Disclosure Dynamics and Investor Learning*

Frank Zhou†

University of Chicago, Booth School of Business

January 25, 2016

Abstract

This paper examines the effects of investor learning on the dynamics of managers’ voluntary disclosure decisions. Using a multi-period voluntary disclosure model and structural estimation, I find that investor learning introduces persistence in disclosure incentives because of the correlation of investors’ beliefs over time. Disclosure is more likely when investors’ beliefs about underlying firm profitability are more pessimistic or investors are more uncertain about their beliefs. This paper enriches the understanding of the dynamics of managers’ disclosure incentives and the role of investor learning in driving the dynamics. In doing so, it also provides an economic explanation for the observed “stickiness” of managers’ disclosure decisions.


†Mailing address: University of Chicago Booth School of Business, 5807 South Woodlawn Avenue, Chicago, IL 60637, United States. E-mail address: frankzhou8674@chicagobooth.edu. Telephone number: +1 (773) 870-2301.
1 Introduction

Managers disclose information facing a dynamic world. Empirical studies document that managers’ voluntary disclosure decisions are “sticky” (e.g., Tang, 2012; Billings et al., 2014), which is consistent with the nature of disclosures being multi-period. However, evidence on the dynamics of disclosures remains scant. In particular, what drives the dynamics of voluntary disclosures? Is the economic force underlying the mechanism empirically important? I propose one mechanism that could help explain the dynamics of disclosures—investor learning about underlying firm profitability, and empirically quantify its effects on the dynamics.

Investors learn by updating their beliefs about underlying firm profitability after seeing managers’ disclosures and firm financial reports (Pástor and Veronesi, 2003; Rogers et al., 2009; Patton and Verardo, 2012). First, many studies examine managers’ disclosure incentives from the supply side, that is, firms or managers. Disclosure incentives induced by the demand side, that is, changes in investors’ beliefs and uncertainty, have received less attention. A key feature of the demand side is that investors, who don’t know firm profitability, learn in a “sticky” way; that is, investors’ beliefs about profitability depend on their prior beliefs. I show analytically and empirically that managers’ disclosure decisions depend on investors’ beliefs and uncertainty. This result, combined with the “stickiness” of investor learning, leads to a positive and economically significant intertemporal correlation in the likelihood of disclosure, that is, “sticky” disclosure incentives. In traditional single-period models, the likelihood of disclosure depends only on firm or manager characteristics. So changes in these characteristics explain changes in the likelihood of disclosure. In contrast, investor learning introduces intertemporal variation in the likelihood of disclosure through changes in investors’ beliefs and uncertainty over time, holding constant firm characteristics.

Second, understanding the effects of investor learning on disclosure decisions is by itself important. Prior studies show that investors learn from disclosures, which explains asset prices and investor uncertainty (e.g., Pástor and Veronesi, 2003; Rogers et al., 2009). A natural question is how the way investors learn in turn affects disclosure decisions.¹ Using structural estimation, I find the effects of investor learning on the likelihood of management earnings forecasts are comparable in economic magnitude to factors such as managers’ forecast accuracy, firm profitability, and earnings

¹Other studies analyze various types of learning by investors (e.g., Kim and Verrecchia, 1991a,b; Xia, 2001; Pástor and Stambaugh, 2012) and focus on how investors interpret disclosures.
persistence (e.g., Miller, 2002; Hutton and Stocken, 2009; Chen et al., 2011).

The first part of the paper incorporates investor learning into a multi-period version of the disclosure model by Verrecchia (1983). The purpose of the model is to specify the mechanism through which investor learning affects disclosure decisions. The model describes a firm that operates for an infinite number of periods. Each period, a firm manager decides whether to incur an exogenous fixed cost to truthfully disclose her private noisy information about unrealized forthcoming earnings. The manager’s disclosure decision depends on its effect on stock price, modeled as the sum of discounted investor expectations about future earnings net of disclosure costs.² At the end of the period, actual earnings are released in a financial report. Central to the model is the assumption that investors update their beliefs about underlying firm profitability after observing (non-)disclosures and then after the release of the firm’s financial reports.

The model features a threshold equilibrium. The manager discloses if and only if her private signal exceeds a threshold. The model shows that investor learning affects disclosure in two ways. First, more pessimistic investor prior beliefs are associated with a higher likelihood of disclosure. Intuitively, pessimistic investor beliefs give the manager more room to convey good news to improve market perception, so the manager is more likely to disclose. The implication is that more profitable firms will not necessarily disclose more; a firm with low profitability could disclose more than a firm with marginally higher profitability if investors’ pessimism is sufficient to justify more disclosure. Bergman and Roychowdhury (2008), who use market sentiment to capture investors’ beliefs, document a similar effect.

Second, greater investor uncertainty about firm profitability increases disclosure. Greater investor uncertainty is associated with larger revision of investors’ beliefs given the same disclosure. In addition, uncertain investors rely more on the manager’s disclosure when updating their beliefs. Both effects increase the importance of the manager’s disclosure by increasing the sensitivity of stock price to the disclosure, which in turn increases the incentive to disclose.

Regarding the dynamics of disclosure incentives, the model shows investor learning induces persistence in the likelihood of disclosure. Persistence results from investors’ belief stickiness. For example, if realized earnings fall short of investors’ expectations, the earnings shortfall will

²Using stock price as the objective function is consistent with many disclosure models (e.g., Verrecchia, 1983; Dye, 1985) and does not require explicit management incentives to align expectations or reduce information asymmetry (e.g., Ajinkya and Gift, 1984; Cotter et al., 2006; Feng and Koch, 2010).
affect investors’ future beliefs and, by extension, managers’ disclosure decisions in future years. In
other words, because managers’ disclosure decisions depend on investors’ beliefs, the stickiness of
investors’ beliefs leads to positive intertemporal correlation in the likelihood of disclosure.

In the second part of the paper, I empirically quantify the effects of investor learning on the
dynamics of management earnings forecast decisions. I focus on two aspects: average future disclo-
sure frequency and the persistence of disclosure incentives. These two aspects are not necessarily
related. Firms could have high disclosure frequency, but disclosure decisions could be independent
over time and vice versa. I measure the average future disclosure frequency as the frequency of
forecast issuance over the following 12 years.³ I measure persistence as the slope coefficients from
regressing future disclosure probabilities on current disclosure probabilities, holding constant firm
characteristics.⁴ I focus on disclosure probabilities rather than disclosure decisions, because the
model does not predict how disclosure decisions are correlated over time.⁵

I use structural estimation for the empirical analysis. Structural estimation avoids measuring
investors’ beliefs using firm fundamentals, which also affect disclosure decisions. Moreover, it
respects the endogenous nature of disclosure decisions, that is, the threshold equilibrium. Using
actual earnings, management earnings forecasts, and stock prices, I recover the cross-sectional
distribution of investors’ prior beliefs, investor uncertainty, actual profitability, earnings persistence,
earnings volatility, firm risk, manager’s information precision, and the fixed cost of disclosure (see
section 2). Recovering the distribution of parameters allows me to simulate firms that resemble those
in the sample but whose economic characteristics are held constant. In light of these other model
parameters that affect disclosure decisions, whether the effects of investor learning are economically
important is an empirical question.

The results suggest that the effects of investor learning on managers’ average future disclosure
frequency are significant but heterogeneous across firms. For around 40% of firms, marginal changes
in investors’ beliefs or uncertainty do not change disclosure frequencies. Examining these firms

³ I obtain frequency of forecast issuance using the simulation described in section 6, which allows me to choose any
arbitrary number of periods while fixing firm characteristics. I choose 12 years, representing the largest number of
years of the sample. The results do not qualitatively depend on this choice.

⁴ I generate disclosure probabilities from the model. Measuring persistence requires variation in disclosure prob-
abilities. I shock firm performance in the initial period, which generates changes in investors’ beliefs, leading to
changes in the likelihood of disclosure. I compute disclosure probabilities for the next five years.

⁵ Allowing investors to learn disclosure costs generates dependence in actual disclosure decisions. Given the focus
on investor learning about profitability, I do not pursue this alternative but leave it as one possible future extension.
Similarly, I acknowledge but do not pursue other multi-period incentives.
reveals they always disclose, which is consistent with 57% of firms in the sample issuing annual earnings forecasts every year.

Investor learning has economically important effects on the average future disclosure frequency for the remaining 60% of firms. A one standard deviation change in investor uncertainty (investors’ prior beliefs) shifts the future forecast frequency by about 0.32 (0.23) standard deviations. For 5% of firms, the effects can be larger than 0.77 standard deviations for investors’ prior beliefs and 1.55 standard deviations for investor uncertainty. The effects are sizable compared with permanent shifts in other parameters. For example, the effect of firm profitability is about 0.44 standard deviations, and the effect of earnings persistence is about 0.55 standard deviations.

Next, the results show that investor learning generates positive intertemporal correlation in the likelihood of disclosure. A 10% increase in disclosure probability induced by investor learning on average increases the disclosure probability for the next year by around 10.3%. The effect declines monotonically to about a 7.5% increase in disclosure probability in five years. The decline occurs because shocks to current beliefs have diminishing effects on disclosure decisions further into the future.

The results highlight the benefits of allowing for cross-sectional variation in parameters. The cross-sectional average of the marginal effects of investor learning is close to zero. Without accounting for the cross-sectional heterogeneity, one might incorrectly infer investor learning has minimal effects. Moreover, biases from individual firm estimates can cancel out over the cross section, producing cleaner estimates of the average effects of investor learning across firms.

This paper makes several contributions. Existing research on the dynamics of managers’ voluntary disclosure decisions remains scant. Most studies examine managers’ disclosure incentives from the supply side, that is, managers or firms (e.g., Bhojraj et al., 2011). I explain the dynamics of managers’ voluntary disclosure incentives from the demand side, that is, the stickiness in investor learning. The results enrich our understanding of how information consumers’ behavior feeds back into information supply incentives and affects the dynamics of information supply.

This paper also adds to the literature that examines how market participants affect managers’ disclosure incentives (e.g., Bushee and Noe, 2000; Cotter et al., 2006; Billings et al., 2014). I

---

6I compute the standard deviation of investor uncertainty from its cross-sectional distribution recovered from the structural estimation. I compute the standard deviation of disclosure frequency using the actual disclosure frequency in the data.
find investor optimism reduces disclosure, whereas prior studies show analyst optimism increases disclosure (e.g., Cotter et al., 2006). The result suggests investor optimism differs from analyst forecasts in their effects on managers’ disclosure decisions (e.g., Bergman and Roychowdhury, 2008). It echoes the need to differentiate types of information users when examining managers’ disclosure incentives.

Unlike many prior studies, the empirical analysis jointly accounts for the theoretical relations among investor learning, firm risk, earnings process, management forecast process, and asset pricing in a multi-period setting. As I impose the model structure on the data, I explicitly account for managers’ endogenous incentives to withhold bad news from the market when making disclosure decisions (e.g. Chen et al., 2011; Houston et al., 2010). This framework allows me to examine the effects of any model component while holding other components constant, and compare their relative empirical magnitude. The empirical results are less subject to omitted cross-sectional heterogeneity, because I allow the model parameters to differ across firms, alleviating biases from firm-level estimates.

