The Dynamics of Concealment: CEO Myopia and Information Withholding

Jeremy Bertomeu, Iván Marinovic, Stephen J. Terry, and Felipe Varas *

November 10, 2015

Abstract

We propose a dynamic model of corporate disclosure in which managers’ information endowment is uncertain. We consider a realistic environment in which corporate disclosures are preceded by analyst’ consensus forecasts and followed by earnings announcements. We estimate the model using a comprehensive sample of annual management forecasts released by public U.S. companies between 2004 and 2014. We find that managers withhold information strategically 26% of the time, leading to an information loss of 6%, relative to a setting where managers never hide bad news from investors. Information endowments are persistent, creating reputation concerns that result in highly sticky disclosure policies.

Keywords: voluntary disclosure, structural estimation
JEL Classification: D72, D82, D83, G20

VERY PRELIMINARY AND INCOMPLETE

*We thank Oeystein Daljord for very helpful comments.
1 Introduction

In the U.S., public firms release their earnings each quarter, as required by the Securities and Exchange Commission. Accounting earnings are limited, to a large extent, to past transactions and exclude most future transactions such as, most notably, expected new revenue and costs. For this reason, many firms supplement their mandatory accounting reports with quantitative disclosures about future earnings, or management forecasts. Such voluntary forecasts appear to carry most of the price-relevant news; for example, Beyer et al. (2010) document that management forecasts explain a significant portion of the firm’s quarterly return variance, about 16%, more than analyst forecasts, earnings announcements and regulatory filings added together!

This paper develops a simple, but rigorous theory, of voluntary management forecasts that can quantitatively explain the empirical distribution of forecasts and earnings. Our theory attempts to jointly explain a number of stylized facts about forecasts. First, while the traditional unravelling theory predicts that all firms should make a forecast (Viscusi 1978; Grossman and Hart 1980; Milgrom 1981), we observe a fairly low propensity to forecast. To be specific, even for the S&P500 index in which firms have well-developed investor relations, a forecast is made only between one half and one third of the time. Second, we observe significant stickiness in propensities to disclose; a firm is four to five times more likely to make a forecast after making a forecast in a past quarter. Third, managers appear to select periods in which information is more favorable to make a forecast, and strategically withhold information (Kothari, Shu and Wysocki 2009). Further, markets react unfavorably to firms that stop giving forecasts, and such decisions usually precede declines in actual earnings.

This paper attempts to rationalize these stylized facts as an optimal response to a disclosure friction and build a simple theory of management forecasts that can quantitatively explain the empirical distribution of quarterly earnings and forecasts. We enrich the model of Dye (1985) and Jung and Kwon (1988), hereafter DJK, a static model in which a manager may not be endowed with information. In this model, the manager is interested in maximizing short-term prices; absent a forecast, the manager is unsure as to whether the manager is truly uninformed or avoiding an unfavorable to forecast. This model implies a threshold below which an informed manager chooses to withhold information.

To make it amenable to an empirical analysis, we make this model fully dynamic, enriching it with several insights from recent disclosure theory: (a) the manager is forward-looking and considers the effect of a disclosure on future prices (Guttman, Kremer and Skrzypacz 2014; Marinovic, Skrzypacz and Varas 2015), (b) the disclosure friction is serially-correlated and markets update their beliefs using realized earnings (Einhorn and Ziv 2008), and (c) a public news process, e.g., the analyst consensus, affects the propensity to disclose (Acharya, Demarzo and Kremer 2011).

These assumptions have several intuitive implications for the time-series properties of earnings and forecasts. A forward-looking manager benefits from a reputation to be uninformed, since the market discounts a non-disclosure less when the market believes strategic withholding is unlikely. By making a forecast, the manager reveals a state of information endowment. The loss of reputation that follows generates, in our model, an endogenous cost of disclosure, especially after long periods without a forecast. That is, the longer the period of time without a forecast, the more costly it is for an informed manager to make a forecast.
Vice-versa, a manager who stops making forecasts is viewed as likely to be informed and, therefore, bears a large negative market reaction conditional on non-disclosure. This feature ties together forecast stickiness to the observed self-selection in forecasts. A firm will tend to self-select more favorable forecasts after long periods without forecasts and when, on average, its forecasts are infrequent.

A more complete empirical understanding of management forecasts has the potential to assess several deep questions about voluntary disclosure in capital markets. To our knowledge, there is limited empirical evidence that voluntary disclosure theory jointly and quantitatively explains many stylized facts about observed voluntary disclosures. Our approach offers new empirical evidence that the theory is useful to explain these facts, within a setting that is of first-order relevance financial markets.

One key implication of the theory is that the information contained in regulated earnings affects voluntary disclosure. Within traditional disclosure theory, voluntary disclosure carries a negative social externality, because while it increases the price of the disclosing party, it reduces the equilibrium non-disclosure price. There tends to be excessive disclosure of information and, if market participants are rational, there is limited scope for regulations that mandate more disclosure (Shavell 1994; Fishman and Hagerty 1990, 2002). Adding to this, a large body of empirical evidence from capital market research suggests that the news contained in mandatory reports is largely known before it is released in regulated statements (Ball and Brown 1968; Beyer et al. 2010). This has led to many long-standing debates about the usefulness of regulated accounting, as imposed in the U.S. after the S.E.C. Act of 1934 (Stigler 1964; Benston 1973; Sunder 1988; Bushee and Leuz 2005).

Our approach fleshes out a new channel through which mandatory disclosure affects the provision of voluntary disclosure. To begin with, we can precisely measure the information known to the manager relative to the information at the earnings release date. Even if actual earnings contain little news, it serves to discipline forecasting behavior and serves as a complement to voluntary disclosure. Actual earnings inform the market about the observed private information by the manager. Hence, they reduce a manager’s ability to build a reputation to be uninformed. When observing low actual earnings, the market knows that the manager was more likely to be informed, and will discount more future non-disclosure; the greater the information contained in earnings, the less the manager will be able to maintain a reputation.

Survey evidence suggests that reputation concerns are a major determinant of a management disclosure policy Graham, Harvey and Rajgopal (2005). In the model, such concerns arise because information endowments are persistent. That is, the manager is more likely to receive information in the future, if he received information in the past. Consistent with this, the market updates its beliefs about the evolution of the managers’ information endowment based on disclosure history. This, fact provides the manager with incentives to manipulate disclosure policies to influence the market perceptions about the likelihood of subsequent disclosures. Indeed, withholding information may lead to a lower price today, but it softens

\[1\]

In their classic text, Ball and Brown (1968) show that the return variance during at typical news release is not greater than in a typical trading day; furthermore, price movements that precede the earnings release predict actual earnings extremely well. Using more recent data, Beyer, Cohen, Lys and Walther (2010) calculate that earnings explain only 2% of stock return variance, which is consistent with the earlier evidence from Ball and Brown.
the negative market reaction to non-disclosure in the future. The presence of reputation concerns thus has the ability to explain, at the same time, the high persistence and low frequency of disclosures. Indeed we strongly reject the hypothesis of an IID disclosure friction. Managers are x times more likely to disclose their signals if they did disclose in the previous period.

In the model reputation concerns are inherently linked to the persistence of managers’ information endowments. In the absence of persistence, the manager uses a myopic disclosure policy. By contrast, with persistence the manager has stronger incentives to withhold information. In fact, the more the manager cares about future stock prices, the more he will withhold information to cash-in his reputation in due time. But the effect of persistence on the likelihood of disclosure is subtle. A higher level of persistence, leads in steady state, to a higher dispersion in the distribution of market beliefs about the information endowment: in the limit, the market is almost certain that the manager is informed or almost certain that the manager is uninformed. This dispersion effect reinforces unraveling pressure sometimes offsetting the reputation effect and leading to more disclosure.

We estimate the model using I/B/E/S data. Our sample represents an unbalanced panel of 1,043 US public firms spanning 2004-2014 with 8,095 firm-fiscal year observations. Our results indicate that managers are informed 85% of the time, but they strategically withhold information 26% of the time: namely they hide the bottom 26% of the information. A back of the envelope calculation reveals that withholding leads to a 6% information loss (measure by the Root Mean Square Error of market earnings expectations) relative to a setting where managers do not strategically withhold bad news.

We also study the asset pricing implications of strategic disclosure and show that a higher incidence of strategic withholding is associated with greater return volatility at the earnings announcement date. Empirically we find that the standard deviation of stock returns is 5% higher if the earnings announcement was not preceded by a management forecast. Furthermore, we show that strategic disclosure cannot explain the asymmetric market reaction to bad news documented by Kothari, Shu and Wysocki (2009) but it does lead, on average, to a higher earnings response coefficient ERC.

**Related literature.** For the most part, the disclosure literature is static (Jovanovic 1982; Verrecchia 1983; Dye 1985; Shin 1994; Shavell 1994). This feature can be understood from the early roots of disclosure theory. In the defining example of this theory, a seller places a product for sale, releases some information prior to a sale, a price forms and the seller leaves with any proceeds from the sale. In his classic treatment of the theory, Milgrom (1981) uses the terminology of **persuasion** to describe disclosure problems. Formally, within the strict definition of persuasion introduced by Milgrom, the seller’s preference can be written only in terms of the posterior beliefs induced by a message. The Dye (1985) model is an example of persuasion because the seller wants to maximize the selling price, regardless of what private signal is received. Put differently, in a persuasion game, the private information received by the seller affects which messages can be feasibly sent, but not the utility of the seller for a given message and a given market posterior.

