sigma^2 + \sigma_e^2} M + \frac{\sigma^2}{\sigma^2 + \sigma_e^2} Y
\]

\[
- \rho \frac{\sigma^2 \sigma^2}{\sigma^2 + \sigma_e^2} \frac{\sigma^2}{\sigma^2 + \sigma_e^2},
\]

which is a well-known formula for pricing a normally distributed payoff based on a normally distributed signal.

The expression in (17) provides more intuition on the substitution \( m = P_1 \) in Proposition 1. Consider the following question: when should there be no price reaction to signal \( Y \)? If \( \rho \to 0 \), then \( P_1 = P_3 \) whenever \( Y = M \), i.e., the signal is equal to its unconditional mean. When investors are risk averse, however, the realization of \( Y \) equal to \( M \) leads to an increase in price, as equation (17) suggests. This is because while \( Y = M \) is neutral news in terms of updating the posterior mean of \( V \), the mere fact that additional information has arrived leads to a decrease in the posterior variance and, accordingly, the risk premium that investors require for holding the stock. Equation (17) shows that price-neutral news are then given
by $Y = P_1 < M$.

The jointly normal case also illustrates why equation (13) only holds in the risk-neutral setting. It is straightforward to see that equation (17) implies that

$$
\mathbb{E}_1 [P_3 - P_1] = \frac{\rho \sigma_Y^4}{\sigma_V^2 + \sigma_\varepsilon^2} > 0,
$$

i.e., in expectation, the firm’s price increases after disclosure. This might suggest that investors face an arbitrage opportunity: instead of buying the stock at date 1, they can wait until date 3 and pay (on average) less for the same number of shares. In fact, it is easy to check that for equilibrium values of $P_1$ and $P_3$, risk averse investors are exactly indifferent, at date 1, between a higher but certain price $P_1$ and a lower in expectation yet risky price $P_3$.

To measure the riskiness of $P_3$, assume again that $\rho \to 0$ and consider equations (15) and (16). We have:

$$
\mathbb{V}_1 [P_3] = \sigma_Y^4 \mathbb{E}_1 \left[ - \frac{\partial^2 \log \mathbb{P} \{ X \}}{\partial M^2} \right] = \frac{\sigma_Y^4}{\sigma_V^2 + \sigma_\varepsilon^2}.
$$

Therefore, the expected increase in the stock price in (18) is equal to $\rho \mathbb{V}_1 [P_3]$, as should be expected from our discussion above.

**Example 2. Voluntary Disclosure of One Signal.** As in Dye (1985) and Jung and Kwon (1988), assume that the firm’s manager has a private signal about $V$ with some publicly known probability $0 \leq \lambda \leq 1$, and with probability $1 - \lambda$ the manager is uninformed. As in our previous example, this signal, $Y$, is distributed normally with mean $V$ and variance $\sigma_\varepsilon^2$:

$$
Y \sim \mathcal{N} (V, \sigma_\varepsilon^2).
$$

The manager, if informed, can disclose the signal truthfully and costlessly to the market participants, yet if the manager is uninformed, then the absence of private information cannot be credibly communicated to outside parties. As is common in this literature, assume that the goal of the manager is to maximize the firm’s post-disclosure stock price (see, e.g., Shin 2003; Guttman et al. 2014; Dye and Hughes, 2018).

In this class of models, the equilibrium disclosure strategy is usually characterized by

---

11 The assumption that $\rho \to 0$ is, in fact, of no importance in the jointly normal setting for the calculation of $\mathbb{V}_1 [P_3]$. [3225448]
a threshold, $C$, with the manager only disclosing the realizations of $Y$ exceeding $C$. An important step in analyzing the manager’s problem is the derivation of the non-disclosure price, i.e., the price that prevails in the market if $Y$ is not received (which happens with probability $1 - \lambda$) or not disclosed (which happens when signal $Y \leq C$ is received). As in our previous example, the a priori distribution of $Y$ is normal with mean $M$ and variance $\sigma_Y^2 + \sigma_\epsilon^2$. Thus, the total marginal likelihood of non-disclosure, an event labeled ’$ND$’ below, can be written as:

$$ P\{ND\} = 1 - \lambda + \lambda F(C; M, \sigma_Y^2 + \sigma_\epsilon^2). $$

Then, it follows that

$$ \frac{\partial \log P\{ND; m\}}{\partial m} = -\frac{\lambda f(C; m, \sigma_Y^2 + \sigma_\epsilon^2)}{1 - \lambda + \lambda F(C; m, \sigma_Y^2 + \sigma_\epsilon^2)}. $$

By equation (12), we have:

$$ \Delta P(ND) = -\sigma_Y^2 \frac{\lambda f(C; P_1, \sigma_Y^2 + \sigma_\epsilon^2)}{1 - \lambda + \lambda F(C; P_1, \sigma_Y^2 + \sigma_\epsilon^2)} $$

$$ = -\frac{\sigma_\epsilon^2}{\sqrt{\sigma_Y^2 + \sigma_\epsilon^2}} \frac{\lambda f(z)}{1 - \lambda + \lambda F(z)}, $$

where

$$ z \equiv \frac{C - P_1}{\sqrt{\sigma_Y^2 + \sigma_\epsilon^2}}, $$

and, as before, $P_1 = M - \rho \sigma_Y^2$.

Equation (19) characterizes the non-disclosure price in the context of Dye’s (1985) model with risk averse investors; it was first obtained in Dye and Hughes (2018).\textsuperscript{12} That paper also derives the condition that determines the equilibrium disclosure threshold $C^*$. Intuitively, a manager who receives signal $Y$ equal to $C^*$ must be indifferent between disclosure and non-disclosure. Using equation (17) in our Example 1, recall that if signal $Y$ is disclosed, then the corresponding price reaction can be written as:

$$ \sigma_Y^2 \frac{Y - P_1}{\sigma_Y^2 + \sigma_\epsilon^2} = \kappa_1^2 (Y - P_1), $$

where $\kappa_1 \equiv \sigma_Y/\sqrt{\sigma_Y^2 + \sigma_\epsilon^2}$\textsuperscript{13} Therefore, a cutoff strategy given by $C^*$ can be sustained

\textsuperscript{12}See equation (14) in that paper.

\textsuperscript{13}The precise meaning of subscript '1' on $\kappa_1$ will become clear later when we generalize this model to
in equilibrium if the price reactions in (19) and (21) are equal. We obtain that $z^* \equiv \frac{\kappa_1 (C^* - P_1)}{\sigma_V}$ must solve:

$$-\frac{\lambda f(z^*)}{1 - \lambda + \lambda F(z^*)} = \frac{\kappa_1 \sigma_V}{\sigma_V} z^*,$$

or, equivalently,

$$-\frac{\lambda f(z^*)}{1 - \lambda + \lambda F(z^*)} = z^*.$$

(22)

It is well-known that the equation above has a unique solution; see, e.g., Acharya et al. (2011) and Dye and Hughes (2018). This solution depends on $\lambda$ alone and none of the other parameters of the model. The absolute value of the disclosure threshold can then be expressed as a function of $z^*$ solving (22) as follows:

$$C^* = \frac{\sigma_V}{\kappa_1} z^* + P_1.$$

(23)

Our Example 2 up to this point has replicated the part of analysis in Dye and Hughes (2018) that characterizes the optimal disclosure policy in a setting with one signal and risk averse investors. It is natural to ask whether the intuition of that paper extends to the case of multidimensional information endowment of the manager coupled with risk aversion on the investors' side. Our Proposition 1 suggests a general way to extend the equilibria obtained under the assumption of risk neutrality in such models to the case of risk aversion: it is sufficient to replace the prior mean, $M$, with the pre-disclosure price, $P_1$, everywhere in the description of the equilibrium in question. Yet, even in the case of risk-neutrality, Dye's (1985) model has only been solved for at most two signals (see, e.g., Pae 2005, Guttman et al. 2014).\textsuperscript{14} Below, we will characterize an equilibrium for this model for any arbitrary number of signals; before going further, however, we need to derive a pricing equation for a more specific class of possible signals, which we do in the next section.

3 Conditioning a Latent Multivariate Normal Signal

According to Proposition 1, the key step in determining the post-disclosure price $P_3$ conditional on some event $X$ is the calculation of the marginal likelihood of that event as a

\textsuperscript{14}Nezlobin (2018) describes an asymptotic equilibrium that arises when the manager's information consists of a large number of imprecise signals.
function of the prior mean $M$, 

$$
\mathbb{P}\{X\} = \int_{-\infty}^{\infty} f(V; M, \sigma_V^2) \ g(X|V) \ dV.
$$

The calculation of the above integral can be further simplified for a particular class of events often encountered in economic applications. To introduce this class, consider a latent (un-observed) multivariate signal $Y = (Y_1, ..., Y_N)$ which is distributed normally with mean $\bar{Y}$ given by a linear function of $V$:

$$
\bar{Y} \equiv \gamma V + \beta,
$$

where $\gamma$ and $\beta$ are two vectors in $\mathbb{R}^N$. Let $\Sigma$ denote the covariance matrix of $Y$.