The paper has several limitations. First, it focuses entirely on investors, and all dynamics come from investors’ responses to firm financial reporting and managers’ voluntary disclosures. My model does not address managers’ reputational concerns (e.g., Einhorn and Ziv, 2008; Beyer and Dye, 2012) or the timing of disclosure (Acharya et al., 2011; Guttman et al., 2014). Second, my model does not include incentives to bias or obfuscate disclosures (Gao, 2013; Fischer and Verrecchia, 2000), litigation risk (Francis et al., 1994; Skinner, 1994; Marinovic and Varas, 2014), preference for aligning investors’ expectations (Ajinkya and Gift, 1984), capital structure (Frankel et al., 1995), corporate governance (Ajinkya et al., 2005), or signaling management types (e.g., Trueman, 1986). Although modeling complexity is one concern, the purpose of the paper is not to discriminate among these incentives but rather to provide one possible channel that explains the dynamic nature of voluntary disclosures. To the extent that learning profitability is still an important aspect of disclosure, the paper can be considered an attempt to isolate this channel.

The paper proceeds as follows. Section 2 presents the model. Section 3 describes the data. Section 4 presents descriptive analysis followed by structural estimation and discussion of results in section 5 and 6. I conclude in section 7.
2 Model

2.1 Investor learning about firm profitability

In principle, learning can be broadly defined as investors updating their beliefs about any metrics of interest upon observing new information. Gao and Liang (2013) show firm managers can learn from stock price in making their investment decisions. The definition of learning in this paper is more narrowly focused and follows from seminal work on (dynamic) Bayesian learning (Erdem and Keane, 1996). In these studies, agents are uninformed about some unknown parameters, for example, product quality (Erdem and Keane, 1996), firm profitability (Pástor and Veronesi, 2003), analysts’ forecasting ability (Chen et al., 2005), or managers’ forecasting ability (Hutton and Stocken, 2009). They update their beliefs after observing noisy signals that contain information about the pertinent parameters. The dynamics of agents’ beliefs affect their behavior over time.

I assume investors do not know underlying firm profitability, modeled as the mean of the earnings distribution. They update their beliefs about profitability upon observing managers’ disclosure decisions and firm financial reports. Focusing on profitability is consistent with prior research on voluntary disclosure that uses stock price as managers’ objective function, where stock price is a function of profitability. It is also consistent with prior research on investor learning (Pástor and Veronesi, 2003; Rogers et al., 2009), and with industry practice, for example, analyst reports (to forecast next period earnings, analysts need to estimate the firm’s profitability).

Uncertainty about the distribution of earnings relates to the literature on parameter uncertainty, which also assumes that agents do not know the distribution (of stock returns). Parameter uncertainty is usually linked to estimation risk. Rather than focusing on portfolio choices under estimation risk, the learning literature typically focuses on how agents update their beliefs after new information arrival and how learning alters the dynamics of agents’ behavior.

2.2 Model set up

2.2.1 Timeline

The model extends the voluntary disclosure model by Verrecchia (1983) to multiple periods and incorporates investor learning as an additional element. Specifically, a firm \( i \) operates for an infinite

\footnote{See Brown (1979), Barry and Brown (1985), Coles and Loewenstein (1988), Coles et al. (1995), and Xia (2001).}
number of periods. Three events occur each period.

1. Information endowment.

At the beginning of each period, the firm’s manager privately observes a signal $s_{it}$. Investors do not observe the signal but know the manager has observed it. The signal is the realization of the end-of-period earnings $y_{it}$ plus independent noise, $s_{it} \equiv y_{it} + \eta_{it}$, where the variance of $\eta_{it}$ captures the manager’s signal quality. I assume $\eta_{it} \sim \mathcal{N}(0, \frac{1}{\tau_i \kappa_i})$, where $\frac{1}{\tau_i}$ represents the earnings volatility of firm $i$, and $\kappa_i$ captures the manager’s information precision (higher $\kappa_i$ implies more precise information).

2. Voluntary disclosure decision.

After observing the private signal $s_{it}$ but before the end of the period, the manager decides whether to disclose $s_{it}$ to investors. Disclosure is truthful and incurs an exogenous cost $c_i$, following Verrecchia (1983). The manager’s utility depends on stock price after the disclosure decision.

3. Mandatory disclosure.

At the end of the period, a financial report containing $y_{it}$ is released to the public and the next period begins.

The existence of a disclosure cost prevents full disclosure (Grossman, 1981; Milgrom, 1981). Verrecchia (1983) interprets it as a proprietary cost. The cost captures factors that make disclosures expensive. Such factors might also include the cost of maintaining an investor relation function, preparing press releases, and the loss of management time. I assume the disclosure cost is fixed and known by both investors and the manager. The choice of a reduced-form disclosure cost is motivated by its theoretical and, more importantly, its empirical simplicity.

Prior literature identifies other factors that prevent full disclosure, such as entry concerns (Gigler, 1994) and probabilistic information endowment (Dye, 1985; Jung and Kwon, 1988).

In multi-period models, incentives such as building a reputation for being forthcoming (Beyer and Dye, 2012) or being uninformed (Einhorn and Ziv, 2008) can also arise. I abstract away from these and focus on how investor learning affects the trade-off between disclosure and nondisclosure.

---

8See Beyer et al. (2010) for a review of voluntary disclosure models.
Another important assumption is truth telling, which is consistent with prior research (Verrecchia, 1983; Dye, 1985). One motivation for truth telling is the existence of an outside monitor and court that can detect and punish lying managers. The release of actual earnings at the end of each period limits the extent to which the manager manipulates voluntary disclosure (e.g., Gigler and Hemmer, 1998). Stein (1989) shows that, if managers cannot commit to truth telling, in equilibrium, they will inflate earnings to sell their stocks at higher prices. However, the market perfectly figures out the amount of manipulation. This type of manipulation does not create information asymmetry and therefore does not affect the mechanism of investor learning. It does, however, require de-biasing the observed earnings process, which complicates the empirical analysis. I do not pursue this more complex structure.

### 2.2.2 Utility

The manager’s **current-period utility** derived from disclosure decision $d_{it}$ is

$$U_{it}(d_{it}) = P^c(d_{it}, \mathcal{H}_{it}),$$

(1)

where $P^c(d_{it}, \mathcal{H}_{it})$ is the stock price inclusive of the expected costs of current and future disclosures conditional on history $\mathcal{H}_{it}$ and $d_{it} \in \{s_{it}, \emptyset\}$, the disclosure decision of firm $i$ at time $t$. History $\mathcal{H}_{it}$ includes all past earnings and disclosures, $\mathcal{H}_{it} \equiv (y_{it_{-1}}, d_{it_{-1}})'$, where $y_{i,t-1} \equiv (y_{i1}, y_{i2}, \ldots, y_{i,t-1})'$ and $d_{i,t-1} \equiv (d_{i1}, d_{i2}, \ldots, d_{i,t-1})'$.

This utility specification says the manager is rewarded for a higher stock price, which many voluntary disclosure models use (e.g., Verrecchia, 1983; Dye, 1985; Acharya et al., 2011). Stock price is the sum of investors’ expectations of discounted unrealized earnings net of current and future disclosure costs,

$$P^c(d_{it}, \mathcal{H}_{it}) = \sum_{t' = t}^{T} \left\{ \beta_i^{t'-1} E(y_{it'} - 1 \{d_{it'} = s_{it'}\} c_i | d_{it}, \mathcal{H}_{it}) \right\},$$

(2)

where $c_i$ is the disclosure cost and $\beta_i$ is the discount rate for firm $i$, both exogenous to managers’ disclosure decisions.

In principle, one can imagine cases in which the manager cares about both current and future
stock prices. In my model, the manager is rationally myopic in that she only cares about the current stock price. Once the actual earnings \( y_{it} \) is revealed, the manager’s private information \( s_{it} \), which is a noisy signal of the actual earnings, does not provide investors with additional information about underlying firm profitability.\(^9\) It follows that the disclosure decision of period \( t \) will not affect investors’ beliefs or stock prices beyond period \( t \).\(^10\) Thus, the manager only needs to be concerned about the effect of her disclosure decision on the current stock price. Note the myopia in my setting does not entail sacrificing long-term benefits for short-term gains (e.g., Stein, 1989). It simply refers to the fact that the manager maximizes her current utility when making the disclosure decision.

### 2.2.3 Investor learning

To price the firm’s shares, investors form beliefs about the earnings distribution. I assume that the earnings of firm \( i \) follow an AR(1) process:

\[
y_{it} = \rho_{0i} + \rho_{1i} y_{i,t-1} + \epsilon_{it}, \quad \text{where} \quad \epsilon_{it} \sim N(0, \frac{1}{\tau_i}),
\]

(3)

where I call \( \rho_{0i} \) “firm profitability” for simplicity, because it is linear with the actual long-run earnings \( \frac{\rho_{0i}}{1-\rho_{1i}} \). An AR(1) process is the simplest way to capture persistent performance.\(^11\)

The key assumption is that investors do not know the firm’s profitability \( \rho_{0i} \) and learn it through the manager’s disclosures and financial reports (recall that the manager’s disclosures come before the release of actual earnings). Investors’ prior beliefs about \( \rho_{0i} \) at the beginning of period \( t \) before observing the manager’s disclosure decision follow a normal distribution:

\[
\rho_{0i} \sim N(q_{0it}, \frac{1}{\lambda_{it} \tau_i}),
\]

where \( q_{0it} \) denotes investors’ average assessment of \( \rho_{0i} \) at the beginning of period \( t \) and \( \lambda_{it} \) denotes

\(^9\)This result follows from \( s_{it} = y_{it} + \eta_{it} \), where \( \eta_{it} \) is random noise. See equation (5) and the appendix for derivation.

\(^10\)As in Verrecchia (1983), voluntary disclosure is not efficient. If the manager can commit not to disclose and to wait for the earnings announcement, both the manager and investors are better off. However, commitment to non-disclosure is impossible because investors interpret nondisclosure as bad news and the manager cares about investors’ responses.

\(^11\)Prior research has found mixed evidence on the time-series properties of the annual earnings process, suggesting its complexity. Ball and Watts (1972) describe annual earnings as a submartingale process. Watts and Leftwich (1977) find evidence that supports a random walk model. Albrecht et al. (1977) find the choice of AR1 or MA process depends on the firm’s industry.
investors’ information precision relative to the precision of earnings shocks. I call \( \lambda_{it} \) investors’ information precision for short and refer to its reciprocal as investor uncertainty. Note \( 1/\lambda_{it}\tau_i \), the variance of investors’ beliefs about \( \rho_{0i} \), has nothing to do with whether their information is close to the truth. In other words, it is possible that investors’ beliefs are wrong. For example, investors might strongly believe (that is, \( \lambda_{it} \) is large) that a firm will generate good earnings (that is, \( q_{0it} \) is high). But more optimistic beliefs do not necessarily reflect higher profitability, that is, higher \( \rho_{0i} \).

Optimistic beliefs could be a result of a history of lucky firm performance.

Conditional on history \( H_{it} \equiv (y'_{i,t-1}, d'_{i,t-1})' \), where \( y_{i,t-1} \equiv (y_{i1}, y_{i2}, ..., y_{i,t-1})' \) and \( d_{i,t-1} \equiv (d_{i1}, d_{i2}, ..., d_{i,t-1})' \), I express investors’ beliefs and information precision at the beginning of \( t \) as

\[
\lambda_{it} = \lambda_{i,t-1} + 1, \quad q_{0it} = (\lambda_{it})^{-1}(\lambda_{i,t-1}q_{i,t-1} + y_{i,t-1} - \rho_{1i}y_{i,t-2}).
\]

One can see posterior beliefs are the weighted average of prior beliefs and new information. The weights are proportional to the relative precision of prior beliefs (\( \lambda_{i,t-1} \)).