2This was later generalized to preferences in which the seller’s preference may also be a function of the message used (e.g., a disclosure cost), see Jovanovic (1982) and Verrecchia (1983).
This theory has only been recently been examined in dynamic context, because most single-period disclosure models no longer fit this definition of persuasion when they are repeated. Three recent examples of two-period disclosure models serve to best illustrate this point. Acharya, Demarzo and Kremer (2011) develop a model in which the manager can delay a disclosure, after the release of public news. In this model, the value of delay is a function of the manager’s expectation about the public news, which depends on the information privately observed by the manager. This form of type-dependent preference is not “allowed” under traditional persuasion theory. Beyer and Dye (2012) consider a model in which some managers may be forthcoming, and disclose all of their information. Managers know their propensity to be forthcoming in the next period, which implies that their type affects their payoff from withholding. Guttman, Kremer and Skrzypacz (2014) examine a model in which more information may be received at some later date. A manager choosing to delay can anticipate the information received at a later date - managers with a low signal know it is more likely that they will be willing to withhold later - thus, the value of delay is type-specific. The traditional methods of persuasion theory fail to apply, or at least are significantly changed, for repeated versions of these models.

One approach has been shown to preserve the basic structure of persuasion with some dynamic aspects. Einhorn and Ziv (2008) and Marinovic (2012) are models in which the market updates dynamically to new information, but the manager is myopic and maximizes only current short-term stock prices. The stock price is the sole channel through which future periods affects current actions. In Einhorn and Ziv (2008), the stock price includes the discounted value of expected disclosures in the future. In Marinovic (2012), the stock price is equal to the value implied by all past disclosures. After observing a current disclosure, the market reassesses the propensity to be truthful and reprices the entire sequence of past disclosures. In this class of models, Bertomeu, Ma and Marinovic is the paper closest to ours, and estimates the DJK model under the assumption that the manager, when informed, maximizes current stock prices. This paper generalizes this approach to a manager with forward-looking motives and adds an exogenous public news process, analysts’ forecasts, to the estimation procedure.

Our model shares its focus with a recent literature on the structural models applied to strategic aspects of financial reporting. Beyer, Guttman and Marinovic (2014) structurally estimate a dynamic model of costly earnings misreporting, in which the manager’s equity incentives are not known to the market (Fischer and Verrecchia 2000; Frankel and Kartik 2014). Zakolyukina (2014) examines a model in which managers trade-off higher stock prices in the present versus a greater probability of prosecution in the future. Both studies focus on earnings only and explicitly model the decision to misreport earnings for a cost. While our approach does not require an explicit that earnings are truthfully reported, that is, reported earnings need not perfectly reflect underlying fundamentals, we do not attempt to structurally explain distortion to earnings reports or interactions between earnings misreporting and voluntary disclosure.
2 Model

This section extends the model in Dye (1985) to a dynamic setting in which the manager has forward-looking motives, and public information flows interact with management disclosure choices. There is a single firm and a mass of risk neutral investors, referred to as the market. Time is discrete and indexed by \( t = \{1, 2, ..., \infty\} \). Following Acharya, Demarzo and Kremer (2011), Benmelech, Kandel and Veronesi (2010) and Beyer (2012) we assume the manager maximizes the present value of the firm’s future stock prices. That is, the manager expected utility in period \( t \) is

\[
U_t = E(\sum_{n=t}^{\infty} \beta^{n-t} P_n | I_t),
\]

where \( \beta \) is the manager’s discount factor, \( P_n \) is the market price of the firm and \( I_t \) is the manager’s information set.

2.1 Timeline

In each period there are three dates. First, the market observes a signal \( c_t \) of the firm’s earnings, \( e_t \). We refer to \( c_t \) as the consensus forecast, but we can think of \( c_t \) more generally as comprising all the public information available to investors prior to the firm’s disclosure. Second, with some probability (described below) the manager observes a private signal \( s_t \) of the firm’s future earnings and chooses whether to disclose this signal or withhold it. Based on the manager’s disclosure choice, the market sets the stock price \( P_t \) as the expected present value of the firm’s future earnings. Finally, the firm’s earnings \( e_t \) are realized and publicly observed.

Figure 1 summarizes the timeline.

Figure 1. Timeline

\[
\begin{array}{ccc}
  t=1 & t=2 & t=3 \\
  \text{The consensus } c_t \text{ is} & \text{The manager privately} & \text{Earnings } e_t \\
  \text{publicly released.} & \text{observes } s_t \text{ and} & \text{are released.} \\
  & \text{chooses } d_t. & \\
  & \text{The price } P_t \text{ is set.} & \\
\end{array}
\]

2.2 Information Structure

Following Dye (1985) we assume that the manager’s information endowment \( \theta_t \in \{0, 1\} \) is random. In particular, when \( \theta_t = 1 \) the manager observes the signal \( s_t \). Formally, following Einhorn and Ziv (2007) we model the manager’s information endowment \( \theta_t \) as a hidden Markov chain with transition matrix

\[
\Pi = \begin{bmatrix}
1 - \lambda_1 & \lambda_1 \\
\lambda_0 & 1 - \lambda_0
\end{bmatrix},
\]
where $\lambda_0$ denotes the probability of transitioning from the informed to the uninformed state, and $\lambda_1$ denotes the probability of transitioning from the uninformed to the informed state. The information endowment is persistent if becoming uninformed is less likely than remaining uninformed, or $\lambda_0 < 1 - \lambda_1$. We will sometimes parametrize the process of $\theta$ using the long-run probability of being uninformed given by $\bar{p} \equiv \frac{\lambda_0}{\lambda_0 + \lambda_1}$, and the persistence of information endowments $r \equiv 1 - \frac{1}{\bar{p}}$, where $\lambda_0 < \bar{p}$.

For simplicity, and following a long tradition in accounting and economics we model the firm’s earnings as an AR1 process (see Ohlson (1995); Gerakos and Kovrijnykh (2013)).

Let the process of earnings, consensus, and manager signal be given by:

\[
\begin{align*}
e_t &= \rho e_{t-1} + u_t \\
c_t &= e_t + n_t \\
s_t &= e_t + v_t,
\end{align*}
\]

where each coordinate of $\varepsilon_t = (u_t, n_t, v_t)$ is iid normally distributed. The assumption that $\varepsilon_t$ is iid is not required but it simplifies the analysis without significantly compromising the model’s realism. We shall consider the case when there is no public signal, namely when $\text{Var}(n_t) \rightarrow \infty$. In this case, we omit $c_t$ from the inputs below. The variance covariance matrix of $\varepsilon_t$ will be denoted $\Sigma$.

### 3 Discussion and Interpretation

**Manager preferences** The manager maximizes the present value of future stock prices. These preferences may capture incentives arising from stock based compensation. Edmans, Goncalves-Pinto, Wang and Xu (2014) show that CEOs strategically release news in months in which their equity vests, to boost the stock price and stock liquidity. In this model, the manager does not care exclusively about the current stock price, as in static models, but also future ones. This concern for future prices, rather than simply the current stock price, is natural given that compensation contracts typically specify long term incentives that are linked to future stock prices (see e.g. Core and Larcker (2002)).

**Information endowment** We model the manager’s information endowment as a hidden Markov chain, following Einhorn and Ziv (2008). This process represents more broadly the possibility that in some years managers don’t observe a precise signals of the firm future earnings, in which case they are forced to withhold their forecasts. Alternatively, we can think of $\theta_t$ as capturing the manager’s trading activity: in some periods managers need to sell shares for liquidity reasons, and hence they have incentives to withhold bad news.

**Public information** Empirical evidence suggests that public information, particularly bad news, trigger an unusually high number of voluntary disclosures. Acharya, Demarzo and Kremer (2011) refer to this phenomenon as disclosure clustering. In a voluntary disclosure context, there are two main sources of public information interacting with the firm’s

---

3 This can be easily generalized, given that earnings are observable, provided that the earnings innovation is iid.
voluntary disclosures: mandatory financial reports such as earnings announcements, and in-
formation released by external parties, notably financial analysts. The former is particulary
important because it allows the market to update perceptions regarding whether managers
who withheld information in the past did it for strategic reasons. The latter source is also
important because it sets market expectations prior to the voluntary disclosure choice.

Reputation concerns Disclosure behavior may be affected by reputation concerns. For
example Beyer and Dye (2012) argue that managers may improve their reputation for being
forthcoming by disclosing the firm’s prospects often. By doing so managers can improve their
ability to hide information in bad times by softening the market reaction to non-disclosure.
In our setting, when the information endowment is persistent, the manager also experiences
reputation concerns, but of opposite sign. The manager seeks to build reputation for being
uninformed so to inflate the market’s perception that the manager will be uninformed in the
future. This captures in a reduced form the notion that managers wish to avoid setting a
precedent of high disclosure frequency that could make them liable before investors in the
future.