We consider events $X$ of the form:

$$
X = \{Y \in \Omega\}
$$

for some measurable set $\Omega \subset \mathbb{R}^N$. While, as a special case, event $X$ can fully reveal the value of $Y$, in general, it does not have to be the case. For example, event $X$ can describe a situation where the true value of $Y$ is only known to exceed or lie below some multivariate threshold - such events often arise in partial-pooling equilibria. Note that while the latent variable $Y$ is jointly normally distributed with $V$, the posterior distribution of $V|X$, in general, will not be normal. For example, for events corresponding to simple multivariate truncations, the distribution of $V|X$ will belong to the skew normal family extensively studied in the statistical literature (e.g., Azzalini and Dalla Valle 1996, Lee and McLachlan 2013).

Let $f_N(Y; \bar{Y}, \Sigma)$ and $F_N(Y; \bar{Y}, \Sigma)$ denote, respectively, the probability density and cumulative distribution functions of an $N$-dimensional normal variable with location parameter $\bar{Y}$ and covariance matrix $\Sigma$. For events of the form (25), the likelihood function is written as:

$$
\mathbb{P}\{X|V\} = \int_{\Omega} f_N(Y; \bar{Y}, \Sigma) \ dY.
$$

The price reaction to event $X$, however, depends on the marginal likelihood of $X$:

$$
\mathbb{P}\{X\} = \int_{-\infty}^{\infty} f(V; M, \sigma_V^2) \ \mathbb{P}\{X|V\} \ dV
$$

$$
= \int_{-\infty}^{\infty} f(V; M, \sigma_V^2) \ \int_{\Omega} f_N(Y; \bar{Y}, \Sigma) \ dY \ dV.
$$

(26)
Using the assumption of joint normality of $Y$ and $V$, we can eliminate the outer integral in (26). To this end, let $\bar{Y}^{(1)}$ and $\Sigma^{(1)}$ denote the mean and covariance matrix of $Y$ in the “a priori” probability measure, i.e., when $V$ is treated as a random variable distributed according to the prior rather than a known parameter. Formally, let

$$\bar{Y}^{(1)} \equiv \gamma M + \beta,$$  \hspace{1cm} (27)

and

$$\Sigma_{i,j}^{(1)} \equiv \Sigma_{i,j} + \gamma_i \gamma_j \sigma_V^2$$  \hspace{1cm} (28)

for all $i, j$. We obtain the following lemma.

**Lemma 1** Assume that event $X$ is of the form (25), with $Y$ distributed normally with mean $\bar{Y}$ given by (24) and covariance matrix $\Sigma$. Then,

$$P\{X\} = \int_\Omega f_N \left( Y; \bar{Y}^{(1)}, \Sigma^{(1)} \right) dY,$$  \hspace{1cm} (29)

where $\bar{Y}^{(1)}$ and $\Sigma^{(1)}$ are given by equations (27) and (28), respectively.

The intuition behind Lemma 1 is straightforward: to calculate the marginal likelihood of event $X$, we can use the marginal density of the latent variable $Y$. When $Y$ and $V$ are jointly normal, then so is the marginal distribution of $Y$. For future reference, note that even if the components of $Y$ are independent conditionally on $V$, $\Sigma_{i,j} = 0$ for $j \neq i$, they will generally not be independent when $V$ is marginalized out, as equation (28) suggests. However, if a certain component is independent of others and as well as of $V$ (i.e., for some $i$, $\gamma_i = 0$ and $\Sigma_{i,j} = 0$ for $j \neq i$), then it will also be independent of other components in the marginal probability measure. Using Lemma 1, we can characterize the stock price reaction to events of the form (25). Let $E\left[ Y|X; \gamma m + \beta, \Sigma^{(1)} \right]$ denote the expectation of a multidimensional normal random variable $Y$ with mean $\gamma m + \beta$ and covariance matrix $\Sigma^{(1)}$, conditional on event $X$ that $Y \in \Omega$.

**Proposition 2** Assume that event $X$ is given by (25). Then, the stock price reaction to event $X$ is given by:

$$\Delta P(X) = \sigma_V^2 \gamma^T \left\{ \Sigma^{(1)} \right\}^{-1} \left( E\left[ Y|X; \gamma m + \beta, \Sigma^{(1)} \right \gamma m - \beta \right) \bigg|_{m=P_1},$$  \hspace{1cm} (30)
where $\Sigma^{(1)}$ is given by (28).

Equation (30) expresses $\Delta P (X)$ as a function of the expected surprise in latent variable $Y$ conditional on $X$ and the marginal covariance matrix of that variable. The right-hand side of (30) can also be viewed as the expected value of the multivariate normal Fisher score, with the expectation taken over all possible locations of $Y$ consistent with event $X$. Even for relatively simple sets $\Omega$, the distribution of $V|X$ can be rather complex. The main advantage of the formula in Proposition 2 is that it only relies on one integral involving the multivariate normal density, and integrals of this form are relatively well understood.

Consider, for instance, the special case in which event $X$ is that each $Y_i$ exceeds a certain threshold $Y_j$.$^{15}$ It is well-known that the left truncated mean of a multidimensional normal variable can be calculated as follows:$^{16}$

$$\mathbb{E} \left[ Y \mid \forall j, Y_j \geq Y_j; \gamma m + \beta, \Sigma^{(1)} \right] = \gamma m + \beta + \Sigma^{(1)} h \left( Y; \gamma m + \beta, \Sigma^{(1)} \right),$$

where

$$Y = (Y_1, \ldots, Y_N),$$

and $h \left( Y; \gamma m + \beta, \Sigma^{(1)} \right)$ is the multivariate normal hazard rate with mean $\gamma m + \beta$ and covariance matrix $\Sigma^{(1)}$. As in e.g. Johnson and Kotz (1975), we define the multivariate hazard rate as the negative of the gradient of the joint survival function:

$$h \left( Y; \gamma m + \beta, \Sigma^{(1)} \right) \equiv -\nabla_Y \log \left( 1 - F_N \left( Y; \gamma m + \beta, \Sigma^{(1)} \right) \right).$$

Then, by Proposition 2, we have:$^{17}$

$$\Delta P (X) = \sigma^2 \gamma^T h \left( Y; \gamma m + \beta, \Sigma^{(1)} \right) \bigg|_{m=P_1},$$

and the price reaction to event $X$ is proportional to the dot product of $\gamma$ and the normal hazard rate calculated at threshold $Y$.

Properties of the multivariate normal hazard rate are studied in, for instance, Johnson and Kotz (1975), Gupta and Gupta (1997), and Ma (2000). For example, these papers have demonstrated that each component of the gradient hazard rate is increasing in the respective

$^{15}$Truncations from above can be treated similarly using change of variables $Y_i' = -Y_i$.
$^{17}$This equation can also be derived directly from Proposition 1 applying Lemma 1.
and this result holds regardless of the covariance matrix $\Sigma$. In the univariate case, the hazard rate coincides with the inverse Mills’ ratio, which is used in the Heckman (1977) correction for selection bias. We note, however, that in the multivariate setting, Mills’ ratio is usually defined as the ratio of the survival function evaluated at some point to the probability density function evaluated at the same point (e.g., Savage 1962), and thus is different from the multivariate hazard rate. In our next example, we demonstrate the link between the stock price reaction to news whose arrival is content-dependent and Heckman’s correction for selection bias.

### 3.1 Example 3. Content-dependent news arrival

Consider a signal, $Y$, given by:

$$Y = \gamma_Y V + \epsilon,$$

where $\epsilon$ is distributed normally with mean zero and variance $\sigma^2_\epsilon$ and $\gamma_Y > 0$. Up to this point, we have considered situations where the signal is disclosed to the market irrespective of its content (Example 1) or disclosed voluntarily by a manager (Example 2). In this section, we focus on signals whose arrival is correlated, not necessarily perfectly, with their content.

Specifically, consider a latent variable $I$ determined by the following equation:

$$I = \gamma_{IY} Y + \gamma_{IV} V + \xi,$$

where $\gamma_{IY}$ and $\gamma_{IV}$ are two constants, and $\xi$ is a random variable distributed normally with mean $\bar{\xi}$ and variance $\sigma^2_\xi$, which is independent of $\epsilon$. The signal $Y$ is released to the market if latent variable $I$ exceeds zero.