After observing the manager’s voluntary disclosure of \( s_{it} \), investors update their beliefs to

\[
\begin{align*}
\lambda_{st}^* &= \frac{\kappa}{\kappa + 1} + \lambda_{it}; \\
q_{0st}^* &= (\lambda_{st}^*)^{-1}\left(\frac{\kappa}{\kappa + 1}(s_{it} - \rho_{1i}y_{i,t-1}) + \lambda_{it}q_{0it}\right),
\end{align*}
\]

where \( \kappa \) measures the manager’s information precision (recall \( s_{it} \equiv y_{it} + \eta_{it} \) with \( \eta_{it} \sim N(0, \frac{1}{\kappa_{ii}\tau_i}) \)).

Modeling investor learning as above assumes disclosure is only informative about firm earnings. Trueman (1986) finds management forecast issuances provide investors with information about the manager’s ability to adjust production in response to economic changes. I do not separately model the manager’s ability. Instead, I use a reduced-form earnings process to capture both firm fundamentals and the manager’s ability. The choice is consistent with management forecasts being mainly informative about firm fundamentals (Ball and Shivakumar, 2008).

Voluntary disclosure accelerates the release of earnings news. Investors interpret nondisclosure

\[\text{12The higher } \kappa \text{ is, the more weight investors put on manager’s disclosure. Note that investors will never put 100\% weight on } s_{it} \text{ unless } \lambda_{it} = 0. \text{ The reason is that the manager’s disclosure is at best one realization from the earnings distribution, and investors’ beliefs always rationally depend on the full history of earnings and disclosures.}\]
as bad news. Subject to this pressure, the manager only releases news that is sufficiently good. I implicitly assume that the manager cannot commit to ignoring investors’ short-term responses before earnings announcements.\textsuperscript{13} Consistent with this assumption, Chen et al. (2011) document that managers stop providing earnings forecasts when firm performance is bad and investors react negatively to the decision to stop providing forecasts. Moreover, I assume the timing of the disclosure decision does not endogenously affect investor learning about profitability and does not affect the disclosure decision.\textsuperscript{14} Finally, disclosures impose costs on investors, and are inefficient from an ex-ante perspective (Verrecchia, 2001). That is, if investors and managers can commit not to disclose, investors will gain by saving disclosure costs. The reasons why such commitment does not occur might be that investors urgently need information, or that managers are subject to the short-term pressure of increasing market value or reducing uncertainty. My model does not address this issue.

### 2.2.4 Summary of the model

The model uses eight parameters to characterize four related times series of firm \( i \), that is, the manager’s disclosure decisions, the contents of disclosure, firm stock prices, and realized earnings.

1. Firm earnings follow an AR1 process, 
   \[
   y_{it} = \rho_{0i} + \rho_{1i} y_{i,t-1} + \epsilon_{it} \sim \mathcal{N}(0, \frac{1}{\tau_i})
   \]

2. Investors’ beliefs at the beginning of period \( t \) before observing the manager’s disclosure decision follow 
   \[
   \mathcal{N}(q_{0it}, \frac{1}{\lambda_{it} \tau_i})
   \]
   Investors obey the Bayes rule.

3. Stock price is investors’ expectation of the sum of discounted future earnings minus disclosure costs. The discount factor, \( \beta_i \), captures how heavily investors discount future earnings.

4. The manager has private information at the beginning of each period, which is a noisy signal of the earnings at the end of the period, 
   \[
   s_{it} = y_{it} + \eta_{it} \sim \mathcal{N}(0, \frac{1}{\kappa_i \tau_i})
   \]
   The manager’s disclosure decision trades off the stock price following disclosure with that following nondisclosure. Disclosure incurs an exogenous cost \( c_i \).

\textsuperscript{13}If the manager cares about stock price only after the revelation of true earnings and does not care about stock price at the time of disclosure, in the context of this model, she will be indifferent between disclosing and not disclosing because current disclosure does not affect future stock price after earnings release. Other disclosure incentives would be needed for the model to work.

\textsuperscript{14}Acharya et al. (2011) and Guttman et al. (2014) study the timing of voluntary disclosure and how it could affect investors’ inference. These models require more structure on the signals and are not studied.
Except for the learning parameters, investors’ average belief $q_{0i}$ and investors’ information precision $\lambda_{it}$, all other parameters are exogenous to the manager’s disclosure decisions and are assumed to be fixed over time. Some parameters are useful mainly for the empirical analysis. They increase the model fit and provide benchmarks for the effects of investor learning. The effects of investor learning do not qualitatively depend on the value of earnings persistence $\rho_{1i}$, the discount factor $\beta_{i}$, or the manager’s information precision $\kappa_{i}$ ($\kappa_{i}$ cannot be zero).

2.3 Equilibrium strategy

In this subsection, I describe the equilibrium. As discussed above, what matters for the manager’s disclosure decision is the current stock price. I re-write stock price $P^{c}(d_{t}, \mathcal{H}_{t})$ as

$$P^{c}(d_{t}, \mathcal{H}_{t}) = P(d_{t}, \mathcal{H}_{t}) - 1\{d_{t} = s_{t}\}c - \beta G_{d_{t}c},$$

(8)

where $P(d_{t}, \mathcal{H}_{t})$ is the stock price excluding the disclosure costs, and $G_{d_{t}c}$ is investors’ expectation of the sum of the discounted future disclosure cost following the current disclosure decision $d_{t}$. I ignore subscript $i$ to save notation.

Proposition 1 characterizes the equilibrium disclosure strategy.

**Proposition 1.** The manager’s disclosure strategy.

The manager follows a threshold strategy in each period and discloses if and only if her private information $s_{t}$ exceeds a threshold $s_{t}^{*}$. The threshold $s_{t}^{*}$ is solved from:

$$P(s_{t}^{*}, \mathcal{H}_{t}) - c = P(s_{t} \leq s_{t}^{*}, \mathcal{H}_{t}).$$

(9)

The result of a threshold strategy depends on the linearity of $P(d_{t}, \mathcal{H}_{t})$ (stock price excluding disclosure costs) and the independence of $d_{t}$ and $G_{d_{t}c}$ (investors’ expectation of future disclosure costs), both of which are proved in the appendix. Specifically, $P(d_{t}, \mathcal{H}_{t})$ is linear in the manager’s private information (when $d_{t} = s_{t}$) or investors’ expectation of the manager’s private information (when $d_{t} = \emptyset$),

$$P(d_{t}, \mathcal{H}_{t}) = A(\kappa, \lambda_{t}, \beta, \rho_{1})s_{t}(d_{t}) + B(\kappa, \lambda_{t}, \beta, \rho_{1}, p_{0t}, y_{t-1}),$$

(10)
where $A$ and $B$ are expressed as

\[
A = \frac{1}{1 - \beta \rho_1} \frac{1}{\kappa + \lambda_1 \lambda_t + 1} \left( 1 + \frac{\beta}{(1 - \beta)(\lambda_t + 1)} \right), \quad (11)
\]

\[
B = \frac{1}{1 - \beta \rho_1} \frac{\lambda_t}{\kappa + \lambda_1 \lambda_t + 1} \left\{ \left[ 1 - \frac{\beta \kappa \rho_1}{(1 - \beta) \lambda_t} \right] \rho_1 y_{t-1} + \left[ 1 + \frac{\beta (\kappa + 1)}{1 - \beta} \right] p_{0t} \right\}. \quad (12)
\]

Next, I derive the disclosure threshold $s^*_t$ and the probability of disclosure $\Pr\{d_t = 1\}$:

\[
s^*_t = \sigma_{st} \tilde{s}^*_t + q_{0t} + \rho_1 y_{t-1}, \quad (13)
\]

\[
\Pr\{d_t = 1\} = 1 - \Phi\left( \frac{\sigma_{st} \tilde{s}^*_t + q_{0t} - \rho_0}{\sigma_s} \right), \quad (14)
\]

where $\tilde{s}^*_t$ is a function of firm characteristics excluding $q_{0t}$ and $\rho_0$ and is always negative, $\sigma_{st}$ is the standard deviation of the manager’s private information including investor uncertainty about $\rho_0$, and $\sigma_s$ is the actual standard deviation of the manager’s private information. Specifically,

\[
\sigma_{st} = \sqrt{\frac{\lambda_t + 1}{\kappa + \lambda_1 \lambda_t + 1}} \frac{1}{\tau}; \quad (15)
\]

\[
\sigma_s = \sqrt{\frac{1}{\kappa \tau}}. \quad (16)
\]

2.4 The effects of investor learning

I examine two effects of investor learning. First, I examine how investor learning affects the disclosure incentive in a given period. I measure the disclosure incentive as the probability of disclosure. Second, I examine how a change in the current disclosure incentive induced by investor learning affects future disclosure incentives, that is, the persistence of disclosure incentives.

The effects of investors’ beliefs and uncertainty on the probability of disclosure directly follow from equation (14), and are summarized in Corollary 1. I discuss the intuition for this result in the next subsection.

**Corollary 1.** The effects of investor learning on disclosure decisions.

More pessimistic investors’ beliefs and greater investor uncertainty increase disclosure.
One might argue that, because earnings are persistent, better past earnings should increase disclosure. The following proposition says this intuition is wrong.

**Corollary 2. Disclosure dynamics and past performance.**

Better past earnings are neither necessary nor sufficient to increase disclosure.

On the one hand, the expectation of future earnings $\rho_0 + \rho_1 y_{t-1}$ increases with $y_{t-1}$, which increases the chances that the manager observes better private information. On the other hand, from (13), increase in $y_{t-1}$ also enhances the disclosure threshold. The net effect is that they cancel each other. Thus the disclosure probability in (14) does not depend on $\rho_1 y_{t-1}$. I discuss the intuition in greater detail in the next subsection.

Next, I show investor learning induces persistent disclosure incentives, that is, an increase in the manager’s current likelihood of disclosure is associated with an expected increase in the likelihood of future disclosure. The persistence results from the stickiness of investor learning. Specifically, disclosure probability in period $t+m$ ($m \geq 1$) depends on investors’ beliefs in period $t+m$ (see (14)), which in turn depend on their beliefs in period $t$ (see (5)). Therefore, shocks to investors’ beliefs $q_{0it}$ will change the disclosure probability in period $t$ and, through the dependence of investors’ beliefs over time, change the disclosure probability $t+m$ in the same direction.

The correlation between disclosure incentives in period $t$ and $t+m$ decreases as $m$ increases, because investors receive more information over time, which reduces the effect of shocks to their current beliefs on beliefs further into the future. I summarize the results in Corollary 3 and discuss the intuitions in the following subsection.

**Corollary 3. The effects of investor learning on disclosure dynamics.**

Investor learning induces persistent disclosure incentives, that is, the correlation between the likelihood of disclosure in period $t$ and $t+m$ is positive. The persistence decreases with the time lag $m$.

### 2.5 The effects of other model parameters

This section presents the effects of other model parameters on the manager’s disclosure incentive, that is, the likelihood of disclosure.
Corollary 4. A lower discount rate $\beta$, higher managerial information precision $\kappa$, a lower disclosure cost $c$, higher profitability $\rho_0$, and higher earnings persistence $\rho_1$ are associated with a higher likelihood of disclosure, holding constant other parameters.

I leave the proof in the appendix and only discuss the intuition for why these parameters affect disclosure. The effects of these parameters provide benchmarks for the effects of investor learning in the empirical analysis.