Disclosure bias Disclosure theory assumes that manager’s private information is verifi-
able, which means that managers can conceal it but they cannot lie about it. This assumption
may seem unrealistic in management forecast settings. We ignore the possibility of forecast
manipulation for two empirical reasons. First, management forecasts are highly informa-
tive and induce important market reactions, even more than earnings announcements, which
suggests forecasts are not cheap talk (Beyer, Cohen, Lys and Walther (2010)). And though
they may be biased, one can easily accommodate a constant bias in the econometric model
without altering the qualitative nature of the model. Second, in practice manager’s forecasting
behavior is relatively irregular/intermitent which is an inherent property of voluntary
disclosure models.

Endogenous timing We ignore the possibility that managers have discretion over the
timing of their disclosures, being able to postpone their forecasts. To mitigate concerns
about the endogeneity of disclosure timing, in our empirical implementation we focus on so
called bundle forecasts which are released in conjunction with the firm’s financial statements.
Also we focus on annual forecasts (which are release at the beginning of the fiscal year) for
which the timing choice is is less relevant than in a quarterly setting.

Litigation risk Ever since Skinner (1997), the literature has debated about the role of
litigation risk as a determinant of the firm’s disclosure incentives. Skinner (1997) argue that
firms use disclosures to preempt litigation, by advancing the release of bad news, whereas oth-
ers contend that voluntary disclosures trigger litigation, especially if the subsequent earnings
realization is too disappointing relative to the disclosure (see Francis, Philbrick and Schipper
(1994)). Modeling litigation risk in depth is beyond the scope of this paper, but one can
easily adapt the model to accommodate a cost of no disclosure.
4 Analysis

Some notation is in order. Let \( z_t = (e_{t-1}, c_t) \) be the public information available at the beginning of period \( t \) and \( z^t = \{ z_0, \ldots, z_t \} \) the public history up to time \( t \). We consider Markov Perfect Equilibria (MPE). In a MPE the payoff relevant public information is given by \( (p_t, z_t) \) and the payoff relevant information of the manager is \( (\theta_t, z_t, s_t) \). For any public state, \( (p, z) \), let \( D(p, z) \equiv \{ s \in \mathbb{R} | d(p, z, s) = 1 \} \) be the manager disclosure set, when the manager disclosure strategy is \( d(p, z, \cdot) \), and \( D^c(p, z) \) its complement.

Let \( P^D(z, s) \) be the price given disclosure and \( P^{ND}(p, z) \) the price of non-disclosure. Now, for the price to be consistent with Bayes’ rule and the manager’s disclosure strategy, we must have:

\[
P^D(z, s) = \frac{\mathbb{E}(e'|z, s)}{1 - \beta \rho}, \tag{1}
\]
\[
P^{ND}(p, z) = \frac{1 - p}{1 - \beta \rho} \frac{p \mathbb{E}(e'|z) + (1 - p) \mathbb{E}(e'|D^c(p, z)|z)}{p + (1 - p) \mathbb{E}(1_{D^c(p, z)}|z)}. \tag{2}
\]

This price function assumes that prices are given by the present value of the firm’s economic earnings. In essence, we are assuming that the firm is not growing, as discussed in Section A.1.

The market reassesses the probability that the manager was informed in the current period based on the earnings realization \( e' \). Conditional on non disclosure and the earnings realization \( e' \), the updated probability that the manager will be uninformed in the next period is given by:

\[
p' = \varphi(p, z, e') \equiv \frac{p(1 - \lambda_1) + (1 - p) \lambda_0 E(1_{D^c(p, z)}|e')}{p + (1 - p) E(1_{D^c(p, z)}|e')}. \tag{3}
\]

When the manager withholding his signal he retains an informational advantage about the firm’s fundamentals, but also about his expected information endowment. Indeed, given non-disclosure, investors cannot tell whether the manager strategically withheld unfavorable information or was uninformed. By contrast, if the manager discloses his signal, the market learns, of course, that the manager was informed and updates the probability that the manager will be uninformed accordingly to \( \lambda_0 \).

We can now define the manager’s optimization problem, when informed, as:

\[
V_1^D(p, z, s) = P^D(s, z) + \beta E \left[ (1 - \lambda_0) V_1(\lambda_0, z', s') + \lambda_0 V_0(\lambda_0, z') \right] | z, s \tag{4}
\]
\[
V_1^{ND}(p, z, s) = P^{ND}(p, z) + \beta E \left[ (1 - \lambda_0) V_1(p', z', s') + \lambda_0 V_0(p', z') \right] | z, s \tag{5}
\]
\[
V_1(p, z, s) = \max_{d \in \{0,1\}} \left[ d V_1^D(p, z, s) + (1 - d) V_1^{ND}(p, z, s) \right] \tag{6}
\]
\[
V_0(p, z) = P^{ND}(p, z) + \beta E \left[ \lambda_1 V_1(p', z', s') + (1 - \lambda_1) V_0(p', z') \right] | z \tag{7}
\]

Consider the market posterior expectations evaluated at alternative information sets. Let \( F(e'\cdot), F(s\cdot) \) denote the cumulative distributions of \( e' \) and \( s \) conditional on \( \cdot \). In order to compute the equilibrium, we need the following distributions: \( F(e'|z), F(e'|z, s), F(s|z) \) and
For any random variable $x$ let $\tau_x \equiv 1/\sigma_x^2$ be the precision; then, by Bayes’ rule, the beliefs are

$$e'|z \sim \mathcal{N}\left(\frac{\tau_u\rho e + \tau_n c}{\tau_u + \tau_n}, \frac{1}{\tau_u + \tau_n}\right)$$

$$e'|z, s \sim \mathcal{N}\left(\frac{\tau_u\rho e + \tau_n c + \tau_v s}{\tau_u + \tau_n + \tau_v}, \frac{1}{\tau_u + \tau_n + \tau_v}\right)$$

$$s|z \sim \mathcal{N}\left(\frac{\tau_u\rho e + \tau_n c}{\tau_u + \tau_n}, \frac{1}{\tau_u + \tau_n}\right)$$

$$s|e' \sim \mathcal{N}\left(e', \frac{1}{\tau_v}\right)$$

We can now formally define an equilibrium.

**Definition 1.** A Markov Perfect Equilibrium is a tuple $\langle P^D, P^{ND}, d, \varphi, V_0, V_1^D, V_1^{ND}\rangle$ such that

1. The market price is

   $$P = \begin{cases} 
   P^{ND}(p, z) & \text{if } d(p, z, s) = 0 \\
   P^D(z, s) & \text{if } d(p, z, s) = 1,
   \end{cases}$$

   where $P^D$ and $P^{ND}$ are given by (1) and (2).

2. The disclosure strategy $d(p, z, s) \in \{0, 1\}$ is

   $$d(p, z, s) \in \arg\max_{d \in \{0, 1\}} \left[ dV_1^D(p, z, s) + (1-d)V_1^{ND}(p, z, s) \right],$$

   if the manager is informed, and is $d(p, z, s) = 0$ if the manager is uninformed.

3. The evolution of market beliefs is

   $$p' = \begin{cases} 
   \varphi(p, z, e') & \text{if } d(p, z, s) = 0 \\
   \lambda_0 & \text{if } d(p, z, s) = 1,
   \end{cases}$$

   where $\varphi(p, z, e')$ is given by (3).

4. The value function of the informed manager solves

   $$V_1(p, z, s) = \max \{ V_1^D(z, s), V_1^{ND}(p, z, s) \},$$

   where

   $$V_1^D(z, s) = P^D(s, z) + \beta E\left[ (1-\lambda_0)V_1(\lambda_0, z', s') + \lambda_0 V_0(\lambda_0, z') | z, s \right]$$

   $$V_1^{ND}(p, z, s) = P^{ND}(p, z) + \beta E\left[ (1-\lambda_0)V_1(p', z', s') + \lambda_0 V_0(p', z') | z, s \right]$$
5. The value function of the uninformed manager solves

\[ V_0(p, z) = P^{ND}(p, z) + \beta E \left[ \lambda_1 V_1(p', z', s') + (1 - \lambda_1) V_0(p', z') | z \right] \]

The equilibrium definition is general. In particular, it allows for non-threshold equilibria: the disclosure set \( D \) can be the union of multiple disconnected sets. This is important, as shown below, because the existence of threshold equilibria is not guaranteed when the information endowment is persistent.

5 Equilibrium

In this section, we discuss the main economic forces behind the equilibrium. First, consider the manager’s disclosure/withholding incentives. Withholding information has two benefits from the manager’s standpoint. First, the classic effect: by hiding bad news, the manager can delay the decline in the stock price, given that the market is always uncertain about the real motivation behind non-disclosures. Second, by withholding information, the manager may influence the market’s perception about his future information endowment. This is valuable too: by pretending to be uninformed, the manager can increase the perceived probability that he will be uninformed in the future. This, in turn, increases the option value of withholding information in the future, because it mitigates the price penalty triggered by non-disclosure. Withholding information entails thus a reputational benefit.

Naturally, when the information endowment is \( iid \), namely \( \lambda_0 = 1 - \lambda_1 \), the reputation benefit of withholding is absent. In this case, market beliefs are constant and independent of the manager’s disclosure choices. For this reason, the manager’s disclosure strategy boils down to that of the static model. We provide this result as a benchmark.