When $\gamma_{IV} = 0$, the specification in (32) corresponds to news whose arrival is governed solely by $Y$. If, in addition, $\sigma^2_\xi = 0$, then signal $Y$ is disclosed only if its value exceeds a certain threshold, $-\bar{\xi}/\gamma_{IY}$, a situation that we studied in Example 2 above. As will become clear soon, $\gamma_{IV}$ is an important parameter of the model – it determines whether news arrival is correlated more strongly with the signal ($\gamma_{IV} > 0$) or noise ($\gamma_{IV} < 0$) component of $Y$. For example, when $\gamma_{IV} = -\gamma_{IY}\gamma_{YV}$, it is only the noise component of $Y$ that determines $I$; and, again, when $\gamma_{IV} = 0$, the two components of $Y$ enter (32) with coefficients proportional to those in (31), and $Y$ is a sufficient statistic for the pair $(Y, I)$ with respect to $V$. In this latter situation, the mere fact of the news arrival is not incrementally informative about $V$.
relative to \( Y \).

We seek to characterize the price reaction to the disclosure of \( Y \). When investors observe such an event \( X \), they know the exact value of \( Y \) as well as the fact that \( I > 0 \). As long as \( \gamma_{IV} \neq 0 \), the price reaction to event \( X \) will be different from what it would have been had news arrival been independent of \( Y \) and \( V \). The probit news arrival model in (32) is similar to the sample selection model of Heckman (1977), and therefore one should expect the stock price reaction to incorporate some form of a “news arrival” correction. There are, however, two important differences between our analysis and the traditional selection bias problem. First, the stock price reaction to \( X \) is determined through a Bayesian updating process; Heckman’s correction, in contrast, is applied under the frequentist approach. Second, investors in our setting are risk averse and thus not just interested in estimating the posterior mean of \( V \), but also in calculating the associated risk premium.

To apply Proposition 2, note that \( I \) can be written as:

\[
I = (\gamma_{IV} \gamma_{YV} + \gamma_{IV}) V + \gamma_{IV} \epsilon + \xi.
\]

Therefore, vector \( \gamma \) (from eq. 24) corresponding to bivariate signal \((Y, I)\) is given by:

\[
\gamma = \begin{pmatrix} \gamma_{YV} \\ \gamma_{IV} \end{pmatrix}.
\]

(33)

It is straightforward to check that the covariance matrix of the marginal joint distribution of \((Y, I)\) is:

\[
\Sigma^{(1)} = \begin{pmatrix} \gamma_{YV}^2 \sigma_V^2 + \sigma_\epsilon^2 & \Sigma_{1,2}^{(1)} \\ \Sigma_{1,2}^{(1)} & (\gamma_{IV} \gamma_{YV} + \gamma_{IV})^2 \sigma_V^2 + \gamma_{IV}^2 \sigma_\epsilon^2 + \sigma_\xi^2 \end{pmatrix},
\]

(34)

where

\[
\Sigma_{1,2}^{(1)} \equiv \gamma_{IV} \gamma_{YV} + \gamma_{IV} \sigma_V^2 + \gamma_{IV} \sigma_\epsilon^2.
\]

The sign of \( \Sigma_{1,2}^{(1)} \) determines whether variables \( Y \) and \( I \) are, a priori, positively or negatively correlated.

We start by characterizing the price reaction in the risk-neutral setting; later, we will apply Proposition 1 to extend our results to the case with risk aversion. Using standard results for normally distributed signals (e.g., our Example 1), one can verify that had signal

\footnote{Proposition 2 can also be applied to determine the price reaction to the event of non-disclosure.}
Y been always disclosed, the price reaction would have been given by:

\[ \pi_Y^0 (Y - \gamma_Y V M), \]

with

\[ \pi_Y^0 \equiv \frac{\gamma_Y V \sigma_Y^2}{\gamma_Y V \sigma_Y^2 + \sigma_e^2}. \]

Furthermore, had the exact values of both Y and I been disclosed simultaneously, the price reaction would have been equal to:

\[ \pi_Y (Y - \gamma_Y V M) + \pi_I (I - (\gamma_I V \gamma_Y V + \gamma_I V) M - \bar{\xi}), \]

where

\[ \pi_Y = \frac{\gamma_Y V \sigma_Y^2 \sigma_e^2 - \gamma_I V \gamma_I V \sigma_Y^2}{(\gamma_Y V \sigma_Y^2 + \sigma_e^2) \sigma_e^2 + \gamma_I V \sigma_Y^2 \sigma_e^2}, \]

and

\[ \pi_I = \frac{\gamma_I V \sigma_Y^2 \sigma_e^2}{(\gamma_Y V \sigma_Y^2 + \sigma_e^2) \sigma_e^2 + \gamma_I V \sigma_Y^2 \sigma_e^2}. \]

For future reference, observe that \( \pi_I \) always has the same sign as \( \gamma_I V \). The intuition for this is straightforward: if \( \gamma_I V > 0 \), then, conditional on \( Y \), \( I \) is incrementally informative about \( V \), and therefore \( \pi_I > 0 \). In contrast, if \( \gamma_I V < 0 \), then \( I \) informs investors about the noise component of \( Y \), and, if observed, it enters the valuation function with a negative coefficient.

**Observation 2** Assume signal \( Y \) and its arrival are characterized by (31) and (32). The price reaction to the disclosure of \( Y \) is given by:

\[ \Delta P(X) = \pi_Y^0 (Y - \gamma_Y V M) + \pi_I \sqrt{\mathbb{V}[I|Y]} h \left( -\frac{\bar{\xi} + (\gamma_I V + \gamma_I V \pi_Y^0) Y}{\sqrt{\mathbb{V}[I|Y]}} \right), \quad (35) \]

where \( h(\cdot) \) is the hazard rate of a standard normal variable, and \( \mathbb{V}[I|Y] \) is the conditional variance of \( I|Y \), given by:

\[ \mathbb{V}[I|Y] \equiv \frac{\gamma_I V \sigma_Y^2 \sigma_e^2}{\gamma_Y V \sigma_Y^2 + \sigma_e^2} + \sigma_e^2. \]

The stock price reaction in (35) consists of two components – the price reaction to the news content and a “news arrival” correction.
\[ \pi_Y^o (Y - \gamma_{YV} M) + \pi_I \sqrt{\nabla [I|Y]} h \left( \frac{-\bar{\xi} + (\gamma_{IY} + \gamma_{IV} \pi_Y^o) Y}{\sqrt{\nabla [I|Y]}} \right) . \]

**Reaction to content of** $Y$

**News arrival correction**

As in the sample selection literature, the news arrival correction is proportional to the hazard rate (inverse Mills’ ratio) evaluated at the arrival surprise conditional on signal $Y$. To glean further intuition for the expression above, recall some of the properties of the standard normal hazard rate $h(t)$: it is everywhere increasing and convex in $t$, approaching zero at $t \to -\infty$ and behaving approximately as $t$ when $t \to \infty$. It immediately follows that $\Delta P(X)$ in (35) is convex in $Y$ if $\pi_I > 0$ (which happens when $\gamma_{IV} > 0$), and concave in $Y$ if $\gamma_{IV} < 0$. Therefore, the price reaction to news is convex (concave) if their arrival is incrementally informative about $V$ (the noise component of $Y$).

It is also interesting to consider the slopes of the asymptotes of $\Delta P(X)$ when $Y \to \pm \infty$. Using straightforward algebra, one can check that

\[ \gamma_{IY} + \gamma_{IV} \pi_Y^o = \frac{\Sigma_{1,2}^{(1)}}{(\gamma_{YV}^2 \sigma_V^2 + \sigma_Y^2)}, \]

and, therefore, the argument of the hazard rate in (35) increases in $Y$ when $\Sigma_{1,2}^{(1)} < 0$ and decreases in $Y$ when $\Sigma_{1,2}^{(1)} > 0$. When the argument of $h(\cdot)$ approaches $-\infty$, the slope of $\Delta P(X)$ with respect to $Y$ approaches $\pi_Y^o$, i.e., the news arrival correction becomes effectively irrelevant. For example, if $I$ and $Y$ are a priori positively correlated, $\Sigma_{1,2}^{(1)} > 0$, and the news is very favorable, $Y \to \infty$, then there is little additional surprise in the fact of news arrival.

When the argument of $h(\cdot)$ in (35) approaches $+\infty$, then, recalling that $h(t) \approx t$ for large $t$, the pricing function approaches:

\[ \pi_Y^o (Y - \gamma_{YV} M) - \pi_I \left( \bar{\xi} + (\gamma_{IY} + \gamma_{IV} \pi_Y^o) Y \right), \]

which after some algebra reduces to

\[ \pi_Y (Y - \gamma_{YV} M) - \pi_I \left( \bar{\xi} + \gamma_{YV} (\gamma_{IY} + \gamma_{IV} \pi_Y^o) M \right). \]

To summarize, the coefficient on $Y$ always varies between $\pi_Y^o > 0$ and $\pi_Y$, the exact relationship between which depends on the sign of $\Sigma_{1,2}^{(1)}$ (determines the a priori correlation between

---

\[^19\] Recall that $\Sigma_{1,2}^{(1)} > 0$ whenever the marginal distributions of $Y$ and $I$ are positively correlated.
Figure 1: $\Delta P(X)$ as a function of signal $Y$. In all panels, $M$ and $\bar{\xi}$ are set to zero. The solid blue line depicts $\Delta P(X)$; the dashed red line - the asymptote $\pi_Y Y$; the dotted brown line - the asymptote $\pi_Y Y$. The panels show the four different cases that can arise depending on the signs of $\Sigma^{(1)}_{1,2}$ and $\gamma_{IV}$.
\( I \) and \( Y \) \( \text{and} \ \pi_I \) (determines whether \( I \) is incrementally informative about \( V \) or noise in \( Y \)). We depict the four possible cases in Figure 1.