With the exception of profitability $\rho_0$ and disclosure cost $c$, other parameters mainly affect the likelihood of disclosure through the sensitivity of stock price to disclosure (see (11)), with higher sensitivity increasing disclosure. This is analogous to higher earnings response coefficient (ERC). For example, Lennox and Park (2006) find higher ERC is associated with more disclosure, which is consistent with the result here. Specifically, a lower discount rate $\beta$ makes future earnings more important to stock price; higher managerial information precision $\kappa$ increases the weight investors put on the manager’s disclosures; higher earnings persistence makes current disclosure more important in affecting investors’ expectation of future earnings. All of these increase the sensitivity of stock price to the manager’s disclosure.

Finally, a higher disclosure cost $c$ makes disclosure more expensive and therefore reduces the disclosure incentive. Higher profitability $\rho_0$ increases the probability that the manager’s information exceeds the threshold.

2.6 Discussion

2.6.1 Investor learning and voluntary disclosure incentives

Investors’ learning consists of two components, their average assessment of underlying firm profitability, and how uncertain they are about their beliefs. Corollary 1 describes how each component affects managers’ disclosure incentives, that is, disclosure probabilities.

First, managers facing more pessimistic (about the firm’s profitability) investors are more likely to disclose. More pessimistic investor beliefs widen the range of news that managers can disclose to improve investors’ beliefs and firm stock price, which increases the likelihood of disclosure. The result differs from the expectation adjustment hypothesis (e.g., Ajinkya and Gift, 1984), which states that managers provide forecasts when “there is strong likelihood of investor dissatisfaction
resulting from their decisions to rely on prevailing unrealistic market expectations." First, the expectation adjustment hypothesis predicts no disclosure when expectations are aligned, whereas in my model, managers are still likely to disclose, because no disclosure is interpreted as bad news. Chen et al. (2011) and Houston et al. (2010) provide evidence consistent with my model. Second, my model has no explicit cost derived from belief misalignment, so managers do not have incentives to guide investors’ expectations downward when investors are optimistic. This result contrasts with some prior findings using analyst forecasts (e.g., Cotter et al., 2006; Feng and Koch, 2010) that managers are more likely to provide guidance when analysts are optimistic. But it is consistent with Bergman and Roychowdhury (2008), who find managers are more likely to provide earnings forecasts when investor sentiment is bad. Analyst optimism and investor optimism are likely to be different. Analyst optimism might result from incentives to cater to client firms (e.g., Lim, 2001), and managers provide guidance to adjust this bias. My model cannot speak to this channel.

The result implies high profitability itself is insufficient to induce more disclosure. A profitable firm with optimistic investors could disclose less than a less profitable firm with pessimistic investors. Equation (14) shows managers’ disclosure incentives decrease with investor optimism, that is, the likelihood of disclosure decreases as the difference between investors’ beliefs about profitability and the actual profitability increases. If investors know profitability, disclosure probability will be independent of profitability.

Second, higher investor uncertainty increases the likelihood of disclosure. For example, suppose the manager does not disclose a forecast given some disclosure cost. The result says an increase in investor uncertainty can result in disclosing the forecast. Specifically, as investor uncertainty increases, investors put less weight on their prior beliefs and more weight on the manager’s disclosure (see (5)), resulting in a larger reaction to the manager’s disclosure decision (see (10)). The stock price following disclosure becomes higher, and the stock price following nondisclosure becomes lower. When investor uncertainty is sufficiently high, the price difference can be large enough to induce the manager to disclose the forecast.

More formally, higher investor uncertainty increases the sensitivity of stock price to disclosure, which in turn increases disclosure incentives. When investors are more uncertain, they rely more on the manager’s disclosure (see (11), sensitivity channel). In addition, because investors do not know firm profitability, the greater their uncertainty is, the more they will revise their beliefs about firm
profitability after observing the manager’s disclosure decision (see (11), future expectation channel). Both channels increase the sensitivity of stock price to the disclosure decision. Higher sensitivity of stock price to disclosure magnifies the difference between stock prices following disclosure and nondisclosure, which in turn increases the likelihood of disclosure.

Finally, Corollary 2 clarifies the role of past earnings in affecting managers’ disclosure incentives. It says better past earnings are neither necessary nor sufficient for more future disclosure. On the one hand, better past earnings are associated with better current earnings, which increases the manager’s disclosure incentive. On the other hand, investors correctly anticipate the increase in current earnings, which increases stock price following nondisclosure, reducing the manager’s disclosure incentive. The two effects cancel each other out. The result underscores the importance of accounting for investors’ beliefs when examining managers’ disclosure incentives.

2.6.2 Investor learning and voluntary disclosure dynamics

Corollary 3 describes how investor learning affects the persistence of managers’ disclosure incentives. Recall that I define disclosure incentives as the likelihood of disclosure. Disclosure incentives are persistent if an increase in the current disclosure probability is associated with increases in future disclosure probabilities.

Before discussing the results, I should note that I examine disclosure persistence holding firm characteristics constant. If a firm changes characteristics such as its earnings persistence, changes in current disclosure incentives due to the change in earnings persistence will perfectly persist into the future. I do not examine this form of persistence. Instead, persistence in the model originates from the way investors learn, holding constant firm characteristics.

Corollary 3 says investor learning induces persistent disclosure incentives. This effect stems from the stickiness of investor learning. For example, suppose investors’ beliefs improve after observing a positive earnings shock. The manager’s disclosure incentive will decrease in response to the increase in investors’ beliefs (Corollary 1). Moreover, because of the dependence of investors’ future beliefs on their current beliefs, increase in investors’ current beliefs leads to expected increases in their future beliefs, implying that the manager’s future disclosure incentives will also decrease in expectation (holding all else constant). Therefore, through the stickiness of investor learning, change in the current disclosure probability induced by investor learning is positively associated
with changes in future disclosure probabilities.

The result indicates that disclosure incentives can be persistent because of investor learning. A caveat is that I do not address and compare other mechanisms that also lead to stickiness outside this framework. Doing so requires incorporating these mechanisms into the model, which is not the primary focus of this paper. Identifying them from data poses another challenge. For instance, Einhorn and Ziv (2008) show past nondisclosure can increase future nondisclosure because investors are more convinced the manager has no information, and punish non-disclosure less. Beyer and Dye (2012) show reputational concern about being forthcoming can motivate managers to provide more disclosure. These incentives are hard to identify from the data.

3 Data and sample selection

I examine management earnings forecast decisions for the empirical analysis. I collect management forecasts from Thomson Reuters I/B/E/S Guidance issued by US firms between January 2003 and December 2014. I only use forecasts issued after January 2003 because the coverage is more comprehensive. Prior research typically uses data from First Call’s Company Issued Guidance (CIG), the predecessor of I/B/E/S Guidance. Chuk et al. (2013) document that about 41% of their hand-collected forecasts are not covered by the CIG database. I show in the appendix that I/B/E/S Guidance is superior to CIG in the coverage comprehensiveness, and that I/B/E/S Guidance is comparable to the hand-collected results of Chuk et al. (2013).\footnote{See Table 1 for the comparison between I/B/E/S Guidance and the hand-collected sample of Chuk et al. (2013) and Figure 1 for the comparison between I/B/E/S Guidance and CIG.}

Following Houston et al. (2010), I drop forecasts issued after fiscal year-end dates, namely pre-announcements. I focus on EPS forecasts, because the model is silent about how to map other types of management forecasts to earnings expectations and stock prices. I keep both bundled and non-bundled forecasts.

For each firm, I assign forecasts to the firm’s fiscal years according the forecast announcement dates. For example, for a firm whose fiscal year ends in December, forecasts issued between Jan. 1, 2011 and Dec. 31, 2011 are assigned to fiscal year 2011. If no forecast can be found for a fiscal year, I assume no forecast was issued within that fiscal year, \( d_{it} = \emptyset \). I perform the empirical analysis at an annual frequency, because I am less likely to incorrectly assign nondisclosure to a fiscal year.
because of poor data coverage. Using annual forecasts avoids the need to model seasonality of earnings and investors’ interpretation of that seasonality, which could result in a complicated stock price formula and a bad model fit. Table 2 describes the sample construction in greater detail.

Finally, to allow a reasonable amount of time for the market to react, I use the stock price three days following the forecast announcement date. In the case of nondisclosure, I use the stock price at the end of the final trading date of the fiscal year. The assumption is that investors should realize the lack of disclosure by then. I require at least five consecutive years of nonmissing earnings (from I/B/E/S) and stock prices (from CRSP).

4 Descriptive analysis

In this section, I establish two empirical regularities regarding the dynamics of management earnings forecast decisions. First, forecast decisions are sticky and mean-reverting over time. Second, firm performance cannot explain the mean-reverting pattern.

The purpose of the descriptive analysis is to demonstrate that disclosure decisions do not necessarily behave like a fixed policy but rather vary over time in a predictable way. It does not aim to support investor learning as the only explanation. Rather, the entire empirical analysis assumes learning exists. For example, Hutton and Stocken (2009), among others, provide evidence that investor reaction to management earnings forecasts is consistent with a Bayesian learning model. The empirical analysis in subsequent sections focuses on whether investor learning can explain management forecast decisions in terms of its empirical (marginal) effects on the likelihood of forecast provision and its intertemporal dynamics.

4.1 Dynamics of management forecast issuance

Table 3 shows managers issue forecasts frequently, about 87% of all firm-years. About 57% of the firms in my sample always issue forecasts in the sample period (untabulated). This finding is not surprising, because I require a firm to issue at least one forecast within a fiscal year. These firms are still useful in understanding the average tendency to issue forecasts. However, the question of how past and future disclosure incentives are related is not meaningful for these firms, because the
Figure 2 plots the cross-sectional density of forecast frequencies between 2003 and 2014 for firms that do not issue forecasts every year. It shows forecast frequencies have a large cross-sectional variance. Some firms issue forecasts only once or twice, whereas other firms are more consistent, implying not all managers can “commit” to a disclosure policy. The finding is consistent with the finding of Rogers et al. (2009) that management forecasts can be regular or sporadic.

Next, I examine times-series properties of forecast decisions. Figure 3 shows management forecast decisions are sticky and mean-reverting. For each year, I sort firms based on their forecast decisions and track their forecast decisions over the following six years. I adjust for the general trend of management forecast issuance by subtracting the average forecasting frequency in each year. The pattern shown in the figure is robust to this adjustment. Firms that issue forecasts in the current year have higher disclosure frequencies over the following six years than firms that don’t issue forecasts. This finding is consistent with stickiness. On the other hand, disclosure frequencies tend to go down (up) for forecasting (non-forecasting) firms, which suggests the decisions are mean-reverting. If providing a forecast is associated with an increase in investors’ beliefs, the subsequent decline in the disclosure frequency is consistent with the pattern predicted by investor learning.

4.2 Firm performance and management forecast dynamics

The patterns that forecast provisions are sticky and mean-reverting could reflect many things. One driver is firm performance. Prior studies show managers are more likely to disclose better performance (e.g., Chen et al., 2011; Houston et al., 2010), and to be persistent in their disclosure decisions due to persistently good performance (Miller, 2002). Because performance (return on assets) is mean-reverting, disclosure decisions can appear mean-reverting.

Figure 4 shows forecast frequencies do not always track performance, measured as return on assets (ROA). I sort firms based on their disclosure decisions and track their ROAs over the following six years. Figure 4 presents the median ROA each year after firms’ disclosure decisions. The upper dashed line shows the median ROA falls over time for firms that issue forecasts, consistent with the disclosure pattern in Figure 3. In contrast, the lower solid line shows that the median ROA among

\[16\] In other words, it is impossible to run a regression of current disclosure decisions on past disclosure decisions for these firms, because disclosure decisions have zero variance.
firms that do not issue forecasts first rises and then falls. The pattern does not follow the disclosure pattern for these firms. Thus, performance does not entirely explain the disclosure pattern.