**Proposition 2.** When the information endowment is \( iid \), \( \lambda_0 = 1 - \lambda_1 \) there is a unique equilibrium where, in each period, the manager uses the myopic threshold \( \tau \) defined as

\[ -\frac{\bar{p}}{1 - \bar{p}} \tau = \int_{-\infty}^{\tau} \Phi(x) dx \]  

where \( \bar{p} = \frac{\lambda_0}{\lambda_0 + \lambda_1} \). That is, in each period the manager discloses his signal \( s_t \) if i) \( \theta_t = 1 \) and ii) \( s_t \geq E[s_t | z_t] + \sqrt{Var(s_t | z_t)} \ast \tau \).

Notice that the threshold \( \tau \) is independent of all parameters of the model except \( p_\infty \). This means that in the static case the frequency of disclosure is independent of most firm characteristics, including the precision of the manager’s signal \( \tau_m \), the precision of the consensus signal \( \tau_c \), and the volatility of the firm fundamentals \( \sigma_u \). We conclude that only in the presence of reputation concerns these properties of the information environment have the potential to affect disclosure behavior.
We next consider the structure of the manager’s payoffs, given disclosure.

**Proposition 3.** In any equilibrium, \( V_D(p, z, s) = \frac{P_D(z, s)}{1 - \rho^2} \).

The manager’s payoff given disclosure is linear in both the public information \( z \) and the manager’s private signal \( s \). On the surface, this property seems to ignore the option value to withhold information available to the manager; in effect, the manager’s payoff given disclosure is the same as if the manager had committed to full disclosure forever. The reason is that given disclosure, the option value of withholding information in the future is zero. Upon disclosure, information becomes symmetric: the manager’s and the market’s expectations about both \( \theta \) and the sequence of future prices are the same. Moreover, expectations about future prices are linear in the current price, given the law of iterated expectations.

The payoff given non-disclosure also increases in the signal \( s \), because a higher \( s \) has a positive expected reputation effect (the market is more likely to give the manager the benefit of the doubt given non disclosure). However, the payoff given non disclosure is non-linear in \( s \). This suggests that the existence of a threshold equilibrium is not guaranteed. Indeed when the information endowment is persistent and the manager signal is very precise, there is no threshold equilibrium in our game.

**Proposition 4.** Assume the information endowment is persistent, \( \lambda_0 < 1 - \lambda_1 \), then there is no threshold equilibrium when the manager is almost perfectly informed \( Var(v_t) = 0 \).
Proof. Consider the case without earnings and suppose there is an equilibrium threshold $k$. Then, we have that
\begin{align}
V_D^1(p, z, k) &= V_{ND}^1(p, z, k) \\
&= P_D^1(z, k) + E[\beta \lambda_0 V_0(\lambda_0) + (1 - \lambda_0)V_1(\lambda_0)|s = k] \\
&= P_{ND}^1(z) + E[\beta \lambda_0 V_0(p') + (1 - \lambda_0)V_1(p')|s = k]
\end{align}
where $p' \in (\lambda_1, \lambda_0)$. Now assume the manager observes $s = k + \varepsilon > k$ for $\varepsilon$ small enough, but chooses to withhold his signal. Then he obtains:
\begin{align}
\hat{V}_1(p, z, k + \varepsilon) &= P_{ND}^1(z) + \beta E[\lambda_0 V_0(\lambda_1) + (1 - \lambda_0)V_1(\lambda_1)|s = k + \varepsilon]
\end{align}
But because $V_\theta(p)$ increases in $p$ (PROVE) and $\lambda_1 < p' < \lambda_0$ it follows that
\begin{align}
\hat{V}_1(p, z, k + \varepsilon) > V_D^1(p, z, k) \approx V_D^1(p, z, k + \varepsilon). 
\end{align}
a contradiction. Now the question is whether this logic extends to the case with earnings but $\sigma_v \to \infty$.

The classic threshold structure characterizing the equilibrium in the static game is not assured in our dynamic setting. To understand this result consider the incentives of a manager whose signal lies slightly above the market’s conjectured threshold. If the manager follows the prescriptions of the equilibrium and discloses his signal he obtains a second order price benefit today (relative to the price of non-disclosure) but he gives up a large benefit in terms of continuation value. In effect, by deviating from the equilibrium and withholding his signal, the manager could lead investors to believe he was almost surely uniformed. This belief would boost the market perception that the manager will be uninformed in the next period, thus increasing the option value of concealment. Such deviations exist when the manager signal is very precise, since the reputation effect of the deviation is very strong in that case. Indeed, in the limit, as the amount of noise in the manager’s signal vanishes, withholding a signal slightly above the threshold would lead investors to believe the manager was uninformed with probability one. Fortunately, this lack of existence issue arises at implausibly low levels of noise in the manager’s signal.

We will thus ignore it as being empirically irrelevant. The following result characterizes linear equilibria.

**Proposition 5.** In any linear equilibrium, the probability of disclosure is independent of the consensus forecast, $c$, and the serial correlation of earnings, $\rho$. In such linear equilibria the manager uses the following disclosure strategy:
\begin{align}
y = E(s|z) + k(p).
\end{align}
where $k(p)$ is defined by $V_D^1(p, y, z) = V_{ND}^1(p, y, z)$. 

12
This result predicts that the probability of disclosure is independent of the evolution of the public information that precedes the manager’s disclosure choice. The reason is that the consensus forecast shifts, in exactly the same way, both the manager disclosure strategy and the market beliefs. This result would seem to contradict the clustering result in Acharya, Demarzo and Kremer (2011) stating that the probability of disclosure is higher in bad times, namely when market expectations are lower. In their paper, this result arises because the timing of disclosure is endogenous. The manager can disclose his signal, either before or after a public signal is released.

The fact that disclosure probabilities are not affected by the realization of the consensus \( z_t \), does not mean that disclosure choices are independent of the presence of public information, as in the static model. In our dynamic setting the variances do affect the probability of strategic withholding. Indeed, at the margin, the strength of reputation concerns depends upon the quality of the public information that precedes the manager disclosure choice.

6 Comparative Statics

Managers’ disclosure choices are affected by several aspects of the information environment. First, disclosure choices depend upon the extent to which the manager cares about the current price. A higher patience \( \beta \) is equivalent to an increase in the manager’s reputation concern. The manager tends to withhold information more often to affect the market perceptions about \( \theta \), thus foregoing a short-term price benefit in order to gain disclosure flexibility in the future.
The distribution of information endowment \( \theta \) is also a key determinant of disclosure behavior. The distribution of \( \theta \) can be characterized by two parameters: the long-run probability of being informed \( \bar{p} = \frac{\lambda_0}{\lambda_0 + \lambda_1} \) and the persistence of the information endowment \( r \equiv 1 - \frac{\lambda_0}{2\bar{p}} \).

The effect of changes in \( \bar{p} \) is intuitive: The higher the probability the manager is informed, the more skeptical is the market about the firm value when the manager does not disclose, which in turn prompts more disclosure; exactly as in the static case.

The effect of persistence is more subtle. First, notice that the absence of persistence (i.e., \( r = \frac{1}{2} \)) renders the model static; the manager uses the static threshold because in the his disclosure choice does not affect market perceptions about tomorrow information endowment. However, when the information endowment is persistence \( r > \frac{1}{2} \) the manager has incentives to distort disclosure choices, relative to the myopic case, to affect the market’s perception about his future information endowment. In particular, the manager is concerned about setting a high disclosure precedent that will subject him to a higher disclosure standard in the future. Such concerns makes the manager less likely to disclose. A higher persistence also affects the dispersion of the long-run distribution of beliefs. The latter effect may counteract the reputation effect sometimes resulting in more disclosure rather than less, especially when the level of persistence is relatively high.
Consider the effect of the variance of fundamentals $\sigma_u$. First, observe that if earnings were not observable, then just like in the static version of the model, the volatility of fundamentals would not affect the probability of disclosure: this variance would simply scale the manager’s payoffs not affecting the marginal trade-off. But the presence of earnings can play a disciplining role in manager’s disclosure behavior because it allows investors to refine their beliefs regarding whether the manager has concealed information in the past. The distribution of beliefs, conditional on non-disclosure, depends on $\sigma_u$. In effect, both the expected value of $p$ and its dispersion are affected by how precise is earnings as a signal of the manager’s own information. This leads to $\sigma_u$ having a non monotone effect on disclosure incentives.

### 6.1 Asset Pricing Implications

The asset pricing implications of disclosure are not well known. A notable exception is Acharya, Demarzo and Kremer (2011). Here we consider whether strategic withholding affects the volatility of returns, and whether the volatility of returns is affected by the history of disclosure. We also examine how the sensitivity of returns to earnings announcements (i.e., $ERC$) is linked to the presence of strategic withholding of information. Furthermore, we investigate the notion that strategic disclosure leads to a skewed distribution of returns at the earnings announcement date, as put forth by Kothari, Shu and Wysocki (2009). We
define returns $R_t$ in levels as

$$R_t \equiv P_t - P_{t-1}.$$  

Because we are interested in the effect of strategic withholding on the properties of asset prices, we keep the information structure constant but vary the manager's incentives, as captured by $\beta$. Since an increase in $\beta$ makes the manager more patient, this leads to a higher probability of withholding; we can thus use $\beta$ as an index of the probability of withholding.