Note that in cases (a) and (d), \( \Delta P (X) \) can be declining in \( Y \) for low and high values of \( Y \), respectively. This non-monotonicity obtains when \( \pi_Y < 0 \), i.e., when

\[
\gamma_{YV} \sigma^2_\xi < \gamma_{IV} \gamma_{IY} \sigma^2_\epsilon.
\]

In case (a), the inequality above holds if, for example, \( \sigma^2_\xi \to 0 \) (i.e., conditionally on \( Y \) and \( V \), there is little uncertainty in news arrival), and \( \sigma^2_\epsilon \) is large (signal \( Y \) is imprecise). In such a situation, when investors observe a very low signal \( Y \), they understand that the value of \( I \) must be sufficiently high to warrant the signal's arrival. The two pieces of information about \( V - Y \) and \( I \) then effectively contradict each other. Since \( I \) attaches a higher weight to the signal component of \( Y \) than to its noise component, the most likely explanation for such contradictory observations becomes a very high realization of \( V \) coupled with a very low realization of \( \epsilon \); and the greater is the dissonance between \( Y \) and \( I \), the higher is the posterior mean of \( V \). Thus, the mere fact of arrival of very bad news can lead to a positive stock price reaction.

To conclude this example, we note that, by Proposition 1, it is straightforward to extend Observation 2 to the setting with risk averse investors. The arrival correction term in (35) is independent of \( M \), and thus the reaction of risk averse investors to event \( X \) will exceed that given in equation (35) by:

\[
\rho \pi^o_Y \gamma_{YV} \sigma^2_V.
\]

As in Example 1, the expected post-disclosure price is now greater than the pre-disclosure price due to uncertainty resolution and the corresponding reduction in the risk premium. This argument illustrates again the power of Proposition 1: it provides a simple way to determine the effect of risk aversion on the firm’s stock price in settings outside the case of joint Gaussianity. For events we consider in this example, the distribution of \( V|X \) is not normal but rather belongs to the skew normal (or power normal) family.

---

\(^{20}\)Recall that \( \pi^o_Y > 0 \) because we assume that \( \gamma_{YV} > 0 \), i.e., higher values of \( V \) lead in expectation to higher values of \( Y \).
4 Multidimensional Voluntary Disclosures

In this section, we consider a voluntary disclosure game with multidimensional information, in which the firm’s manager has up to \( N \) signals that can be disclosed to the market. The model builds on Dye (1985) in that the informational endowment of the manager is uncertain; specifically, the signals are received by the manager independently of each other, each with probability of \( \lambda < 1 \). This uncertainty about the manager’s actual information set prevents full unraveling from happening in equilibrium. As in, for instance, Fishman and Hagerty (1990), Pae (2005) and Guttman et al. (2014), we further assume that the maximum number of signals \( (N) \) is common knowledge, as is \( \lambda \), and that the event that a certain signal is received by the manager is independent of the signal’s value and of all other random variables of the model. The manager cannot participate in the stock market as either a buyer or a seller.

It will be convenient to represent the manager’s information endowment by a \( 2N \)-dimensional normal variable, \( G \). The first \( N \) components of \( G \) are signals \( Y_1, ..., Y_N \), each one of which is distributed normally with mean \( V \) and variance \( \sigma^2 \). Conditionally on \( V \), the signals are independent of each other. The second \( N \) components of \( G \), denoted by \( I_1, ..., I_N \), indicate whether the corresponding signals are received by the manager and thus can be disclosed to the market. Random variables \( I_1, ..., I_N \) have a variance of one, are independent of \( \{Y_i\} \) and of each other, and have a mean of \( \bar{X} = F^{-1} (\lambda) \), so that \( \mathbb{P} \{I_j > 0\} = \lambda \). Accordingly, signal \( Y_j \) is assumed to be received whenever \( I_j > 0 \).

As in much of the literature on persuasion games, the signals can be costlessly disclosed to the market, but such disclosures must be truthful. We also make the usual assumption that the manager cannot credibly communicate that she did not get to observe certain information. Finally, all signals are assumed to be disclosed to the market simultaneously, which appears to be descriptive of at least some real-life situations, such as investor earnings calls and new product presentations.

Note that the disclosure game starts once nature draws \( G \), and there are no other random variables drawn from the physical probability measure before price \( P_3 \) is realized. We will, therefore, first describe an equilibrium of the disclosure game under the risk-neutral probability measure, i.e., assuming that \( \rho \rightarrow 0 \). Then, applying Proposition 1, we will make the substitution \( M = P_1 \) everywhere in the description of the equilibrium to characterize the effect of investors’ risk aversion on the manager’s disclosure decisions.

To adapt our result in Proposition 2 to the current problem, note that the mean of \( G \)
can be written as:
\[
\begin{pmatrix}
V, ..., V, \bar{\lambda}, ..., \bar{\lambda}
\end{pmatrix}_{\text{N entries}}^T = \gamma_G V + \beta_G,
\]
where \( \gamma_G = (1_{1 \times N}, 0_{1 \times N})^T \), \( \beta_G = (0_{1 \times N}, 1_{1 \times N})^T \), and \( 1_{i \times j} \) and \( 0_{i \times j} \) are \( i \)-by-\( j \) matrices of ones and zeros, respectively. It is also straightforward to check that the covariance matrix of the marginal distribution of \( G \) (calculated according to Lemma 1) has the following structure:
\[
\Sigma^{(1)}_G = \begin{pmatrix}
\sigma^2 \mathbf{1}_{N \times N} + \sigma^2 \mathbf{1}_{N \times N} & 0_{N \times N} \\
0_{N \times N} & \mathbf{I}_{N \times N}
\end{pmatrix},
\]
(36)
where \( \mathbf{I}_{N \times N} \) is the \( N \times N \) identity matrix.

Let \( Y[d] \equiv (Y_1, ..., Y_d) \) denote the vector of the first \( d \) signals, and let
\[
G[-d] \equiv (Y_{d+1}, ..., Y_N, I_{d+1}, ..., I_N)
\]
denote the components of \( G \) that are related to the remaining \( N - d \) signals. Using a standard Bayesian updating formula, for \( i > d \), we can calculate the expectation of \( Y_i \) conditional on observing \( Y[d] \):
\[
\mathbb{E}[Y_i|Y[d]] = M + \kappa^2_d \sum_{i=1}^{d} (Y_i - M),
\]
(37)
where
\[
\kappa_d \equiv \frac{\sigma_V}{\sqrt{d\sigma^2_V + \sigma^2_e}}.
\]
In the following lemma, we characterize the price reaction to events in which \( Y[d] \) is known and there is also some additional information about \( G[-d] \). This lemma is a special case of Proposition 2 for the values of \( \gamma_G \) and \( \Sigma^{(1)}_G \) described above.

**Lemma 2** Let \( X \) be the following event: \( Y[d] \) is disclosed, and it is also known that the vector \( G[-d] \) belongs to some set \( \Omega \subset \mathbb{R}^{2(N-d)} \). Then, the price reaction to event \( X \) is given by:
\[
\Delta P(X) = \kappa^2_d \sum_{i=1}^{d} (Y_i - M) + \kappa^2_N \mathbb{E} \left[ \sum_{i=d+1}^{N} Y'_i | G[-d] \in \Omega \right],
\]
where
\[
Y'_i \equiv Y_i - \mathbb{E}[Y_i|Y[d)],
\]
(38)
and \( \mathbb{E} [Y_i | Y [d]] \) is given by equation (37).

In the equilibrium we present in this section, the firm’s value is increasing in each reported signal, and, as a consequence, the manager always discloses some number of the highest received signals. Let \( T_i \) denote the \( i \)-th best signal received by the manager, with \( T_i \equiv -\infty \) if the manager observed fewer than \( i \) signals. Then, the equilibrium is characterized by two main conditions. First, if the market observes the disclosure of \( d \) signals (and such disclosure is admissible, as defined below), the price reaction to this information, \( \Delta P \left( \{T_i\}_{i=1}^d \right) \), is given by:

\[
\Delta P \left( \{T_i\}_{i=1}^d \right) = \kappa^2_d \sum_{i=1}^d (T_i - M) + L_{N-d},
\]

where \( L_1, \ldots, L_N \) are some constants that depend on \( \lambda, \sigma_e, \sigma_V \), but not on \( M \) or any of the disclosed \( T_i \). Intuitively, \( L_i \) reflects a fixed penalty for withholding information about \( i \) signals.