### 4.3 Summary statistics of variables used for estimation

I now describe the data used for the empirical analysis. Each firm has four time series: earnings, management forecasts, disclosure decisions (that is, the existence of a management forecast), and stock prices. To increase comparability across firms, I scale all variables by stock price at the beginning of the sample period. I present the summary statistics in Table 3.

Earnings are on average 12% of the initial firm stock price and have a large standard deviation of 18%, suggesting the need to account for the cross-sectional differences among firms. Earnings forecasts are close to earnings but are slightly larger, consistent with managers’ preference for disclosing good news (e.g., Chen et al., 2011).

Firms in the sample experience significant growth. The average buy-and-hold return from the beginning of the sample period is about 114%. The result is consistent with the fact that earnings per share at the end of the sample period are on average 3.52 times as large as earnings per share at the beginning of the sample period. The observation raises the question of whether a fixed AR1 earnings process is appropriate in capturing such a level of growth. Rationality dictates that investors assign equal weights to all past earnings when forming their posterior beliefs. For a firm that experiences high growth and whose earnings distribution might change over time, one might assume earnings further in the past become increasingly less relevant in predicting future earnings. To capture this effect, one can propose more complicated behavioral models such as assuming a hidden long-run earnings process that evolves over time (e.g., Pástor and Stambaugh, 2012). Doing so requires identifying the hidden process, which is hard given the length of the time series (12 years), and is not further pursued. Assuming equal weighting is a low-cost starting point. If earnings tend to rise over time, investors’ prior beliefs are likely to be overestimated, because higher prior beliefs can “compensate” for the excessive weights that are put on low past outcomes.
5 Structural estimation

In this section, I structurally estimate the model presented in section 2. Structural estimation differs from a linear regression in that, rather than imposing a linear relation among variables, I assume the structure of the model holds for the data. I choose structural estimation for four reasons. First, the model suggests performance, its persistence, and investor learning interact in a nonlinear way, which makes a linear OLS regression misspecified. Second, measuring beliefs using analyst forecasts, implied volatility from option prices or stock returns could confound the effects of firm fundamentals and investor learning. Third, identifying the causal mechanism of investor learning requires finding an instrument that is related to investors’ beliefs but not firm fundamentals. Finding an instrument is next to impossible, because belief changes necessarily result from changes in firm fundamentals. More importantly, even if such an instrument existed, pinning down how the instrument affects beliefs is hard. Fourth, the empirical analysis requires simulating many samples for a given firm, which linear regressions cannot achieve.

Structural estimation can address these aforementioned issues. It directly imposes the theory model on the data and thus explicitly accounts for the non-linear relations described by the model. Investors’ beliefs are estimated from the data as parameters, which bypasses the need for measurements and IVs. The causal mechanism is transparent from the model structure. And structural estimation recovers the data-generating process for all the variables in the data, which can be used to simulate new samples.

One drawback of structural estimation is its reliance on the model’s assumptions. Assuming investors are Bayesian and learn about firm profitability is not an unreasonable starting point and has some empirical support (e.g., Pástor and Veronesi, 2003; Hutton and Stocken, 2009). Nevertheless, agents will never behave exactly as the model predicts. What matters is whether the model reasonably describes the behavior of economic agents.

Another drawback is that structural estimation excludes many incentives identified by prior literature on management forecasts, which include but are not limited to equity issuance (Frankel et al., 1995), warning investors (Kasznik and Lev, 1995), litigation risk (Skinner, 1994), aligning investors’ expectations (Ajinkya and Gift, 1984), and corporate governance (Ajinkya et al., 2005). Structural estimation prohibits incorporating all of them because they should be explicitly modeled.
For instance, incorporating equity issuance requires modeling managers’ capital structure decisions. Inferences will not be affected if these incentives do not affect how investors learn from managers’ disclosure decisions. To the extent that they interact with investor learning, the current learning model can be viewed as a benchmark, and other incentives can modify investors’ or managers’ behavior.

Using structural estimation, I examine two aspects of the intertemporal dynamics of managers’ disclosure incentives. As in the model, I define disclosure incentive as the probability of providing a management forecast. First, I estimate how investor learning affects the average future disclosure frequency measured over a given period. Quantifying this effect is of first order importance in light of the high disclosure frequency observed in the data. The extent to which investors’ prior beliefs set disclosure incentives for the ensuing years is by itself an interesting research question. Second, I quantify how a change in the likelihood of current disclosure induced by investor learning relates to the likelihood of future disclosure, which is relevant for how we think about the observed persistence of disclosure decisions.

Because I impose the model structure, investor learning by construction affects managers’ disclosure decisions. However, even if we accept that investor learning exists and that the model holds, whether investor learning is associated with economically important effects is unclear. Other model parameters such as earnings persistence, profitability, managers’ information precision, and disclosure cost could all be the main drivers of disclosure decisions, which could make investor learning de-facto useless. Quantifying the effects of investor learning thus allows researchers to evaluate the importance of investor learning by itself and relative to other model parameters.

5.1 Identification

Model parameters are identified through how they affect the joint likelihood of stock prices (see (8)), management earnings forecast decisions (see (14)), the contents of management forecasts, and realized earnings (see (3)). To capture the heterogeneity across firms, I allow for the parameters to differ across firms.\textsuperscript{17} I present the joint likelihood function and the model structures and discuss

\textsuperscript{17}Allowing firm-level parameters to vary over the cross section is called hierarchical modeling. The idea resembles conditional logit, where economic agents make sequential choices, for example, choosing whether to take buses or taxis to work and, conditional on choosing buses, which bus routes to take. Here the parameters of each firm are distributed according to a population distribution, which is parameterized to be normal. Hierarchical modeling is useful when one expects the economic relations of interest differ cross-sectionally and aims to estimate such heterogeneity.
the estimation procedure in the appendix. I discuss parameter identifications below.

First, differences between actual earnings and management forecasts of firm $i$ identify $\kappa_i$, the information precision of firm $i$’s manager. Larger differences imply smaller $\kappa_i$, that is, lower precision. The identification of $\kappa$ relies on the relation between actual earnings and management forecasts, which is assumed to be $s_{it} = y_{it} + \eta_{it}$, where $\eta_t \sim \mathcal{N}(0, \frac{1}{\kappa_i \tau_i})$.

Second, the disclosure cost $c_i$ is identified through two channels. The disclosure cost enters stock price directly when a management forecast is issued (see (8)). Differences between stock prices following management forecast issuance and nondisclosure are informative about the cost. Moreover, a higher disclosure cost is associated with lower stock prices. Identifying $c_i$ requires the additive separability of $c_i$ from other components of stock price. Identifying $c_i$ is harder in cases in which managers always or never issue forecasts. For example, if a manager always discloses, then $c_i = 0$ can always rationalize the disclosure decisions. In such cases, identifying $c_i$ occurs through the functional form of stock prices.

Third, investors’ belief precision $\lambda_{i0}$ and the discount factor $\beta_i$ are identified from the sensitivity of stock prices to management forecasts. The relation between prices and earnings identifies the discount factor. Investors’ belief uncertainty varies over time according to the functional form of (4). The resulting time-varying relations between stock prices and management forecasts identify $\lambda_i$.

Fourth, investors’ prior belief $q_{0i0}$ is identified through the (time-varying) intercept of the price-forecast relation (see (10)). Identifying $q_{0i0}$ depends on the assumption that investors follow the Bayes rule. Finally, the time series of earnings is informative about $\rho_{0i}$, $\rho_{1i}$, and $\tau_i$.\(^{18}\)

6 Results

Structural estimation recovers the cross-sectional distribution of each model parameter. I first present the distribution of these parameters and discuss their magnitude. I then examine the model fit of disclosure decisions and stock prices. Finally, I examine the effects of investor learning on the average disclosure frequency and the persistence in the likelihood of disclosure. Because each firm has a different set of parameters, the effects of investor learning are heterogeneous across

---

\(^{18}\)Because identifying these parameters does not depend solely on disclosure frequency, no mechanical relation exists between disclosure decisions and the value of these parameters. Moreover, due to the existence of other parameters that also affect managers’ disclosure decisions, the economic magnitude of the effects of investors’ beliefs and their uncertainty do not mechanically vary with observed disclosure frequency, either.
firms. I focus on the average effects as well as the cross-sectional heterogeneity in these effects.

6.1 Parameter estimates

Structural estimation recovers the joint distribution of model parameters for each firm. For each firm, I compute the median of each parameter (see section 2.2.4 for the list of all parameters) and report in Table 4 the cross-sectional distribution of these medians.

First, investors’ beliefs about firm profitability (the first row) are different from the actual firm profitability identified using earnings data (the sixth row). For about 61% of firms, investors’ prior beliefs are lower than the estimated profitability from the data. Investors’ beliefs are not necessarily irrational. They might be justified by soft information not easily conveyed by management forecasts or by historical information not included in the sample. Exploring the data suggests firms with high investor prior beliefs have high growth.

Next, investors are confident about their prior beliefs. The average (median) belief precision is 5.572 (5.183), compared with the average (median) management forecast precision of 4.032 (2.918). Investors will put significant weight on their own prior beliefs after observing new information. Exploring the data suggests high belief precision is associated with low volatility of firm earnings, consistent with the intuition that investor uncertainty is lower for stable firms.

For other parameter estimates, the average (median) firm earnings persistence $\rho_{1i}$ is 0.63 (0.64) with a relatively heavy left tail. The magnitude is consistent with industry average profit persistence estimated by prior research (e.g., Richardson et al., 2005; Dichev and Tang, 2009; Frankel and Litov, 2009). The average (median) expected return, computed using $\beta_i = \frac{1}{1+r_i}$, is 11% (8%) with a standard deviation of 9%. The magnitude is comparable to the expected return estimated by Easton (2004), about 13%, with a standard deviation of 3.9%.

Finally, the average (median) disclosure cost implied by the model is 5.7% (4.6%) of the initial stock price. This cost is very large compared with the distribution of earnings (exclusive of disclosure cost) with a mean of 12.3% and a median of 9.1% of the initial stock price. The disclosure cost can be interpreted as frictions that prevent the manager from disclosing. Its magnitude suggests the model misses a significant portion of cross-sectional or time series variation of disclosure decisions or both. Any residual variation of forecast decisions that the model does not capture will be treated as reasons for nondisclosure and will appear in the disclosure cost estimates. Alternatively,
part of the cost of disclosure could actually be non-monetary, which do not affect stock prices. The former issue requires more rigorous theoretical models. To the extent that omitted variable problems certainly exist, future research can extend the current model and consider those omitted factors. To explore the issue of monetary versus nonmonetary costs, I estimate the structural model separating the disclosure cost into a component that only enters the manager’s disclosure decision process (nonmonetary) and a second component that also enters the stock price (monetary). The median monetary disclosure cost (untabulated) is 2.6% of the initial stock price, compared with the median disclosure cost (sum of both components) of 7.1%. Both are still very high relative to realized earnings (i.e., almost 20% of earnings). Subsequent analysis relies on results from the original model that does not differentiate between monetary and nonmonetary disclosure costs.

6.2 Model fit

Table 5 indicates the structural model on average fits disclosure decisions and stock prices well. As a benchmark for the model fit of disclosure decisions, I run a simple linear regression of disclosure decisions on past performance (ROA) (e.g., Chen et al., 2011), pooling all firm-year observations. I then examine whether the structural model explains the variation of disclosure decisions better than this linear model.