We simulate returns for 10,000 years. Our simulations reveal the following patterns. First, increased disclosure is associated with a lower volatility of returns at the earnings announcement date (but a higher volatility of returns at the disclosure date). Intuitively, when managers are less likely to disclose a forecast, the earnings announcement becomes more informative. This substitution effect also explains why the correlation between returns and earnings (ERC) goes up as $\beta$ increases.

Second, the average return, computed at the earnings announcement date, is independent of the manager’s disclosure behavior. In fact the probability of positive returns at the earnings announcement date equals that of negative returns; which contradicts the notion that the strategic delay of negative news leads to an asymmetrically strong reaction to bad news at the earnings announcement date (see Kothari, Shu and Wysocki (2009)). Notice that the volatility of returns is time-varying. Figure 6 shows that the volatility of returns, at the disclosure date, conditional on disclosure is higher than that conditional on non-disclosure. Disclosure history matters for the volatility of returns, insofar as the history determines the likelihood of disclosure in the future.

7 Data

Our sample is from the I/B/E/S earnings announcement database for fiscal years ending between January 1st 2004 and December 31st 2014. Details of the sample selection are
reported in Table 1.

We start the sample in 2004 because of significant changes in the regulation in 2000 and 2002 that have significantly change incentives to disclose information and the collection of forecasts. Since August 2000, Regulation Fair Disclosure prohibits most private communications which, by shutting down this alternative communication channel, have significantly increased the frequency of public managerial forecasts. Since July 2002, the Sarbanes-Oxley Act has dramatically increased internal controls and management responsibilities and, from a data-collection perspective, require all conference calls to be in transcript form. We construct a sample of raw earnings per share (EPS). Earnings in I/B/E/S are reported as pro-forma (“street”) earnings calculated under the same accounting principles as analysts’ and management forecasts. The initial sample includes 59,819 firm-years from 10,621 unique firms. We require non-missing announcement and lagged announcement dates, in order to create a window for management forecasts; this shrinks the sample to 53,993 firm-years from 9,420 unique firms.

We obtain management forecasts from the I/B/E/S management forecast guidance (CIG) database. We merge to the I/B/E/S earnings database using I/B/E/S unique tickers and forecasted period end date, retaining only annual forecasts where the forecasted period end date can be matched to the I/B/E/S earnings announcement sample. This selection yields a sample of 52,273 forecasts from 2,814 unique firms. We require a forecast to be made after the prior earnings announcement date but made at least six months before the period end date, reducing the sample to 23,950 forecasts from 2,478 firms. We remove forecasts made after the fiscal year end, because they are less likely to be consistent with a model of incomplete information endowment. The majority of our forecasts are made bundled with the prior earnings announcement, and are made between 10 and 11 months before a fiscal year end. For periods with multiple forecasts, we use the earliest forecasts, which yields a final sample of forecasts of 10,950 forecasts from 2,478 unique firms. We calculate the raw forecast, as it was made, by multiplying the forecast in I/B/E/S adjusted for the number of shares, with the I/B/E/S adjustment factor. We recover the later using the ratio of raw earnings to adjusted earnings in the I/B/E/S earnings database.

For each firm-year with or without a forecast, we require a measure of market expectation about realized earnings. For any firm-year with a forecast, we use the I/B/E/S CIG analyst consensus; this number is provided adjusted for stock splits by I/B/E/S and we unadjust it, as for management forecasts, by multiplying by the adjustment factor. This consensus measure reflects the consensus before a management forecast is made and presents the benefit of being highly visible to market participants. For firm-years without a forecast, we calculate a consensus using the I/B/E/S analyst file, which reports all annual forecasts made by financial analysts. We calculate a consensus using the average of all analyst forecasts for the current fiscal year earnings made during the window between one and two prior earnings announcements. This implies that the consensus is always constructed from analyst forecasts made prior to the management forecast. When this information was unavailable, that is, for firms analysts who rarely update their forecasts, we used the EPS at the prior earnings announcement as the market expectation. Dropping all firm-years for which no market

---

4For a few cases in which this information was missing (for example, if earnings are zero, we used the last available adjustment factor or, if unavailable as well, the earliest adjustment factor in a following period.
expectation could be constructed, the sample has 51,563 firm-years from 8,890 unique firms, including 10,922 management forecasts from 2,459 unique firms. We obtain price and number of outstanding shares from CRSP, and accounting fundamentals from Compustat, merging using the WRDS link between the identifiers I/B/E/S ticker, compustat gvkey and CRSP permno, and then perform an additional merge using fiscal year end and CUSIP. After the merge, the sample shrinks to 43,604 firm-years from 6,913 unique firms, and 10,601 forecasts from 2,331 unique firms.

To ensure a complete time-series for each firm, we require at least 10 earnings announcements, which may include firms with a complete time-series of 11 firm-years, as well as some additional firms that may have changed their fiscal year end and, thus, may have 10 or 12 earnings announcements. The final sample has 23,578 firm-years, from 2,177 unique firms. There are 7,139 management forecasts from 1,106 unique firms.

Table 2 reports several descriptive statistics. The overall disclosure frequency in the sample is 30%. An average management forecast is about $2.35, and about half of the forecasts lie within $1.20 and $3.03. Management forecasts tend to be significantly greater than realized earnings which are, on average, $1.68 per share. This appears to be consistent with some self-selection in forecasting behavior. When correcting for the market expectation, we find, however, that forecast surprises are not positive, at about −4 cents. Forecast errors are small, at about 5 cents per share on average. A typical firm in our sample has a full time-series of 11 earnings announcements. Average total assets are $23 billion USD and average market capitalization is $6.3 billion USD. These distributions, however, are highly right-skewed. The median firm has assets equal to $1.8 billion USD and $1.136 billion USD in market cap. The median book-to-market is 0.51, although the average 1.66 is much greater primarily due to a few outliers. Leverage is on average 3.24 but has a median of 1.16.

### 8 Estimation

We fix the values of the parameters of the quantitative dynamic model in two main steps. First, we calibrate four of the parameters ($\beta, \rho, \sigma_u, \sigma_n$) externally, drawing on both comparable quantitative exercises from the literature on dynamic corporate finance as well as a moment-matching exercise for the parameters of some purely exogenous series. Second, conditional upon these calibrated parameters, we estimate each of the three remaining disclosure-related parameters ($\sigma_v, \lambda_0, \lambda_1$) through a straightforward application of the generalized method of moments (GMM). Throughout this section, we rely on an annual dataset of firm earnings, consensus analyst forecasts, and selectively disclosed manager forecasts from IBES. Our sample represents an unbalanced panel of 1,043 US public firms spanning 2004-2014 with 8,095 firm-fiscal year observations.

### Outside Calibration of Earnings and Consensus Parameters

We assume a model period of one year, corresponding to the frequency of our IBES data. We set the value of the manager subjective discount rate to $\beta = \frac{1}{1.04}$, approximately in the
### Table 1. Outside Calibration

<table>
<thead>
<tr>
<th>Parameter, Role</th>
<th>Value</th>
<th>Targeted Moment</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$, Earnings Persistence</td>
<td>0.85</td>
<td>Corr($e_t, e_{t-1}$)</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>$\sigma_u$, Earnings Volatility</td>
<td>0.45</td>
<td>St Dev(IHS($e_t - \mathbb{E}e_t$))</td>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td>$\sigma_n$, Analyst Precision</td>
<td>0.68</td>
<td>St Dev(IHS($c_t - e_t$))</td>
<td>0.59</td>
<td>0.59</td>
</tr>
</tbody>
</table>

Note: The first two columns report the name, role, and value of each of the three parameters which are calibrated externally through a moment-matching exercise. The final three columns report the data and model values of the three targeted moments. “Corr” refers to correlation, “St Dev” refers to standard deviation, and “IHS” refers to the inverse hyperbolic sine transformation $IHS(x) = \log(x + \sqrt{1 + x^2})$. The moments are computed from IBES data on manager forecasts and realized fiscal year earnings spanning 2004-2014 for a sample of 1,043 US public firms in an unbalanced panel with 8,095 firm-year observations. In the model, moments are computed from a simulation of the earnings $e_t$ and consensus $c_t$ processes for 10,000 years, discarding the first 500 years of data. The minimization of the distance between model and data moments was performed numerically via particle swarm optimization.

The values $\rho$, $\sigma_u$, and $\sigma_n$ govern the persistence and volatility of the fundamental earnings process as well as the signal precision for analyst forecasts. The realized earnings process $e_t$ and consensus forecasts $c_t$ are directly observable in our data for all years. Both series are exogenous to the model disclosure choice, so to economize on the number of computationally costly estimated parameters we fix the values of these parameters by matching three data moments directly.

The model assumes a homogenous underlying driving process with $e_t = \rho e_{t-1} + u_t$ and $c_t = e_t + n_t$, where $u_t \sim N(0, \sigma_u^2)$ and $n_t \sim N(0, \sigma_n^2)$. Before taking this to the data directly, it’s useful to note that the reported values within our dataset for earnings per share may vary across firms in levels and scale as the share structure and size of firms varies in the cross section. In order to compute transformations which are approximately unit free, we target the autocorrelation of earnings as well as the dispersion of the inverse hyperbolic sine of firm-demeaned earnings and analyst forecast errors.