Second, the manager chooses the number of signals to disclose to maximize \( \Delta P \left( \{T_i\}_{i=1}^d \right) \) given in (39), subject to the additional constraint that the lowest one of the disclosed signals is not too low relative to all the other disclosed signals:

\[
T_d \geq M + \kappa^2_d \sum_{i=1}^d (T_i - M) + C_d,
\]

where the cutoffs \( C_d \) also depend only on \( \lambda, \sigma_e, \) and \( \sigma_V \), but not \( M \) or any \( T_i \). We will call disclosures satisfying (40) admissible.\(^{21}\) Full disclosure is always admissible, \( C_N = -\infty \). In the proof of Proposition 2, we specify the off-equilibrium beliefs ensuring that the manager never wants to make an inadmissible disclosure.

**Proposition 3** There exists an equilibrium, characterized by constants \( (L_1, \ldots, L_N) \) and \( (C_1, \ldots, C_{N-1}) \), in which the manager discloses the top \( d \) received signals if (i) \( T_d \) satisfies the

\(^{21}\)While the current form of inequality (40) is convenient for our future analysis, it can also be written in a more intuitive fashion:

\[
T_d \geq \frac{\sigma^2_e}{(d-1)\sigma^2_e + \sigma^2_V} M + \frac{\sigma^2_V}{(d-1)\sigma^2_e + \sigma^2_V} \sum_{i=1}^{d-1} (T_i - M) + \frac{d\sigma^2_e + \sigma^2_V}{(d-1)\sigma^2_e + \sigma^2_V} C_d.
\]

The first two terms in the right-hand side above are effectively the posterior mean of a normal distribution with a prior of \( M \) after updating on signals \( \{T_i\}_{i=1}^{d-1} \); the third term is a constant independent of \( M \) and \( \{T_i\}_{i=1}^{d-1} \).
admissibility constraint in (40), and (ii) among all admissible disclosures, \(d\) maximizes the firm’s price given by (39).

Before proceeding, we must note that the equilibrium described in Proposition 3 is not unique. In general, it is well-known that in disclosure games with multidimensional information, there are often multiple equilibria supported by different off-equilibrium beliefs (see, e.g., Shin 2003, and Guttman et al. 2014). In particular, by imposing sufficiently high off-equilibrium punishments, it is easy to reduce the amount of partial disclosures in any given equilibrium. As we discuss below, in our setting, the manager faces an off-equilibrium punishment if she chooses to disclose \(d\) signals violating the admissibility constraint in (40). Such off-equilibrium beliefs, however, appear to be reasonably sparing, at least in the sense that any given number of signals is disclosed with strictly positive probability in our equilibrium.

Why is it necessary to have some off-equilibrium punishment for a manager who makes an inadmissible disclosure? Consider a situation where the number of signals is large, and the manager observes relatively high realizations of \(T_1, \ldots, T_{d-1}\). Let \(\mathbb{E}[V | \{T_i\}_{i=1}^{d-1}]\) denote the expected value of \(V\) conditional on those signals alone, ignoring any strategic considerations.\(^{22}\) Since we are assuming that the top \(d - 1\) signals happen to be relatively large, so is \(\mathbb{E}[V | \{T_i\}_{i=1}^{d-1}]\).

Now assume that the manager discloses those \(d - 1\) signals and some very small signal \(T_d'\), thus violating constraint (40). Investors would then be led to believe that either (i) all the remaining received signals are less than \(T_d'\), which is an unlikely event given the large value of \(\mathbb{E}[V | \{T_i\}_{i=1}^{d-1}]\), or (ii) no more signals were received by the manager. The conditional likelihood of event (ii) increases as the gap between \(\mathbb{E}[V | \{T_i\}_{i=1}^{d-1}]\) and \(T_d'\) gets larger.\(^{23}\) Furthermore, it is intuitive that the non-disclosure penalty should be lower in cases when it is due to the absence of signals rather than withholding of unfavorable information. The off-equilibrium punishment that we specify in the proof of Proposition 3 prevents the manager from adding one unfavorable signal to an otherwise strong report with the goal of minimizing the penalty for the non-disclosure of remaining signals.

To shed more light on the equilibrium presented in Proposition 3, let us now discuss the manager’s disclosure incentives. When will the manager prefer to disclose more signals \((j)\)

\(^{22}\)In other words, this expected value is calculated under the assumption that investors somehow got to observe the top \(d - 1\) of the received signals, without any involvement from the manager.

\(^{23}\)Note also that since we assume that \(d\) is large, the presence of one low signal does not in itself significantly move the posterior mean relative to \(\mathbb{E}[V | \{T_i\}_{i=1}^{d-1}]\).
to fewer \((d)\)? The following corollary provides the answer that follows from equation (39) by straightforward algebra.

**Corollary 1** Assume that \(j > d\) and both disclosures \(\{T_i\}_{i=1}^d\) and \(\{T_i\}_{i=1}^j\) satisfy the admissibility constraint (40). Then, inequality

\[
\Delta P(\{T_i\}_{i=1}^j) > \Delta P(\{T_i\}_{i=1}^d)
\]

is equivalent to:

\[
\sum_{i=d+1}^{j} T_i > \frac{(j-d)\sigma^2\kappa^2}{\sigma^2 V} M + (j-d)\kappa^2 \sum_{i=1}^{d} T_i + \frac{1}{\kappa^2_j} (L_{N-d} - L_{N-j}).
\]

(41)

According to equation (41), the manager prefers to disclose an additional \(j - d\) signals if the sum of those signals exceeds a threshold determined by \(M\) and the sum of the best \(d\) signals. Three observations are in order. First, note that the right-hand side of (41) increases in \(\sum_{i=1}^{d} T_i\), that is, the higher is the sum of the first \(d\) signals, the less likely are the remaining signals to be disclosed. On the margin, each signal is evaluated not only relative to the prior mean of \(V\), but also relative to all the information that will be disclosed “ahead” of that signal; thus firms with great news to share are less likely to disclose “just fine” news.

Second, the right-hand side of (41) increases in \(M\), which suggests that investors’ risk-aversion generally strengthens disclosure incentives. Indeed, applying out result in Proposition 1 to the equilibrium described in Proposition 3, we obtain that the manager discloses \(d\) signals under two conditions. First, disclosure \(\{T_i\}_{i=1}^d\) must be admissible and satisfy the following inequality:

\[
T_d \geq M - \rho\sigma^2_V + \kappa^2_d \sum_{i=1}^{d} (T_i - (M - \rho\sigma^2_V)) + C_d
\]

\[
= \frac{\sigma^2\kappa^2}{\sigma^2 V} (M - \rho\sigma^2_V) + \kappa^2_d \sum_{i=1}^{d} T_i + C_d.
\]

(42)

It can be immediately seen that any disclosure that is admissible with risk-neutral investors is also admissible with risk-averse investors. Second, among admissible strategies, the disclosure of \(d\) signals must be optimal. Note that after substituting \(M\) for \(M - \rho\sigma^2_V\), inequality
(41) becomes:

$$\sum_{i=d+1}^{j} T_i > \frac{(j-d) \sigma^2 \kappa_i^2}{\sigma^2} \left( M - \rho \sigma^2 \right) + (j-d) \kappa_i^2 \sum_{i=1}^{d} T_i + \frac{1}{\kappa_j^2} (L_{N-d} - L_{N-j}),$$

which suggests that the threshold for disclosing additional signals is unequivocally lower in the setting with risk averse investors.\footnote{In fact, the only situation in which the manager will prefer to disclose less to risk-averse investors than to risk-neutral investors arises when the manager’s information set is such that there is an inadmissible, in the risk neutral setting, disclosure with a small number of signals yet a large price reaction that becomes admissible with risk averse investors. This only happens for a relatively narrow band of signal values, i.e., when the value of $d$ maximizing the expression in (39) with $M$ substituted for $M - \rho \sigma^2$ satisfies inequality (42) but not (40).}

Finally, equation (41) allows us to glean intuition regarding the shape of the firm’s non-disclosure region. Pae (2005) demonstrates that in a model with two signals, a fully informed firm does not disclose anything a region of $\mathbb{R}^2$ constrained by three hyperplanes, so that $Y_1 \leq D_1$, $Y_2 \leq D_1$, and $Y_1 + Y_2 \leq D_2$.\footnote{See Figure 3 in Pae (2005).} Our results suggest a natural generalization of this non-disclosure region: a fully informed firm remains silent if, for all $j \in [1, N]$, the sum of any $j$ signals satisfying the admissibility constraint is bounded from above by some constant:\footnote{This inequality follows from (41) by setting $d = 0$.}

$$\sum_{i=1}^{j} T_i < j M + \frac{1}{\kappa_j^2} (L_N - L_{N-j}).$$

Again, the condition above becomes more restrictive if investors are risk averse, thus making firms more likely to voluntarily disclose at least some of their signals.