I use the predictive (non-)disclosure probability to measure the model fit of disclosure decisions. Specifically, if a manager issues (does not issue) a forecast in a particular period, I compute the (non)disclosure probability using the model. Then for each firm, I average the predictive probabilities across all periods. Higher predictive probabilities indicate a better within-sample prediction and thus a better model fit.

In the first row of Table 5, I present the distribution of the ratio of the model fit of the structural model to the benchmark linear model. Figure 5 plots the distribution of model fit for disclosure decisions across firms. On average, the structural model fits better than the benchmark model.

To measure the model fit of stock prices, I compute the ratio of the standard deviation of the model residual to the standard deviation of stock prices. The ratio is similar to the $R^2$. The difference is that the standard deviation of the model residual can be larger than the standard deviation of stock prices, which makes the measure larger than 1. In other words, it can happen that using the average stock price to explain stock prices produces better model fit than using
the structural model. The higher the ratio is, that is, the higher the standard deviation of model residual relative to the standard deviation of stock prices, the worse the model fits stock prices.

In the second row of Table 5, I present the cross-firm distribution of the model fit of stock prices. For most firms, the structural model has a better fit than using the average stock price. The average (median) ratio is 0.704 (0.687).

6.3 Investor learning and the expected disclosure frequency

I now examine the effects of investor learning on the average future disclosure incentive. I use the frequency of forecast issuance over the subsequent 12 years to measure the average future disclosure incentive. The purpose is to understand whether investors’ prior beliefs and uncertainty have economically significant effects in the presence of other model parameters.

I compute disclosure frequencies via simulation for two reasons. First, I do not have a close-form solution for the marginal effects of investor learning on the average future disclosure frequency. Second, simulation allows me to generate stationary firms, which resemble firms from the estimation sample but do not experience shocks to other parameters that could confound the effects of investor learning. For example, earnings persistence $\rho_1$ could vary for a given firm because of random shocks, which changes its disclosure frequency. Simulation allows me to control for such random variation by fixing $\rho_1$.

I first generate 2,000 firms using the cross-sectional distributions of firm level parameters (see section 2.2.4 for the list of all parameters). These firms resemble the firms in the estimation sample. For each simulated firm, I simulate 1,000 samples each with 12 years, using the corresponding firm level parameters. Then for each sample, I compute the percentage of the total number of disclosures out of the 12 years. I obtain the disclosure frequency for each firm by averaging over the 1,000 sample disclosure frequencies.\(^{19}\)

To ease comparison among the marginal effects, I vary each parameter by one standard deviation (computed from its cross-sectional distribution) while holding other parameters constant. I measure the marginal effects as the change in the future disclosure frequency as a proportion of the standard deviation of the disclosure frequency in the data, which is 22%. For example, the standard deviation

\(^{19}\)I could use parameters of firms in the sample to perform the simulation instead of simulating new firms. One benefit of the latter is that simulation allows researchers to generate firms that belong to the same population but whose parameters do not appear in the sample. Another benefit is that it can produce large sample size.
of $\lambda_i$ is 3.057. For each firm, I vary $\lambda_i$ by 3.057 holding other parameters constant, and compute the change in the future disclosure frequency, divided by 22%. For investors’ beliefs, I also compute the marginal effects of investors’ beliefs by varying them by one standard deviation of the firm earnings shock. The reason is that the standard deviation of investors’ prior beliefs is 30%—too large to be considered “marginal,” especially given the standard deviation of earnings of 17%. Using earnings shock is consistent with the idea that investors’ beliefs change should relate to firm earnings shocks.

Figure 6 and Table 6 present the distribution of the simulated disclosure frequencies and the marginal effects of selected model parameters. The simulated disclosure frequencies are more dispersed than the data. Although the cross-sectional median of the simulated disclosure frequency (98.5%) is close to the data (100%), the first quartile (45.1%) is much lower than the data (83%). One reason is that the simulation generates firms that differ from those in the estimation sample, although their parameters belong to the same cross-sectional distribution implied by the sample.

Table 6 reveals the effects of investor learning, that is, the effects of investors’ prior beliefs and uncertainty, are heterogeneous across firms. The marginal effects are zero for about 40% of firms (the final column). The reason is that these firms always disclose, and a marginal change in one parameter is not sufficient to change their disclosure incentives.

Next, conditional on the marginal effects of investor learning being non-zero, the results demonstrate that investor learning has economically significant effects on managers’ future disclosure incentives. A one standard deviation change in investors’ prior beliefs on average changes future disclosure frequency by 15-25 percentage points, which is about one standard deviation of the empirical disclosure frequency. As discussed before, the change in investors’ prior beliefs might be too large relative to the standard deviation of earnings. The third row presents the change in disclosure frequency from varying investors’ prior beliefs by one standard deviation of each firm’s earnings shock. The effects become more moderate, with an average of 5 percentage points, or 0.23 standard deviations of the empirical disclosure frequency. Finally, about 5% of firms have effects larger than 17 percentage points, about 0.77 standard deviations of the empirical disclosure frequency.

Investor uncertainty has similar effects on the average disclosure frequency to investors’ prior beliefs. The average marginal effect for a one standard deviation change in investors’ belief precision (the reciprocal of investors’ belief uncertainty) is about 6-8 percentage points or about 0.32 standard deviations of the empirical disclosure frequency. For 5% of firms, the marginal effects are larger
than 35 percentage points or about 1.60 standard deviations of the empirical disclosure frequency.

To understand the magnitude of the effects above, I compute the marginal effects of other parameters and present them in Table 6. I examine the marginal effects of firm risk, earnings persistence, firm profitability, managers’ information precision, and disclosure cost. Similar to investor learning, about 30%-40% of firms have zero marginal effects associated with a one standard deviation change in one of these parameters.

Conditional on the marginal effects being non-zero, the effects of investor learning are about half the size of earnings persistence and firm risk and resemble the effects of firm profitability and managers’ information precision. Specifically, earnings persistence and firm risk are associated with an average marginal effect of 12 percentage points—larger than those of investor learning.20 Firm profitability and managerial information precision have an average of 4-10 percentage points across firms—close to the effects of investor learning. Disclosure costs have the largest effect with a cross-sectional average of 19-24 percentage points.

Overall, changes in investors’ prior beliefs and investor uncertainty significantly shift managers’ disclosure incentives over the following 12 years for about 60% of firms. The effects are expected to be larger for a shorter window, because changes in investors’ beliefs have larger effects on the likelihood of disclosure in the near future. The results illustrate that investors’ initial perception about the firm has a meaningful impact on managers’ disclosure decisions over a reasonably long time.

### 6.4 Investor learning and persistent disclosure incentives

In this section, I examine the effects of investor learning on the persistence of managers’ disclosure incentives. As in the previous sub-section, I measure disclosure incentives by simulating managers’ disclosure probabilities. I focus on the ex-ante incentive to disclose, not actual disclosure decisions, because the model does not have any empirical predictions regarding the correlation between current and future disclosure decisions. I consider disclosure incentives to be persistent if changes in the current disclosure incentive are positively associated with changes in future disclosure incentives.

To estimate the persistence in the likelihood of disclosure, I generate 1,000 firms using the

---

20 Recall that firm risk affects how heavily the future is discounted. Heavier discounting, or a smaller discount factor, decreases the effect of the manager’s disclosure on investors’ expectation about future performance. The effect reduces the price reaction to the manager’s disclosure, which decreases her disclosure incentive.
distribution of parameter estimates. For each firm, I simulate 500 samples each with 7 years, totaling 3,500 observations for each firm. I compute the probability of disclosure as in (14). All samples of a given firm share the same set of model parameters. The simulation procedure ensures variation in investors’ beliefs (caused by random shocks to earnings) is solely responsible for the variation in disclosure probabilities over time.

To understand how investor learning affects the dynamics of disclosure incentives, for each firm, I regress the disclosure probabilities of years 3-7 each on the disclosure probabilities of year 2. The slope coefficients of the five regressions capture the associations between the disclosure probability in year 2 and the disclosure probabilities in years 3-7. I have the following model:

\[
\Pr\{d_{i,t+m} = s_{i,t+m} | I_{i,t-1}\} = \alpha_{i,m} + \beta_{i,m} \Pr\{d_{i,t} = s_{i,t} | I_{i,t-1}\} + \varepsilon_{i,t+m}
\]

I report in Table 7 the cross-sectional distribution of the regression slopes, \(\beta_{im}\).

Table 7 shows investor learning on average generates persistent disclosure incentives. The first row suggests that an increase of 10% in the current disclosure probability is associated with an 10.3% increase in the disclosure probability of the next year. The effect is not significantly different from 10%. So changes in the current disclosure incentive almost perfectly persist into the following year.

Table 7 also suggests persistence monotonically declines over time but remains economically significant. A 10% increase in the current disclosure probability is associated with a 7.5% increase in the disclosure probability in five years, which is still an economically meaningful effect.

Overall, the evidence suggests investor learning affects the persistence of managers’ disclosure incentives, measured as disclosure probabilities. The empirical magnitude of the effects is non-trivial, making investor learning an important factor driving the intertemporal dynamics of managers’ voluntary disclosure decisions.

\footnote{Beliefs only start to vary starting from the second year, because the initial beliefs are assumed to be the same among all the samples of a given firm.}
7 Conclusion

I model and empirically quantify the effects of investor learning on the intertemporal dynamics of managers’ voluntary disclosure decisions. My model suggests that more pessimistic investor prior beliefs and higher investor uncertainty increase disclosure and that investor learning induces persistence in the likelihood of disclosure.

Empirically, I structurally estimate the model and recover the cross-sectional distribution of the model parameters. I find investor learning has economically meaningful effects on managers’ future disclosure incentives for 60% of firms. A one standard deviation change in investors’ prior beliefs and investor uncertainty are associated with a change of around 0.23-0.32 standard deviation in the average disclosure frequency for the following 12 periods. The economic effects of investor learning are comparable to those of other model parameters, such as firm profitability and managers’ information precision.

I also find investor learning induces persistent disclosure. A 10% increase in the current disclosure probability is associated with an increase the disclosure probability of the following year by around 10.3%, and the disclosure probability in five years by about 7.4%.

Overall, this paper shows the importance of investor learning in explaining managers’ disclosure decisions. Future work can extend the current paper along two dimensions. First, limited research exists on how investors’ information processing affects disclosure decisions. The topic is important because investors are consumers of information. One extension might be to model more realistic learning. For example, Pástor and Stambaugh (2012) find stock returns are more volatile in the long run because of investor learning. In their model, the distribution of firm earnings changes over time, and investors do not perfectly observe the changes. Rogers et al. (2009) find implied volatility increases after the release of sporadic earnings forecasts. How managers disclose when uncertainty could increase after disclosure is theoretically and empirically interesting.

Another extension might be to incorporate other reporting incentives. For instance, managers could have incentives to bias disclosures (e.g., Fischer and Verrecchia, 2000). Bias could alter disclosure incentives because extreme disclosures might indicate misreporting. Moreover, in addition to its effect on stock price, the possibility of missing her own forecast can be costly to the manager (e.g., Feng and Koch, 2010). How this incentive arises as an equilibrium outcome and how it affects
the dynamics of disclosures are both interesting questions.