Although the mapping between moments and parameters here is joint and not one-to-one, the autocorrelation of earnings is informative for the value of persistence $\rho$, and the dispersion of earnings and analyst forecast errors help to pin down the values of $\sigma_u$ and $\sigma_n$, respectively.

---

5 See, for comparison, the dynamic firm-level investment model calibrations studied in Hennessy and Whited (2007) or Cooper and Ejarque (2003).

6 The inverse hyperbolic sine $IHS(x)$ is a convenient function which is approximately equal to $\log(2x)$ for large values of $x$ and hence is approximately unit free. The $IHS$ function is also defined for both positive and negative values, convenient given the presence of losses and negative analyst forecast errors in the observed data. See MacKinnon and Magee (1990) for further information on this transformation.
Table 2. Targeted Moments for the GMM Estimation

<table>
<thead>
<tr>
<th>Moment</th>
<th>Data Value (Standard Error)</th>
<th>Model Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(d_t = 1)$</td>
<td>0.59 (0.0117)</td>
<td>0.63</td>
</tr>
<tr>
<td>$P(d_t = 1</td>
<td>d_{t-1} = 1)$</td>
<td>0.88 (0.0054)</td>
</tr>
<tr>
<td>$P(d_t = 1</td>
<td>d_{t-2} = 1)$</td>
<td>0.83 (0.0072)</td>
</tr>
<tr>
<td>$P(d_t = 1</td>
<td>d_{t-3} = 1)$</td>
<td>0.78 (0.0088)</td>
</tr>
<tr>
<td>St Dev($s_t - e_{t</td>
<td>d_{t}=1}$)</td>
<td>0.45 (0.0122)</td>
</tr>
</tbody>
</table>

Note: The table reports the data values of moments computed from IBES data on manager forecasts $s_t$ conditional upon disclosure $D_t$ and realized fiscal year earnings $e_t$ spanning 2004-2014 for a sample of 1,043 US public firms with 8,095 observations. $D_t$ refers to a disclosure event in time $t$, “St Dev” refers to standard deviation, and “IHS” refers to the inverse hyperbolic sine transformation $IHS(x) = \log(x + \sqrt{1 + x^2})$. Standard errors based on a block bootstrap procedure, resampling at the firm level. The third column of the table reports the values of the targeted moments in the estimated model.

Internal Estimation of Disclosure Parameters

Armed with values for each of the other parameters in the model, we now turn to the estimation of $\sigma_v$, $\lambda_0$, and $\lambda_1$. These three parameters relate directly to the unobserved series of manager signals and information endowments in the model, and we cannot observe either of these series directly in the dataset. Instead, we turn to a GMM procedure. We minimize the sum of squared deviations between the model and data values of a selected set of moments relating to the disclosure process.

To ensure identification, moment-based structural estimation strategies rely crucially on the selection of an appropriate set of targeted moments which are informative for the underlying parameters of interest. Within our GMM estimation exercise, just as in the case of our external calibration, the mapping between parameters and moments is joint and not one-to-one. However, in the model the disclosure rate and dynamics rely directly upon the parameters $\lambda_0$ and $\lambda_1$ of the manager information endowment. This dependence reflects a direct effect through the information endowment itself as well as an indirect effect through the dynamics of the market belief $p_t$ surrounding the manager’s information state at any point in time. To capture this relationship, we target the unconditional probability of disclosure as well as the probability of disclosure given prior disclosure in each of the past three years. Note that conditional upon disclosure, the dispersion of manager forecast errors depends directly upon the precision of the manager signal $\sigma_v$. Following the same strategy we implemented for analyst forecast errors above, we target the standard deviation of transformed manager forecast errors in the data. Table 2 reports the values and standard errors of the five targeted moments from our sample of IBES data.

With the targeted moments from the data in Table 2 in hand, and given values for the parameters $\sigma_v$, $\lambda_0$, and $\lambda_1$, we compute comparable model moments directly from the ergodic or stationary distribution of the model. We numerically minimize the GMM objective function, the sum of squared deviations between data and model moments, to obtain point estimates. The standard errors of the estimated parameters follow standard GMM formulas and rely upon a bootstrap procedure to calculate the empirical moment covariance matrix.\footnote{In particular, we minimize the GMM objective using particle swarm optimization, a numerical global}
Table 3. GMM Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter, Role</th>
<th>Estimate (Standard Error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda_0, \mathbb{P}(\theta_t = 0</td>
<td>\theta_{t-1} = 1) )</td>
</tr>
<tr>
<td>( \lambda_1, \mathbb{P}(\theta_t = 1</td>
<td>\theta_{t-1} = 0) )</td>
</tr>
<tr>
<td>( \sigma_v, \text{St Dev Manager Signal} )</td>
<td>0.45 (0.0178)</td>
</tr>
</tbody>
</table>

Note: Parameters estimated via GMM, targeting the five moments described in the text and implemented with a diagonal weighting matrix and numerical minimization. The standard errors are based on an empirical moment covariance matrix computed from a block bootstrap procedure, resampling at the firm level, together with a moment Jacobian computed from numerical forward differentiation averaging over step sizes of 0.25%, 0.5%, and 1%. “St Dev” refers to standard deviation.

Table 3 reports the parameter estimates. We estimate substantial persistence in manager information endowments, with only around 4% probability of future information loss given a manager signal today (\( \hat{\lambda}_0 \)) and less than a 20% chance of becoming informed for a currently uninformed manager (\( \hat{\lambda}_1 \)). We reject the iid hypothesis at 1% significance level. Indeed our point estimate of persistence \( r \) is 0.88. The high degree of persistence confirms the importance of forward looking considerations when it comes to explaining disclosure behavior.

The manager signal dispersion \( \hat{\sigma}_v \approx 0.45 \) implies more precise information for managers than analysts. As documented in the final column of Table 2 the estimated model reproduces the targeted moments well. We cannot expect an exact match between data and model moments given our nonlinear model and the overidentified nature of the estimation exercise with three parameters and five targeted moments. However, the estimated model closely reproduces the unconditional disclosure probability of around 60%, the high persistence of disclosure observed in the data, and the dispersion of manager forecast errors.

9 Application

9.1 Strategic withholding and information loss

In the last two decades, policy makers in the U.S. have encouraged firms to disclose forward looking information. Indeed, in 1995 U.S. regulators enacted the Private Securities Litigation Reform act (PSLR) which includes a safe harbor provision that protects managers from litigation arising from unattained projections of forward looking information.\(^8\) Regulation Fair Disclosure to prohibit managers from selectively disclosing information to some market participants while excluding others. Arguably, these regulations partially explain the rapid expansion of management forecasts observed in the last two decades.

\(^8\)Following a similar logic, in 2002 U.S. regulators enacted the Regulation Fair Disclosure to prohibit managers from selectively disclosing information to some market participants while excluding others.
tors’ preference for more disclosure emphasize the importance of quantifying the magnitude of firms’ strategic withholding of information.

### Table 4. Strategic Withholding

<table>
<thead>
<tr>
<th>Name</th>
<th>Data</th>
<th>Model</th>
<th>Model Full Disclosure</th>
<th>Model IID Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathbb{P}(d = 1) )</td>
<td>0.59</td>
<td>0.63</td>
<td>0.85</td>
<td>0.60</td>
</tr>
<tr>
<td>( \mathbb{P}(d = 1</td>
<td>d_{−1} = 1) )</td>
<td>0.88</td>
<td>0.81</td>
<td>0.96</td>
</tr>
<tr>
<td>( \mathbb{P}(d = 1</td>
<td>d_{−2} = 1) )</td>
<td>0.83</td>
<td>0.75</td>
<td>0.94</td>
</tr>
<tr>
<td>( \mathbb{P}(d = 1</td>
<td>d_{−3} = 1) )</td>
<td>0.78</td>
<td>0.72</td>
<td>0.92</td>
</tr>
<tr>
<td>St Dev(( IHS(s − e)</td>
<td>d = 1 ))</td>
<td>0.45</td>
<td>0.47</td>
<td>0.52</td>
</tr>
<tr>
<td>( \mathbb{P}(d = 0</td>
<td>\theta = 1) )</td>
<td>0.26</td>
<td>0.00</td>
<td>0.29</td>
</tr>
<tr>
<td>( \mathbb{P}(\theta = 1) )</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
</tr>
</tbody>
</table>

Note: The table above reports data and model moments. The data moments in the second column, computed only when observable, are drawn from IBES data on manager forecasts \( s_t \) conditional upon disclosure and realized fiscal year earnings \( e_t \) spanning 2004-2014 for a sample of 1,043 US public firms with 8,095 firm-fiscal year observations. The model moments, in the remaining column, are computed numerically directly from the theoretical stationary distribution. The third column reports moments from the estimated baseline model. The fourth column reports moments from an otherwise identical counterfactual model with no withholding allowed. The fifth column reports moments from an otherwise identical counterfactual model with no persistent in information endowments.

The estimates in Table 4 show that, on average, managers are informed 85% of the time, but conditional on being informed, they withhold information strategically 26% of the time (the unconditional probability of withholding is roughly \( .85 \times .26 = 22\% \)). The amount of withholding implied by a model with IID information endowment, but otherwise identical, is 29%. The greater probability of withholding implied by the IID model speaks to the importance of the dispersion effect discussed in Section 6. The empirically high level of persistence \( (r = .88) \) explains the relatively low frequency of strategic withholding.