\section{5 Conclusion}

In this paper, we have shown that the stock price reaction to a piece of news is proportional to its Fisher score calculated under the risk-neutral measure, and this result holds for a large class of news distributions. We have also constructed an equilibrium in a multidimensional voluntary disclosure game, and characterized the price reaction to news with content-dependent arrival.

There are at least two natural directions for future study. First, it would be interesting to extend our analysis to settings outside the CARA-Normal case. It is clear that the steps
leading up to Proposition 1 have relied substantially on this structure, but it is also possible that similar, if approximate, results can be obtained in other settings. Second, one might consider dynamic settings in which information is released to the market gradually. For the case of risk-neutral investors, such models have been analyzed in, e.g., Acharya et al. (2011) and Guttman et al. (2014), yet less is known about the dynamic pricing of non-Gaussian news by risk-averse investors.

References


Appendix A

Proof of Proposition 1.

In this proof, we verify that for all values of $\alpha$ satisfying the first-order condition in (9), the second derivative of the target function is negative. Since the target function is everywhere twice continuously differentiable in $\alpha$, this will imply that there is only one extremum and it is the function’s global maximum.$^{27}$

When the first-order condition (9) is satisfied, the second derivative can be written as:

$$\exp \left\{-\rho \left( \alpha M - \frac{\rho \alpha^2}{2} \sigma^2_V - \alpha P_3 \right) \right\} \times \left( -\rho \sigma^2_V \mathbb{P} \{X; m\} - \rho \sigma^2_V (M - \rho \alpha \sigma^2_V - P_3) \frac{\partial \mathbb{P} \{X; m\}}{\partial m} \right.$$

$$\left. - \rho \sigma^4_V \frac{\partial^2 \mathbb{P} \{X; m\}}{\partial m^2} \right|_{m=M-\rho \alpha \sigma^2_V} = 0. \quad (43)$$

Again, due to (9), we have:

$$\left( M - \rho \alpha \sigma^2_V - P_3 \right) \frac{\partial \mathbb{P} \{X; m\}}{\partial m} \bigg|_{m=M-\rho \alpha \sigma^2_V} = -\sigma^2_V \frac{\partial \log \mathbb{P} \{X; m\}}{\partial m} \bigg|_{m=M-\rho \alpha \sigma^2_V} \quad (44)$$

Therefore, the derivative in (43) is of the same sign as:

$$\left( -\mathbb{P} \{X; m\} + \frac{\sigma^2_V \left( \frac{\partial \mathbb{P} \{X; m\}}{\partial m} \right)^2}{\mathbb{P} \{X; m\} - \frac{\sigma^2_V \frac{\partial^2 \mathbb{P} \{X; m\}}{\partial m^2}}{m=M-\rho \alpha \sigma^2_V} \right) \right.$$  

Now note that

$$\frac{\partial \mathbb{P} \{X; m\}}{\partial m} = \int_{-\infty}^{\infty} \frac{(V-m)}{\sigma^2_V} f(V; m, \sigma^2_V) g(X|V) \, dV = \frac{1}{\sigma^2_V} \mathbb{E} [V - m | X; m] \mathbb{P} \{X; m\}, \quad (45)$$

where the second equality is obtained using the definition of the expected value and the

$^{27}$The existence of at least one extremum is guaranteed for $P_3$ given by (10).
density in (5). Similarly,
\[
\frac{\partial^2 \mathbb{P}\{X;m\}}{\partial m^2} = \frac{\partial}{\partial m} \int_{-\infty}^{\infty} \frac{(V-m)}{\sigma^2_V} \int_{-\infty}^{\infty} f(V;m,\sigma^2_V) \, g(X|V) \, dV \\
= \frac{1}{\sigma^2_V} \mathbb{E}[(V-m)^2|X;m] \, \mathbb{P}\{X;m\} - \frac{1}{\sigma^2_V} \mathbb{P}\{X;m\}. \tag{46}
\]
Substituting (45) and (46) into (44) yields
\[
-P\{X;m\} + \frac{1}{\sigma^2_V} \mathbb{E}[V-m|X;m]^2 \mathbb{P}\{X;m\} - \frac{1}{\sigma^2_V} \mathbb{E}[(V-m)^2|X;m] \mathbb{P}\{X;m\} + \mathbb{P}\{X;m\} < 0
\]
for any \(m\) by Jensen’s inequality.

\[\text{Proof of Lemma 1.}\]
Recall that
\[
\mathbb{P}\{Y \in \Omega\} = \int_{-\infty}^{\infty} f(V;M,\sigma^2_V) \int_{\Omega} f_N(Y;\bar{Y},\Sigma) \, dY \, dV \\
= \int_{-\infty}^{\infty} f(z) \int_{\Omega} f_N(Y;\gamma M + \beta + \gamma \sigma_V z,\Sigma) \, dY \, dz, \tag{47}
\]
where the second equality follows after the change of variables \(z \equiv (V-M) / \sigma_V\).
Now observe that the right-hand side of (47) is equal to:\[\text{Proof of Proposition 2.}\]
\[\text{Proof of Proposition 2.}\]

\[\text{Proof of Proposition 2.}\]

\[\text{Proof of Proposition 2.}\]

\[\text{Proof of Proposition 2.}\]
By Lemma 1,
\[ \mathbb{P} \{ Y \in \Omega \} = \int_{\Omega} f_N \left( Y; \bar{Y}^{(1)}, \Sigma^{(1)} \right) dY. \]

Therefore,
\[
\frac{\partial \mathbb{P} \{ Y \in \Omega \}}{\partial M} = \int_{\Omega} \frac{\partial f_N \left( Y; \bar{Y}^{(1)}, \Sigma^{(1)} \right)}{\partial M} dY \\
= \gamma^T \left\{ \Sigma^{(1)} \right\}^{-1} \int_{\Omega} \left( Y - \bar{Y}^{(1)} \right) f_N \left( Y; \bar{Y}^{(1)}, \Sigma^{(1)} \right) dY \\
= \gamma^T \left\{ \Sigma^{(1)} \right\}^{-1} \left( \mathbb{E} \left[ Y \mid X; \bar{Y}^{(1)}, \Sigma^{(1)} \right] - \bar{Y}^{(1)} \right) \mathbb{P} \{ Y \in \Omega \}.
\]

Then,
\[
\frac{\partial \log \mathbb{P} \{ X \}}{\partial M} = \frac{\partial \mathbb{P} \{ Y \in \Omega \}}{\partial M} / \mathbb{P} \{ Y \in \Omega \} \\
= \gamma^T \left\{ \Sigma^{(1)} \right\}^{-1} \left( \mathbb{E} \left[ Y \mid X; \bar{Y}^{(1)}, \Sigma^{(1)} \right] - \bar{Y}^{(1)} \right),
\]
and the claim of Proposition 2 now follows directly from Proposition 1.

**Proof of Observation 2.**

Applying Proposition 2 to \( \gamma \) and \( \Sigma^{(1)} \) in equations (33) and (34), we obtain:
\[
\Delta P (X) = \pi_o^Y (Y - \gamma_Y V M) + \pi_I \left( \mathbb{E} [I \mid Y, I > 0] - (\gamma_I Y \gamma_Y + \gamma_I V) M - \bar{\xi} \right).
\]

The distribution of \( I \mid Y \) is normal, with mean
\[
\mathbb{E} [I \mid Y] = (\gamma_I Y \gamma_Y + \gamma_I V) M + \bar{\xi} + \Sigma^{(1)} \left\{ \Sigma^{(1)} \right\}^{-1} (Y - \gamma_Y V M),
\]
and variance
\[
\mathbb{V} [I \mid Y] = \Sigma^{(1)}_{2,2} - \Sigma^{(1)}_{1,2} \left\{ \Sigma^{(1)} \right\}^{-1} \Sigma^{(1)}_{1,2} \\
= \frac{\gamma^2_I \sigma_V^2 \sigma^2_{\xi}}{\gamma^2_Y V^2 \sigma^2_V + \sigma^2_{\xi}} + \sigma^2_{\xi}.
\]

The claim of the Observation then follows by straightforward algebra from the standard expression for the mean of a truncated normal variable.
Proof of Lemma 2.