Second, future research can incorporate other incentives that arise in the multi-period setting and show how they affect the dynamics of managers’ disclosure decisions. For example, future research can model managerial reputational concern, the timing of disclosure, or managerial types and compare their effects with investor learning. The general issue is that investor learning or other multi-period incentives such as reputation are usually not directly observable. Econometricians must rely on observable data to identify them. Structural estimation is a good way to deal with this issue.
References


A Appendix Proofs

Lemma 1. Call \( h(s) \equiv E(\tilde{s}|\tilde{s} \leq s) \), where \( \tilde{s} \) follows a standard normal distribution, and call \( g(s) \equiv s - h(s) \). Then \( h'(s) > 0 \) and \( g'(s) > 0 \).

Proof.

\[
\begin{align*}
    h'(s) &= -\frac{\phi'(s)\Phi(s) - \phi^2(s)}{\Phi^2(s)} \\
         &= \frac{s\phi(s)\Phi(s) + \phi^2(s)}{\Phi^2(s)}.
\end{align*}
\]

When \( s \geq 0 \), \( h'(s) > 0 \). Suppose \( \exists s \) s.t. \( h'(s) < 0 \), which implies at least one \( s \) exists such that \( s = -\frac{\phi(s)}{\Phi(s)} = E(\tilde{s}|\tilde{s} \leq s) \), a contradiction.

The proof of \( g'(s) > 0 \) can be found in the appendix of Verrecchia [1983].

Lemma 1 says both the nondisclosure payoff and the payoff difference between disclosing at the threshold and nondisclosure are increasing with the threshold. Uniqueness can be proved by using the lemma after defining \( \tilde{s}_t \equiv (s_t - \mu)\sqrt{\tau} \) and rewriting the equilibrium condition as

\[
\tilde{s}_t^* - E(\tilde{s}_t|\tilde{s}_t \leq \tilde{s}_t^*) = c\sqrt{\tau}.
\]\n
A.1 Multi-period Bayesian updating

In this subsection, I describe the Bayesian updating process following any arbitrary history. Recall that I have firm profit following an autoregressive process, that is,

\[
y_{t+1} = \rho_0 + \rho_1y_t + \epsilon_{t+1}, \quad \epsilon_{t+1} \sim \mathcal{N}(0, \frac{1}{\tau}),
\]

where \( \tau \) is the precision parameter known by all, and \( \rho_0 \) is unknown to firm investors who learn them from the manager’s disclosures and financial reports that are described in the next paragraph. Investors’ prior belief about \( \rho_0 \) at the beginning of \( t \) is given by \( \rho_0 \sim \mathcal{N}(q_t, \frac{1}{\lambda_0\tau_t}) \) with \( \lambda_0\tau_t \) denoting the precision of investors’ beliefs, that is, the inverse of the variance of investors’ beliefs.

In period \( t \), the sequence of events is as follows.
1. At the beginning of each period, the firm manager observes a private signal \( s_t \sim \mathcal{N}(y_t, \frac{1}{\kappa t}) \) about unrealized period-\( t \) profit. If the signal is disclosed, investors will observe \( s_t \). If not, investors rationally expect \( s_t \leq s_t^* \), where \( s_t^* \) is an endogenous threshold to be calculated.

2. At the end of the period, \( y_t \) is released in a financial report and becomes public knowledge.

The objective is to describe investor learning given any history, where history at the beginning of each period is denoted by \( \mathcal{H}_t \), with \( \mathcal{H}_t \equiv (s'_{t-1}, x'_{t-1}, y'_{t-1}), y_{t-1} \equiv (y_{t-1}, y_{t-2}, ..., y_1)', x_{t-1} \equiv (y_{t-2}, y_{t-3}, ..., y_0)' \) and \( s_{t-1} \equiv (s_{t-1}, s_{t-2}, ..., s_1)' \). Below, I derive the posterior distribution of \( \rho_0 \) given history \( \mathcal{H}_t \), the prior predictive distribution of \( s_t \) given \( \mathcal{H}_t \), and the posterior predictive distribution of \( y_t \) given \( s_t, \mathcal{H}_t \).

### A.1.1 Learning \( \rho \)

The joint density of \( y_{t-1}, x_{t-1}, s_{t-1}, \rho \) is

\[
f(y_{t-1}, x_{t-1}, s_{t-1}, \rho) \propto e^{-\frac{1}{2} \left( (y_{t-1} - X_{t-1} \rho)' (y_{t-1} - X_{t-1} \rho) + \kappa (s_{t-1} - y_{t-1})' (s_{t-1} - y_{t-1}) + \lambda_0 (\rho_0 - q_0)^2 \right)}
\]

where \( X_t \equiv (1, x_t) \) and \( \rho \equiv (\rho_0, \rho_1)' \). The prior distribution of \( \rho_0 \) following history \( \mathcal{H}_t \) can be obtained by completing the square.

\[
f(\rho_0 | \mathcal{H}_t) = f(\rho_0 | y_{t-1}, x_{t-1}, s_{t-1})
\]

\[
\propto e^{-\frac{1}{2} \left( (t-1 + \lambda_0) \rho_0^2 - 2 (t-1)(\bar{y}_{t-1} - \rho_1 \bar{x}_{t-1}) + \lambda_0 \rho_0 \right)}
\]

where \( \bar{y}_t \) and \( \bar{x}_t \) refer to the average of \( y_t \) and \( x_t \) up to time \( t \).

The posterior distribution of \( \rho_0 \) is normal:

\[
\rho_0 \sim \mathcal{N}\left( \frac{(t-1)(\bar{y}_{t-1} - \rho_1 \bar{x}_{t-1}) + \lambda_0 q_0}{t-1 + \lambda_0}, \frac{1}{(t-1 + \lambda_0)\tau} \right),
\]  

(18)
I denote the mean and precision (exclusive of \(\tau\)) of the prior distribution of \(\rho_0\) after \(H_t\) as:

\[
\lambda_t = t - 1 + \lambda_0 = \lambda_{t-1} + 1; \quad (19)
\]

\[
q_t = (\lambda_t)^{-1}((t-1)(\bar{y}_{t-1} - \rho_1 \bar{x}_{t-1}) + \lambda_0 q_0) = (\lambda_t)^{-1}((\lambda_{t-1} q_{t-1} + y_{t-1} - \rho_1 y_{t-2})). \quad (20)
\]

After observing \(s_t\), investors update their beliefs about \(\rho_0\). The wrinkle is that \(s_t\) is a noisy version of \(y_t\). The posterior belief about \(\rho_0\) can be obtained by integrating \(y_t\) from the joint distribution of \(y_t, x_t, s_t, \rho_0\):

\[
f(y_t, x_t, s_t, \rho_0) \propto e^{-\frac{1}{2} \left\{ (y_{t-1} - X_{t-1} \rho)'(y_{t-1} - X_{t-1} \rho) + (y_t - X_t \rho)^2 + \kappa (s_t - y_t)^2 + \lambda_0 (\rho_0 - q_0)^2 \right\}}.
\]

Integrating away \(y_t\) gives

\[
f(y_{t-1}, x_{t-1}, s_t, \rho_0) \propto e^{-\frac{1}{2} \left\{ (y_{t-1} - X_{t-1} \rho)'(y_{t-1} - X_{t-1} \rho) + \kappa (s_t - X_t \rho)^2 + \lambda_0 (\rho_0 - q_0)^2 \right\}}.
\]

The posterior distribution is then obtained by completing the square for \(\rho_0\):

\[
\rho_0 \sim N \left( \frac{\kappa}{\kappa + 1} \left( \frac{\kappa}{\kappa + 1} (s_t - \rho_1 y_{t-1}) + \lambda_t q_t \right), \frac{1}{\left( \frac{\kappa}{\kappa + 1} + \lambda_t \right) \tau} \right). \quad (21)
\]

I denote the mean and precision (exclusive of \(\tau\)) of the posterior distribution of \(\rho_0\) after \(H_t\) as

\[
\lambda^*_t = \frac{\kappa}{\kappa + 1} + \lambda_t; \quad (22)
\]

\[
q^*_t = (\lambda^*_t)^{-1} \left( \frac{\kappa}{\kappa + 1} (s_t - \rho_1 y_{t-1}) + \lambda_t q_t \right). \quad (23)
\]

### A.1.2 Prior predictive distribution of \(y_t\) before observing \(s_t\)

In this section, I derive the prior predictive distribution of \(y_t\) given \(H_t\) and before observing \(s_t\). We have:
\[ f(y_t|\mathcal{H}_t) = \int f(y_t|\rho, \mathcal{H}_t) f(\rho_0|\mathcal{H}_t) d\rho_0 \]
\[ = \int f(y_t|\rho, y_{t-1}) f(\rho_0|\mathcal{H}_t) d\rho_0 \]
\[ \propto \int e^{-\frac{1}{2} \left\{ [y_t - (\rho_0 + \rho_1 y_{t-1})]^2 + \lambda_t (\rho_0 - q_t)^2 \right\}} d\rho_0. \]

Integrating away \( \rho_0 \) shows that the posterior predictive of \( y_t \) after history \( \mathcal{H}_t \) is normal:
\[
y_t|\mathcal{H}_t \sim N\left(q_t + \rho_1 y_{t-1}, \frac{1}{\left(\frac{\lambda_t}{\lambda_t+\tau}\right)}\right).
\]

I denote the mean and precision (exclusive of \( \tau \)) as
\[
\Lambda^y_t = \frac{\lambda_t}{\lambda_t + 1};
\]
\[
\mu^y_t = q_t + \rho_1 y_{t-1}.
\]

### A.1.3 Prior predictive (marginal) distribution of \( s_t \)

With the prior predictive distribution of \( y_t \) after \( \mathcal{H}_t \), I can now derive the marginal distribution of \( s_t \) as
\[
f(s_t|\mathcal{H}_t) = \int f(s_t, y_t, \mathcal{H}_t) dy_t = \int f(s_t|y_t, \mathcal{H}_t) f(y_t|\mathcal{H}_t) dy_t \]
\[ \propto \int e^{-\frac{1}{2} \left\{ \kappa (s_t - y_t)^2 + \Lambda^y_t (y_t - \mu^y_t)^2 \right\}} dy_t,
\]
where the second line follows because \( s_t \) depends only on \( y_t \). Integrating away \( y_t \) gives the marginal distribution of \( s_t \) given history \( \mathcal{H}_t \):
\[
s_t|\mathcal{H}_t \sim N\left(\mu^y_t, \frac{\kappa + \Lambda^y_t}{\kappa\Lambda^y_t\tau}\right).
\]
Substituting the expressions of $\mu^y_t$ and $\Lambda^y_t$, we have

$$s_t|\mathcal{H}_t \sim \mathcal{N}\left(q_t + \rho_1 y_{t-1}, \left(\frac{\lambda_t + 1}{\lambda_t} + \frac{1}{\kappa}\right) \frac{1}{\tau}\right). \quad (28)$$

Note as $k \to \infty$, the precision term in (28) converges to $\frac{\lambda_t}{\lambda_t + 1}$, which is the same as the precision term of the prior predictive distribution of $y_t$ in (27). This argument verifies (28) because they should be exactly the same when $s_t$ is essentially the same as $y_t$.