We can also use our estimates to assess the information loss associated with the manager’s strategic withholding. To evaluate the amount of information loss caused by strategic withholding we compute the Mean Square Error MSE of market expectations and contrast it with the MSE arising in an otherwise identical model where managers do not withhold information strategically. We simulate a panel of 1,000 firms and 5,000 years. The results from the simulation are reported in Table 5. The root mean square error RMSE of the market prediction for earnings is about 6% higher in the baseline model than in the model with no strategic withholding. Strategic withholding thus leads to substantial information loss.
Table 5. Strategic Withholding and Information Loss

|                | # Firms | # Years | # Yrs. Discarded | RMSE($E(e'|D)$) St. Dev Earnings |
|----------------|---------|---------|------------------|---------------------------------|
| Baseline Model | 1,000   | 5000    | 50               | 0.3565742 0.8937673              |
| No Strategic Non-Disclosure | 0.3350457 0.8937673 |

Note: The results are obtained simulating a panel of 1000 firms for 5000 periods, based on the estimated annual parameters.

10 Conclusion

This paper develops a realistic model of voluntary disclosure in which managers have forward looking motives and firms’ disclosure choices interact with both analysts consensus forecasts and mandatory earnings announcements.

Building on Dye (1985) we assume that markets are uncertain about managers’ information endowment. We show that small levels of persistence in the information endowment can fuel reputation concerns capable of generating significant disclosure stickiness. This result accords with the notion, originally advanced by the survey of Graham, Harvey and Rajgopal (2005), that managers’ reluctance to disclose is driven by the fear of setting a disclosure precedent that might expose the stock price to negative market reactions during non-disclosure periods.

The persistence of information endowments has a subtle effect on disclosure policy. On one hand, persistence boosts managers’ reputation concerns: by manipulating disclosure policies the managers may attempt to influence market perceptions regarding the manager’s information endowment. Indeed, by withholding information —thereby pretending to be uninformed— a manager can increase the option value of withholding information in the future. The presence of this reputation effect makes managers more reluctant to disclose information. Persistence also has another effect, running in the opposite direction. Higher persistence leads, in the long-run, to a greater dispersion in the distribution of market beliefs about the manager’s information endowment. On average, this belief dispersion effect causes stronger unraveling pressure thus inducing extra disclosure. In some cases, the belief dispersion effect dominates the reputation effect. A high persistence is consistent with very expansive disclosure policies.

In contrast with static models, the probability of strategic withholding depends on the relative quality of managers’ private information. The reason is that in the dynamic setting the periodic release of earnings announcements discipline the manager’s disclosure behavior weakening the manager’s reputation concerns.

We estimate the model using an unbalanced panel of 1,043 US public firms spanning 2004-2014 with 8,095 firm-fiscal year observations. Our estimates suggest that, conditional on being informed, firms withhold information 26% of the time, thereby causing an information loss that amounts to 6%, measured in root mean-square error, relative to a setting where managers do not withhold information strategically.
A Appendix

A.1 Cash flows versus earnings

Thus far, we have assumed that manager preferences depend on the market’s perception of the firm’s discounted earnings. In rigor, stock prices are the present value of future cash flows, which generally differ from earnings. Next we provide a simple method that allows us to infer cash-flows based on the firm’s earnings as an intermediate step toward computing stock stock prices and returns.

Assume as in Hennessy and Whited (2007) that output for a firm is given by
\[ y = k^\alpha, \]
where \( \alpha \in (0, 1) \). Here, \( z \) is idiosyncratic productivity. But assume that investment is flexibly adjusted so that a firm can statically choose
\[ k = i + (1 - \delta)k_{-1}. \]
A firm therefore solves
\[ \max_k y - p_I i \]
\[ k = i + (1 - \delta)k_{-1}. \]

You have that
\[ k = z^{\frac{1}{1-\alpha}} \left( \frac{\alpha}{p_I} \right)^{\frac{1}{1-\alpha}} \]
\[ y = z^{\frac{1}{1-\alpha}} \left( \frac{\alpha}{p_I} \right)^{\frac{\alpha}{1-\alpha}} = \text{Revenue} \]
\[ y - p_I i = (y - p_I k) + p_I (k - i) = (1 - \alpha)y + p_I (1 - \delta)k_{-1} = \text{Cash Flow} \]
\[ y - p_I \delta k = (y - p_I k) + p_I (k - \delta k) = (1 - \alpha)y + p_I (1 - \delta)k = \text{Earnings} \]

So unless we are in steady state, earnings differ from cash flow, because firms that are growing have higher earnings than cash flow. However, we have that
\[ \text{Earnings} - \text{Cash Flow} = p_I (1 - \delta)(k - k_{-1}) = p_I (1 - \delta) \left( \frac{\alpha}{p_I} \right)^{\frac{1}{1-\alpha}} (z^{\frac{1}{1-\alpha}} - z_{-1}^{\frac{1}{1-\alpha}}) \]

If we assume that \( \hat{z} \equiv z^{\frac{1}{1-\alpha}} \) follows \( \hat{z} = \rho \hat{z}_{-1} + \varepsilon \), then we get that earnings inherit the AR(1) property because they are a constant multiple of \( \hat{z} \):
\[ \text{Earnings} = \hat{z} \left( \frac{\alpha}{p_I} \right)^{\frac{\alpha}{1-\alpha}} \left[ (1 - \alpha) + p_I (1 - \delta) \left( \frac{\alpha}{p_I} \right) \right]. \]

We also get that the difference between earnings and cash flow are given by
\[ \text{Earnings} - \text{Cash Flow} = p_I (1 - \delta) \left( \frac{\alpha}{p_I} \right)^{\frac{1}{1-\alpha}} [\hat{z} - \hat{z}_{-1}] \]

We can use this method to simulate stock prices and stock prices defined as
\[ P_t = \frac{E_t(Cash Flow_t | D_t)}{1 - \beta \rho}. \]
We calibrate the model assuming $\alpha = 0.5$, $\delta = 0.1$, $p_I = 1$, roughly standard values in the economics literature. Based on that, we compute stock prices based on the estimated distribution of cash flows using a discount factor of $\frac{1}{1.04}$. In the baseline model the variance of returns is 10% higher, at the earnings announcement date, when non-disclosure occurs than when disclosure occurs. The histograms of the two return distributions in the baseline model are shown in Figure A.1. The standard deviation of returns measured at the earnings announcement date, conditional on non-disclosure (resp. disclosure) is 1.96 (resp. 1.86).

### A.2 Computation

The computational strategy is straightforward policy iteration. Basically, the steps are as follows:

1. Discretize the state space.

2. On the $s$-th iteration of the solution algorithm, guess a disclosure policy $d^{(s)}(p, e_{-1}, c, s)$.

   (a) Assume that market beliefs as well as manager actions are governed by $d^{(s)}$, and iterate forward on the system of Bellman equations above until the implied $V_1^{(s)}$, $V_0^{(s)}$ converge to some tolerance.

   (b) Compute the stationary distribution $\mu_1^{(s)}(p, e_{-1}, c, s)$ and $\mu_0^{(s)}(p, e_{-1}, c)$ of the model given $d^{(s)}$ as well as the exogenous distributions in the model. This involves repeatedly pushing forward weight on a histogram given the policies and exogenous transitions until the distributions stabilizes to within some tolerance.

   (c) Compute a new policy $d^{(s+1)}(p, e_{-1}, c, s)$, simply given by

   $$\arg \max_d \left( d V_1^{D(s)} + (1 - d) V_1^{ND(s)} \right).$$

   (d) Then, compute an error measure given by the mean absolute difference between $d^{(s+1)}$ and $d^{(s)}$, weighted by the ergodic distributions $\mu_1^{(s)}$ and $\mu_0^{(s)}$. This error is exactly equal to the probability of disclosure policy deviation given assumed market beliefs. When this error is sufficiently small, you’ve computed an equilibrium.

3. Once we have solved the model, we can compute moments as desired for input into the structural estimation routine.

### A.3 Data
Figure 6. Return heteroskedasticity.

We simulate 1000 firms for 5000 periods, based on the estimated annual parameters. We calibrate the baseline model assuming $\alpha = 0.5$, $\delta = 0.1$, $p_I = 1$, which are standard values in the economics literature. Based on that, we compute stock prices based on the estimated distribution of cash flows using a discount factor of $\frac{1}{1.04}$.
Table 6. Sample Selection.

This table summarizes the sample selection criteria. Annual earnings announcements (EA) and management forecasts (MF) are obtained from I/B/E/S. Firms in our sample must be present in the CRSP and Compustat database using the I/B/E/S ticker, gvkey and permno matching tables from Wharton Research Data Services. Market expectation is calculated as the I/B/E/S analyst consensus in CIG (for observations with a MF), or as a researcher-created consensus from the I/B/E/S analyst forecast file (for observations without a MF); if any of these is not available, we use the inflation-corrected lag EPS.