It is straightforward to verify that

\[
\left( \Sigma_G^{(1)} \right)^{-1} = \begin{pmatrix}
\frac{1}{\sigma^2} (I_{N \times N} - \kappa_N 1_{N \times N}) & 0_{N \times N} \\
0_{N \times N} & I_{N \times N}
\end{pmatrix}.
\]

Recalling that \( \gamma_G = \left( \frac{1}{N} \ldots , \frac{1}{N} \right) \), we have:

\[
\gamma_G^T \left( \Sigma_G^{(1)} \right)^{-1} = \frac{\kappa_N}{\sigma^2} \left( \frac{1}{N} \ldots , \frac{1}{N} \right).
\]

Then, by Proposition 2,

\[
\Delta P (X) = \kappa_N^2 \sum_{i=1}^{d} (Y_i - M) + \kappa_N^2 \mathbb{E} \left[ \sum_{i=d+1}^{N} (Y_i - M) \mid G[-d] \in \Omega \right].
\]

Equation (37) follows by a well-known formula for the posterior mean in the Bayesian updating problem with a normal prior on the mean and normally distributed signals. Finally, straightforward algebra yields

\[
\kappa_N^2 = \kappa_d^2 \left( 1 - (N - d) \kappa_N^2 \right),
\]

which implies that the right-hand side in (48) can be rewritten as:

\[
\kappa_d^2 \sum_{i=1}^{d} (Y_i - M) + \kappa_N^2 \mathbb{E} \left[ \sum_{i=d+1}^{N} \left( Y_i - \kappa_d^2 \sum_{i=1}^{d} (Y_i - M) - M \right) \mid G[-d] \in \Omega \right].
\]

Proof of Proposition 3.

We will construct an equilibrium by induction, beginning with the last non-disclosure penalty \( L_1 \) and cutoff \( C_{N-1} \). Note that the penalty \( L_1 \) is relevant only if the firm discloses all but one signal \( - \ldots , T_{N-1} \), and if such disclosure satisfies

\[
T_{N-1} \geq M + \kappa_{N-1}^2 \sum_{i=1}^{N-1} (T_i - M) + C_{N-1}.
\]
As we construct the equilibrium, we will assume that the manager never reports \( d \) signals if the lowest reported signal is below its respective cutoff given by (40); once the full equilibrium is constructed, we will specify off-equilibrium beliefs which indeed ensure that this is case.

To identify \( C_{N-1} \) and \( L_1 \), consider the following event \( ND_1 \): the manager reports \( T_1, \ldots, T_{N-1} \) and does not disclose only the last remaining signal. Without loss of generality, assume that the undisclosed signal is \( Y_N \). Consider random variable \((Y'_N, I_N)\) with \( Y'_N \) defined according to (38). Conditional just on the values \( Y_1, \ldots, Y_{N-1} \), \((Y'_N, I_N)\) is a bivariate normal with independent components and mean \((0, \bar{\lambda})\).

Conditional on the fact that if \( Y_N \) is received by the manager then it is the lowest signal, \((Y'_N, I_N)\) must belong to the following feasible set \( \Omega^f_1(T_1, \ldots, T_{N-1}) \):

\[
\Omega^f_1(T_1, \ldots, T_{N-1}) \equiv \left\{ \left( Y'_N, I_N \right) \mid Y'_N < T_{N-1} - M - \kappa^2_{N-1} \sum_{i=1}^{N-1} (T_i - M) \text{ or } I_N < 0 \right\}.
\]

For some values of \((Y'_N, I_N) \in \Omega^f_1(T_1, \ldots, T_{N-1})\) with \( I_N > 0 \), the manager prefers to disclose all signals. Let \( \Omega^{ND}_1(T_1, \ldots, T_{N-1}) \) denote the subset of \( \Omega^f_1(T_1, \ldots, T_{N-1}) \) where non-disclosure of \( Y_N \) is optimal or that signal is not received by the manager \((I_N < 0)\).\(^{29}\) By Lemma 2, the price reaction to event \( ND_1 \) must be given by:

\[
\Delta P(ND_1) = \kappa^2_{N-1} \sum_{i=1}^{N-1} (Y_i - M) + \kappa^2_N E\left[ Y'_N \mid (Y'_N, I_N) \in \Omega^{ND}_1(T_1, \ldots, T_{N-1}) \right]. \tag{49}
\]

The price reaction to full disclosure, \( FD \), is:

\[
\Delta P(FD) = \kappa^2_{N-1} \sum_{i=1}^{N-1} (Y_i - M) + \kappa^2_N Y'_N. \tag{50}
\]

Expression (49) must exceed (50) for \((Y'_N, I_N) \in \Omega^{ND}_1\) and the opposite inequality must hold for \((Y'_N, I_N) \in \Omega^f_1 \setminus \Omega^{ND}_1\).

From or analysis of the univariate Dye (1985) model in Example 2, we know that if \((Y, I)\) is distributed as a bivariate normal with mean \((0, \bar{\lambda})\) and independent components, then there exists some cutoff \( C^* \) such that

\[
E[Y \mid Y < C^* \text{ or } I < 0] = C^*.
\]

\(^{29}\)We will sometimes omit the parameters of \( \Omega^{ND}_1\) and \( \Omega^f_1\) to simplify notation.
Given this constant $C^*$, let

$$\Omega_{1}^{ND}(T_1, ..., T_{N-1}) \equiv \left\{ \left( Y'_N, I_N \right) \mid Y'_N < C^* \text{ or } I_N < 0 \right\}.$$  

Observe that as long as $T_{N-1} > C^* + M + \kappa_{N-1}^2 \sum_{i=1}^{N-1} (T_i - M)$, the requirement that

$$\Omega_{1}^{ND}(T_1, ..., T_{N-1}) \subseteq \Omega_1^f(T_1, ..., T_{N-1})$$

is satisfied. Therefore, we set $C_{N-1} \equiv C^*$. The expression in (49), the price reaction to non-disclosure of one signal, then becomes:

$$\kappa_{N-1}^2 \sum_{i=1}^{N-1} (Y_i - M) + L_1,$$

with $L_1 \equiv \kappa_{N}^2 C^*$. The manager then indeed prefers to disclose $Y'_N$ if and only if $(Y'_N, I_N) \in \Omega_1^f \setminus \Omega_{1}^{ND}$. We have thus specified $L_1$ and $C_{N-1}$.

Now assume that we have constructed the non-disclosure penalty amounts $L_{N-d-1}, ..., L_1$ and cutoffs $C_{d+1}, ..., C_{N-1}$, which characterize the firm’s value on the equilibrium path if it discloses anywhere from $d+1$ to $N-1$ signals. We will now simultaneously characterize the situations in which the manager chooses to disclose exactly $d$ highest signals and the firm’s value based on such disclosures. Again, without loss of generality, let $Y_{d+1}, ..., Y_N$ denote the undisclosed signals.

Conditional on the values of the first $d$ signals, $Y[d]$, consider random variable

$$Y'[−d] \equiv Y[−d] | Y[d] − E[Y[−d] | Y[d]],$$

where $Y[−d] = (Y_{d+1}, ..., Y_N)$. In other words, this random variable is defined according to equations (38) and (37), has a mean of zero and a covariance matrix which depends only on $\sigma_2^2$ and $\sigma_1^2$. Let $G'[−d]$ denote the 2 $(N-d)$-dimensional variable with the first $(N-d)$ coordinates equal to $Y'[−d]$ and the second $(N-d)$ coordinates equal to $I[−d]$:

$$G'[−d] = \left( Y'[−d], I[−d] \right),$$

where $I[−d] \equiv (I_{d+1}, ..., I_N)$. Conditional on $T_1, ..., T_d$ being the highest signals, $G'[−d]
belongs to the following set \( \Omega^f_{N-d}(T_1, \ldots, T_d) \):

\[
\Omega^f_{N-d}(T_1, \ldots, T_d) \equiv \left\{ G' [-d] | \forall d + 1 \leq j \leq N, Y'_j < T_d - M - \kappa^2_d \sum_{i=1}^{d} (T_i - M) \text{ or } I_j < 0 \right\}
\] (51)

Let \( \Omega^{ND}_{N-d}(T_1, \ldots, T_d) \subset \Omega^f_{N-d}(T_1, \ldots, T_d) \) denote the region in which, in equilibrium, the manager leaves \( N - d \) signals undisclosed (we label this event \( ND_{N-d} \)). By Lemma 2, the price reaction to this event is then given by:

\[
\Delta P(ND_{N-d}) = \kappa^2_d \sum_{i=1}^{d} (T_i - M) + \kappa^2_N \mathbb{E} \left[ \sum_{i=d+1}^{N} Y'_i | G' [-d] \in \Omega^{ND}_{N-d}(T_1, \ldots, T_d) \right].
\]

Therefore, the value of \( L_{N-d} \) we are looking for must satisfy:

\[
L_{N-d} = \mathbb{E} \left[ \sum_{i=d+1}^{N} Y'_i | G' [-d] \in \Omega^{ND}_{N-d}(T_1, \ldots, T_d) \right]. \tag{52}
\]