<table>
<thead>
<tr>
<th></th>
<th>Earnings announcements (EA)</th>
<th>Management forecasts (MF)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nb. of EA</td>
<td>Unique firms</td>
</tr>
<tr>
<td>I/B/E/S EA sample 2004-2014</td>
<td>58,819</td>
<td>10,621</td>
</tr>
<tr>
<td>non-missing current or prior announcement date</td>
<td>53,993</td>
<td>9,420</td>
</tr>
<tr>
<td>I/B/E/S CIG sample 2004-2014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>matched to I/B/E/S EA</td>
<td>53,993</td>
<td>9,420</td>
</tr>
<tr>
<td>after prior EA date but prior period end</td>
<td>53,993</td>
<td>9,420</td>
</tr>
<tr>
<td>minimum 6 month before period end</td>
<td>53,993</td>
<td>9,420</td>
</tr>
<tr>
<td>retain only earliest MF</td>
<td>53,993</td>
<td>9,420</td>
</tr>
<tr>
<td>must have market expectation</td>
<td>51,563</td>
<td>8,890</td>
</tr>
<tr>
<td>merged to CRSP</td>
<td>43,798</td>
<td>6,956</td>
</tr>
<tr>
<td>merged to Compustat, non-missing assets</td>
<td>43,604</td>
<td>6,913</td>
</tr>
<tr>
<td>at least 10 EA</td>
<td>23,578</td>
<td>2,177</td>
</tr>
<tr>
<td><strong>Full sample</strong></td>
<td><strong>23,578</strong></td>
<td><strong>2,177</strong></td>
</tr>
</tbody>
</table>
Table 7. Descriptive Statistics.

This table reports descriptive statistics. Forecast frequency is computed as the frequency of a management forecast over all firm-years. Management forecast (MF) is the unadjusted annual management forecast. Realized I/B/E/S is the (pro-forma) earnings reported by I/B/E/S. Forecast surprise is the difference between the MF and the market expectation (from I/B/E/S or estimated as lagged EPS when a consensus is unavailable). Forecast error is the difference between MF and the realized I/B/E/S earnings. Firm characteristics are obtained from Compustat and measured at the lagged earnings announcement date. Market capitalization is obtained from CRSP and measured as the closing price (variable prc) multiplied by the number of shares outstanding (variable shrout) one day before the lagged earnings announcement.

<table>
<thead>
<tr>
<th>Forecast characteristics</th>
<th>mean</th>
<th>sd</th>
<th>min</th>
<th>1%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>99%</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forecast frequency</td>
<td>30%</td>
<td>46%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Management forecast</td>
<td>2.35</td>
<td>2.35</td>
<td>-3.85</td>
<td>-0.33</td>
<td>1.20</td>
<td>1.98</td>
<td>3.03</td>
<td>8.00</td>
<td>103.05</td>
</tr>
<tr>
<td>Realized I/B/E/S earnings</td>
<td>1.68</td>
<td>3.76</td>
<td>-196.12</td>
<td>-3.22</td>
<td>0.44</td>
<td>1.32</td>
<td>2.50</td>
<td>9.78</td>
<td>89.61</td>
</tr>
<tr>
<td>Forecast surprise</td>
<td>-0.04</td>
<td>0.24</td>
<td>-3.43</td>
<td>-0.77</td>
<td>-0.08</td>
<td>-0.02</td>
<td>0.03</td>
<td>0.17</td>
<td>5.40</td>
</tr>
<tr>
<td>Forecast error</td>
<td>0.05</td>
<td>0.93</td>
<td>-13.23</td>
<td>-1.50</td>
<td>-0.19</td>
<td>-0.04</td>
<td>0.14</td>
<td>3.10</td>
<td>14.28</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Firm characteristics</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of years</td>
<td>11</td>
<td>0.36</td>
<td>10</td>
<td>10</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Total assets</td>
<td>23,108</td>
<td>146,390</td>
<td>2</td>
<td>25</td>
<td>491</td>
<td>1,812</td>
<td>6,675</td>
<td>401,433</td>
<td>3,065,553</td>
</tr>
<tr>
<td>Market Capitalization</td>
<td>6,330</td>
<td>21,846</td>
<td>1</td>
<td>29</td>
<td>249</td>
<td>1,136</td>
<td>3,863</td>
<td>101,449</td>
<td>568,178</td>
</tr>
<tr>
<td>Book to Market</td>
<td>1.66</td>
<td>14.14</td>
<td>-74.33</td>
<td>-0.19</td>
<td>0.30</td>
<td>0.51</td>
<td>0.83</td>
<td>24.50</td>
<td>785.16</td>
</tr>
<tr>
<td>Leverage</td>
<td>3.24</td>
<td>62.85</td>
<td>-2,556.17</td>
<td>-6.58</td>
<td>0.53</td>
<td>1.16</td>
<td>2.44</td>
<td>23.25</td>
<td>4,800.72</td>
</tr>
</tbody>
</table>
B Proofs

Proof of Proposition 2. This result comes from noting that when $\lambda_0 = 1 - \lambda_1$ we get

$$V^D(p, z, s) - V^{ND}(z, s) = P^D(p, z, s) - P^{ND}(p, z).$$

which means that the manager maximizes his myopic price gain, when choosing whether or not to disclose. \hfill \square

Proof of Proposition 3. By assumption the value function can be expressed as

$$V^D(p, s, z) = E[\sum_{t=0}^{\infty} \beta^t P_t | s, z] = \sum_{t} \beta^t E[P_t | z, s] = \frac{P^D(z, s)}{1 - \rho \beta},$$

where the last equality follows by the law of iterated expectations and AR(1) property of earnings. \hfill \square

Proof of Proposition 5. Recall that

$$e' = \rho e + u; c = e' + n; s = e' + v$$

Let’s define

$$\tilde{e}' \equiv e' - E[s | z]$$

$$\tilde{c} \equiv c - E[s | z]$$

$$\tilde{s} \equiv s - E[s | z]$$

Notice that by construction $\tilde{c}$ and $\tilde{s}$ are independent of $z$. However $\tilde{c}$ is correlated with $z$. Now, the price given disclosure can be written as

$$P(s, z) = \frac{E(e'| z, s)}{1 - \rho \beta} = \frac{E(e'| z, s)}{1 - \rho \beta} = \frac{E(s | z) + \phi \tilde{s}}{1 - \rho \beta}$$

for some constant $\phi$. Now, assume the disclosure threshold is given by

$$y = E(s | z) + k$$

then the price of no disclosure can be written as

$$P^{ND}(z) = \frac{E(e'| z, s < y)}{1 - \rho \beta} = \frac{E(s | z) + \phi \kappa(k)}{1 - \rho \beta}$$

for some function $\kappa(\cdot)$. Finally, note that
\[
\Pr(s < y|z, e') = \Pr(s < k|z, e') \\
= \Phi\left(\frac{y - E(s|z, e')}{\sigma(s|z, e')}\right) \\
= \Phi\left(\frac{y - [E(s|z) + \hat{\varphi}(e' - E(s|z))]}{\sigma(s|z, e')}\right) \text{ for some constant } \hat{\varphi} \\
= \Phi\left(\frac{k - \hat{\varphi}e'}{\sigma(s|z, e')}\right)
\]

This implies that \(p'\) is independent of \(z\), it only depends on \(\tilde{s}\). Also \(E[p'|z, s]\) only depends on \(\tilde{s}\). We conjecture and verify that the value functions can then be written as

\[V_0(p, z) = \alpha E[s|z] + \nabla_0(p),\]

and

\[V_1(p, s, z) = \alpha E[s|z] + \nabla_1(p, \tilde{s})\]

for some constant \(\alpha\) and functions \(\nabla_0, \nabla_1\) that depend only on \(\tilde{s}\). Based on the above conjecture we can write

\[V_1^D(p, z, s) = P^D(p, z, s) + \beta \alpha \rho E(s|z) + \beta E\left[(1 - \lambda_0) V_1(p', \tilde{s}') + \lambda_0 V_0(p')|\tilde{c}, \tilde{s}\right] \]

where we used

\[E[E(s'|z')]|c, s] = \rho E(s|z).\]

By the same token, we can write

\[V_1^{ND}(p, z, s) = P^{ND}(p, z) + \beta \alpha \rho E(s|z) + \beta E\left[(1 - \lambda_0) V_1(p', \tilde{s}') + \lambda_0 V_0(p')|\tilde{c}, \tilde{s}\right] \]

Now, by definition

\[V_1(p, z, s) = \max_d \left[dV_1^D(p, z, s) + (1 - d)V_1^{ND}(p, z, s)\right] \]

\[= \frac{E(s|z)}{1 - \rho \beta} + \beta \alpha \rho E(s|z) + \Lambda(\tilde{s})\]

(13)

where \(\Lambda(\cdot)\) is a function that does not depend on \(z\). Equating the coefficients of the original conjecture \(V_1(p, z, s) = \alpha E(s|z) + \nabla_1(p, \tilde{s})\) and those of Equation (9) yields

\[\frac{1}{1 - \rho \beta} + \beta \alpha \rho = \alpha,\]

hence

\[\alpha = \frac{1}{(1 - \rho \beta)^2}.\]

So we can express the value function as the sum of a linear function of \(E(s|z)\) and another term that is independent of \(z\).

\[V_1(p, s, z) = \frac{E(s|z)}{(1 - \rho \beta)^2} + \nabla_1(p, \tilde{s})\]

\[\square\]
References


Beyer, A. (2012). Conservatism and aggregation: The effect on cost of equity capital and the efficiency of debt contracts, rock Center for Corporate Governance at Stanford University Working Paper No. 120.