For each point in \( \Omega^{ND}_{N-d} \), the disclosure of \( d \) signals must be preferred to the disclosure of any \( j > d \) signals. By the induction hypothesis, for \( j > d \), the price reaction to the disclosure of \( j \) highest signals (event \( ND_{N-j} \)) is given by:

\[
\Delta P(ND_{N-j}) = \kappa^2_j \sum_{i=1}^{j} (T_i - M) + L_{N-j}. \tag{53}
\]

Observe that

\[
\kappa^2_d - (j - d) \kappa^2_j \kappa^2_d = \frac{\sigma^2_V}{d\sigma^2_V + \sigma^2_\varepsilon} \left( 1 - (j - d) \frac{\sigma^2_\varepsilon}{j\sigma^2_V + \sigma^2_\varepsilon} \right) = \frac{\sigma^2_V}{j\sigma^2_V + \sigma^2_\varepsilon} = \kappa^2_j.
\]

Therefore, \( \Delta P(ND_{N-j}) \) in (53) can be rewritten as:

\[
\Delta P(ND_{N-j}) = \kappa^2_d \sum_{i=1}^{d} (T_i - M) + \kappa^2_j \sum_{i=d+1}^{j} (T_i - M - \kappa^2_d \sum_{i=1}^{d} (T_i - M)) + L_{N-j}
\]

\[
= \kappa^2_d \sum_{i=1}^{d} (T_i - M) + \kappa^2_j \sum_{i=d+1}^{j} T_i + L_{N-j}, \tag{54}
\]

38
where $T'_{d+1}, \ldots, T'_{N}$ are the values of $Y'_{d+1}, \ldots, Y'_{N}$ ordered from largest to smallest and with the values not received by the manager (the corresponding $I_i < 0$) replaced by $-\infty$.

Recall further that we are constructing the equilibrium under the assumption that the manager only discloses $j$ signals if

$$T_j \geq C_j + M + \kappa^2_j \sum_{i=1}^{j} (T_i - M),$$

which is equivalent to:

$$T'_j \geq C_j + \kappa^2_j \sum_{i=d+1}^{j} T'_i. \quad (55)$$

Using expressions (54) and (55), we obtain that the disclosure of $d$ signals leads to a higher payoff than that of any greater number of signals if $L_{N-d}$ satisfies:

$$L_{N-d} \geq \max_{\Omega^{ND}_{N-d}(T_1, \ldots, T_d)} \left\{ \kappa^2_j \sum_{i=d+1}^{j} T'_i + L_{N-j} \mid j > d \text{ and } T'_j \geq C_j + \kappa^2_j \sum_{i=d+1}^{j} T'_i \right\}. \quad (56)$$

Furthermore, $\Omega^{ND}_{N-d}(T_1, \ldots, T_d)$ must be the maximal set for which the inequality above holds: for any point in $\Omega^{I}_{N-d} \setminus \Omega^{ND}_{N-d}$ there must exist a $j > d$ such that

$$L_{N-d} < \kappa^2_j \sum_{i=d+1}^{j} T'_i + L_{N-j}$$

and $T'_j \geq C_j + \kappa^2_j \sum_{i=d+1}^{j} T'_i$. We now need to prove that there exist a number $L_{N-d}$ and set $\Omega^{ND}_{N-d}(T_1, \ldots, T_d)$ such that (56) is satisfied simultaneously with equation (52), which we reproduce below for readers’ convenience:

$$L_{N-d} = \mathbb{E} \left[ \sum_{i=d+1}^{N} Y'_i \mid G' [-d] \in \Omega^{ND}_{N-d}(T_1, \ldots, T_d) \right]. \quad (57)$$

In addition, $\Omega^{ND}_{N-d}(T_1, \ldots, T_d)$ must be a subset of $\Omega^{I}_{N-d}(T_1, \ldots, T_d)$ defined in (51).

We will first show that a solution to (56) and (57) exists in the unconstrained problem, i.e., when $T_d \to \infty$ and $\Omega^{I}_{N-d}(T_1, \ldots, T_d)$ is $\mathbb{R}^{2(N-d)}$. Then, we will show that there is a
constant $C_d$ such that our candidate solution is contained in $\Omega^f_{N-d}(T_1, \ldots, T_d)$ as long as

$$T_d \geq C_d + \kappa_d^2 \sum_{i=1}^d T_i.$$

For a given value of $L_{N-d}$, let $\Omega^{ND}_{N-d}(L_{N-d})$ denote the maximal subset of $\mathbb{R}^{2(N-d)}$ satisfying

$$L_{N-d} \geq \max_{\Omega^{ND}_{N-d}(L_{N-d})} \left\{ \kappa_j^2 \sum_{i=d+1}^j T'_i + L_{N-j} | j > d \text{ and } T'_j \geq C_j + \kappa_j^2 \sum_{i=d+1}^j T'_i \right\}. \quad (58)$$

Three observations are in order about $\Omega^{ND}_{N-d}(L_{N-d})$. First, since the distributions of $Y_{d+1}', \ldots, Y_N'$, and therefore, of $T_{d+1}', \ldots, T_N'$ do not depend on $M$ or $\{T_i\}_{i=1}^d$, $\Omega^{ND}_{N-d}(L_{N-d})$ does not depend on these values either.

Second, when $L_{N-d} \to -\infty$,

$$\Omega^{ND}_{N-d}(L_{N-d}) \to \left\{ \mathcal{G}' [-d] | \forall j > d, T'_j < C_j + \kappa_j^2 \sum_{i=d+1}^j T'_i \right\}.$$

We will denote the set in the right-hand side above $\Omega^{ND}_{N-d}(-\infty)$. Intuitively, if the penalty for disclosing exactly $d$ signals is very high, then the only non-disclosing firms will be those who either do not have any more signals to disclose ($T_j = -\infty$) or who have only “non-disclosable” signals, which are below their respective cutoffs derived from the values of $\{C_j\}_{j=1}^N$. Finally, when $L_{N-d} \to \infty$, $\Omega^{ND}_{N-d}(L_{N-d}) \to \mathbb{R}^{2(N-d)}$: if the firms who disclose exactly $d$ signals are assigned infinite valuations, then no firm will disclose more than $d$ signals.

Now note that both $\mathbb{E} \left[ \sum_{i=d+1}^N Y_i' | \mathcal{G}' [-d] \in \Omega^{ND}_{N-d}(-\infty) \right]$ and $\mathbb{E} \left[ \sum_{i=d+1}^N Y_i' | \mathcal{G}' [-d] \in \Omega^{ND}_{N-d}(\infty) \right]$ are finite numbers. By continuity, it follows that equation

$$L_{N-d} = \mathbb{E} \left[ \sum_{i=d+1}^N Y_i' | \mathcal{G}' [-d] \in \Omega^{ND}_{N-d}(L_{N-d}) \right]$$

has a solution, which we will denote by $L^*_{N-d}$. This number does not depend on $M$ or $\{T_i\}_{i=1}^d$.

It remains to verify that $C_d$ can be chosen such that $\Omega^{ND}_{N-d}(L_{N-d}) \subset \Omega^f_{N-d}(T_1, \ldots, T_d)$. To this end, note that the definition of $\Omega^{ND}_{N-d}(L_{N-d})$ in (58) implies that for any point in
this set, either $T_{d+1}' < C_{d+1} + \kappa_{d+1}^2 T_{d+1}'$ or

$$L_{N-d}^* \geq \kappa_{d+1}^2 T_{d+1}' + L_{N-d-1}.$$  

Then,

$$T_{d+1}' \leq \max \left\{ \frac{C_{d+1}}{1 - \kappa_{d+1}^2}, \frac{L_{N-d}^* - L_{N-d-1}}{\kappa_{d+1}^2} \right\},$$

and, therefore, setting $C_d$ equal to the right-hand side above ensures $\Omega_{N-d}^{ND} (L_{N-d}) \subset \Omega_{N-d}^f (T_1, ..., T_d)$. This completes the construction of $(C_1, ..., C_{N-1})$ and $(L_1, ..., L_N)$.

To conclude the proof, we need to specify off-equilibrium pricing that ensures that the manager never reports $d < N$ signals if

$$T_d < M + \kappa_d^2 \sum_{i=1}^{d} (T_i - M) + C_d.$$  

Suppose that $d$ signals are reported and the inequality above holds. Let $P$ denote the firm’s price had it reported the best subset of $d$ signals using an equilibrium strategy:

$$P (T_1, ..., T_d) = \max \left\{ M + \kappa_j^2 \sum_{i=1}^{j} (T_i - M) + L_{N-j} | j < d \& T_j < M + \kappa_j^2 \sum_{i=1}^{j} (T_i - M) + C_j \right\}.$$  

Assume that if investors observe an off-equilibrium report consisting of $d$ signals, they believe that the remaining $N-d$ signals are sufficiently low so as to warrant a price of $P (T_1, ..., T_d) - 1$. Then, the manager will never find it optimal to use an off-equilibrium disclosure strategy.

---

30Recall that, by construction, $C_N = -\infty$, so full disclosure is always admissible in equilibrium.