**A.1.4 Posterior predictive of $y_t$ given $s_t$**

This section derives the posterior predictive of $y_t$ given $\mathcal{H}_t$ and $s_t$:

$$f(y_t|s_t, \mathcal{H}_t) \propto f(s_t, y_t, \mathcal{H}_t) = f(s_t|y_t)f(y_t|\mathcal{H}_t)$$

$$\propto e^{-\frac{1}{2} \left\{ \kappa(s_t-y_t)^2 + \Lambda^y_t(y_t-\mu^y_t)^2 \right\}},$$

where the second line follows from $s_t$ depending only on $y_t$. Completing the square gives the posterior predictive of $y_t$ given $\mathcal{H}_t$ and $s_t$:

$$y_t|s_t, \mathcal{H}_t \sim \mathcal{N}\left(\frac{\kappa s_t + \Lambda^y_t \mu^y_t}{\kappa + \Lambda^y_t}, \frac{1}{\tau(\kappa + \Lambda^y_t)}\right).$$

Substituting the expressions of $\mu^y_t$ and $\Lambda^y_t$, we have

$$y_t|s_t, \mathcal{H}_t \sim \mathcal{N}\left(\frac{\kappa s_t + \Lambda^y_t (q_t + \rho_1 y_{t-1})}{\kappa + \frac{\lambda_t}{\lambda_t + 1}}, \frac{1}{\tau(\kappa + \frac{\lambda_t}{\lambda_t + 1})}\right). \quad (29)$$

The higher $s_t$ is, the larger the posterior mean is.

**A.2 Derivation of stock price exclusive of disclosure cost**

To derive stock price exclusive of disclosure cost, the first step is to compute investors’ expectation of all future profits following $s_t$ and $\mathcal{H}_t$. 

44
First, the expected profit \( m \) period after observing \( s_t \) in period \( t \) is

\[
E(y_{t+m} | s_t) = E \left( E(y_{t+m} | \rho, y_t) | s_t \right) \\
= E \left( \rho_0 + \rho_1 \rho_0 + \rho_1^2 \rho_0 + \ldots + \rho_1^{m-1} \rho_0 + \rho_1^m y_t | s_t \right) \\
= q_t^s \left( 1 + \rho_1 + \rho_1^2 + \ldots + \rho_1^{m-1} \right) + \rho_1^m \frac{\kappa s_t + \Lambda_t^y \mu_t^y}{\kappa + \Lambda_t^y}. \tag{30}
\]

Then the sum of the discounted expected future profits (infinite horizon) is

\[
\sum_{m=1}^{\infty} \beta^m \left\{ q_t^s \left( 1 + \rho_1 + \rho_1^2 + \ldots + \rho_1^{m-1} \right) + \rho_1^m \frac{\kappa s_t + \Lambda_t^y \mu_t^y}{\kappa + \Lambda_t^y} \right\} \\
= \sum_{m=1}^{\infty} \beta^m \left\{ q_t^s \left( \frac{1 - \rho_1^m}{1 - \rho_1} \right) + \rho_1^m \frac{\kappa s_t + \Lambda_t^y \mu_t^y}{\kappa + \Lambda_t^y} \right\} \\
= \sum_{m=1}^{\infty} \beta^m \left\{ \frac{q_t^s}{1 - \rho_1} + \rho_1^m \left( \frac{\kappa s_t + \Lambda_t^y \mu_t^y}{\kappa + \Lambda_t^y} - \frac{q_t^s}{1 - \rho_1} \right) \right\} \\
= \beta \left\{ \frac{q_t^s}{(1 - \rho_1)(1 - \beta)} + \frac{\rho_1}{1 - \beta \rho_1} \left( \frac{\kappa s_t + \Lambda_t^y \mu_t^y}{\kappa + \Lambda_t^y} - \frac{q_t^s}{1 - \rho_1} \right) \right\} \\
= \frac{\beta}{1 - \beta \rho_1} \left\{ \frac{q_t^s}{1 - \beta} + \rho_1 \frac{\kappa s_t + \Lambda_t^y \mu_t^y}{\kappa + \Lambda_t^y} \right\} \\
= \frac{\beta}{1 - \beta \rho_1} \left\{ \frac{\kappa}{\kappa + 1} (s_t - \rho_1 y_{t-1}) + \lambda_t q_t \right\} + \rho_1 \frac{\kappa s_t + \lambda_t (q_t + \rho_1 y_{t-1})}{\kappa + \lambda_t}. \tag{31}
\]

which is increasing in \( s_t \) and may increase or decrease with respect to \( \rho_1 \).

Adding \( E(y_t | H_t, s_t) \) to the expression above and regrouping terms produces stock price exclusive of disclosure cost. Stock price is linear in \( s_t \):

\[
P_t(s_t) = A_t s_t + B_t,
\]

45
where

\[
A_t = \frac{\beta}{1 - \beta \rho_1} \left\{ \frac{\kappa}{\kappa + 1} + \rho_1 \frac{\kappa}{\kappa + 1 + \lambda_t} \right\} + \frac{\kappa}{\kappa + 1 + \lambda_t}
\]

\[
= \frac{\beta}{1 - \beta \rho_1} \frac{\kappa}{\kappa + 1 + \lambda_t} (1 - \beta) + \frac{1}{1 - \beta \rho_1} \frac{\kappa}{\kappa + 1 + \lambda_t}
\]

\[
= \frac{1}{1 - \beta \rho_1} \frac{\kappa}{\kappa + 1 + \lambda_t} \left[ 1 + \frac{1}{(1 - \beta)(\lambda_t + 1)} \right],
\]

and

\[
B_t = \frac{1}{1 - \beta \rho_1} \left\{ \left[ \frac{\beta}{1 - \beta} \frac{\lambda_t}{\kappa + 1 + \lambda_t} + \frac{\lambda_t}{\kappa + 1 + \lambda_t} \right] q_t + \left[ \frac{\lambda_t}{\kappa + 1 + \lambda_t} - \frac{\beta}{1 - \beta} \frac{\kappa}{\kappa + 1 + \lambda_t} \right] \rho_1 y_{t-1} \right\}
\]

\[
= \frac{1}{1 - \beta \rho_1} \frac{\lambda_t}{\kappa + 1 + \lambda_t} \left\{ \left[ 1 + \frac{\beta}{1 - \beta}(\kappa + 1) \right] q_t + \left[ 1 - \frac{\beta}{1 - \beta} \frac{\kappa}{\lambda_t} \right] \rho_1 y_{t-1} \right\}.
\]

A.3 Proof of Proposition 1

The optimal disclosure strategy would be a threshold strategy if the stock price did not include future expected disclosure cost \( G_{d,c} \), because \( P(d_t, H_t) \) is linear and increasing in the signal \( s_t \) (see equation (10)). However, investors’ expectation of future disclosure costs following disclosing a signal could differ from that following not making disclosures. Below, I prove that the expected future discounted disclosure costs are the same following any disclosure decision, which is summarized in Lemma 2.

Lemma 2. The expected future disclosure costs do not depend on the current disclosure decision, that is, \( G_{d_t=s_t} = G_{d_t=\emptyset} \).

The independence of expected future disclosure costs from the current disclosure decision follows from (1) the assumption of normal distribution, and (2) that investors observe realized earnings \( y_t \) at the end of the period. I provide a sketch of the proof below and a more formal one following the sketch.

Proof. Because investors do not know the actual distribution of the manager’s private signals, their expectations of future disclosure probabilities depend only on their perceived distribution of the manager’s future private signals, which is normal conditional on the history of earnings and
disclosures (see equations (4) and (5)). In principle, one needs to simulate paths of realized future earnings to obtain the expected distribution of manager’s future signals. Fortunately, investor uncertainty about $\rho_0$ evolves deterministically according to equation (4). Because other variance terms are public knowledge, investors’ perceived variances of the manager’s future signals are also deterministic and, in particular, do not depend on the manager’s current disclosure decision. Moreover, with a normally distributed signal, the probability of disclosure depends only on the signal variance.\(^{22}\) Therefore, future disclosure probabilities do not depend on the current disclosure decision.

I prove Lemma 2 formally below.

Suppose there are $T$ periods. Let $G_d^m$ denote the expected total discounted disclosure costs $m$ periods before $T$ following disclosure decision $d \in \{0, 1\}$.

First, suppose $m = 1$. In period $T$, the probability of disclosure conditional on $H_T = (H_{T-1}, s_{T-1}, y_{T-1})'$ is obtained from

\[
E(y_T|s_{T}^{**}, H_T) - c = E(y_T|s_T \leq s_{T}^{**}, H_T)
\]

\[
\Rightarrow \frac{\sum^0_T}{\kappa + \sum^0_T} \left( s_{T}^{**} - E(s_T|s_T \leq s_{T}^{**}) \right) = c
\]

\[
\Rightarrow \frac{\sum^0_T}{\kappa + \sum^0_T} \left( s_{T}^{*} - \mu_T^0 \right) \sqrt{\frac{\kappa \sum^0_T}{\kappa + \sum^0_T}} - E \left( \frac{\kappa \sum^0_T}{\kappa + \sum^0_T} s_T \leq s_{T}^{**} \right) = c \sqrt{\frac{\kappa \sum^0_T}{\kappa + \sum^0_T}}.
\]

Note that I use investors’ perceived distribution in (28) to compute the probability of disclosure rather than the true distribution, because what matters here is investors’ perceived probability of disclosure given their information set. From the calculation, one can see that the disclosure probability at $T$ does not depend on realization of $y_{T-1}$ and $s_{T-1}$. At $T - 1$, the expected cost of disclosure is

\[
G_d^1 = cE \left[ Pr \{ s_T \geq s_{T}^{**}|H_T \} \right]
\]

\[
= c \int \int Pr \{ s_T \geq s_{T}^{**}|s_T-1, y_T-1 \} f(s_{T-1}, y_{T-1}|d_{T-1}) dy_{T-1} ds_{T-1}
\]

\[
= c \eta \left( \frac{\kappa \sum^0_T}{\kappa + \sum^0_T} \right),
\]

\(^{22}\)Recall a normal distribution can always be expressed relative to the mean of the distribution.
with the final line following from the probability of disclosure depending only on the precision. It immediately follows that \( G_d^1 \) does not depend on \( d \). Therefore, the disclosure decision at period \( T - 1 \) can be modeled without incorporating the expected future disclosure cost.

The rest of the proof uses backward induction. Suppose the expected disclosure cost \( m \) periods ahead, \( G_d^m \), does not depend on \( d_{T-m} \). I prove below that the expected disclosure cost \( m + 1 \) periods ahead, \( G_d^{m+1} \), does not depend on \( d_{T-m-1} \).

In period \( T-m \), the probability of disclosure conditional on \( H_{T-m} = (H_{T-m-1}, s_{T-m-1}, y_{T-m-1})' \) is:

\[
E(x_{T-m}|s_{T-m}^{**}, H_{T-m}) + \sum_{t=T-m+1}^{T} \{ \beta^{t-(T-m)} x_t | s_{T-m}^{**}, H_{T-m} \} - c - \beta G_d^m
\]

\[
= E(x_{T-m}|s_{T-m} \leq s_{T-m}^{**}, H_{T-m}) + \sum_{t=T-m+1}^{T} \{ \beta^{t-(T-m)} x_t | s_{T-m} \leq s_{T-m}^{**}, H_{T-m} \} - \beta G_0^m
\]

\[
\Rightarrow \Gamma_{T-m} \left( s_{T-m}^{**} - E(s_{T-m}|s_{T-m} \leq s_{T-m}^{**}) \right) = c,
\]

\[
\Rightarrow \Gamma_{T-m} \left( s_{T-m}^{**} - \mu_{T-m}^0 \right) \sqrt{\frac{\kappa \Sigma_{T-m}^0}{\kappa + \Sigma_{T-m}^0}} - E \left( \left( s_{T-m} - \mu_{T-m}^0 \right) \sqrt{\frac{\kappa \Sigma_{T-m}^0}{\kappa + \Sigma_{T-m}^0}} | s_{T-m} \leq s_{T-m}^{**} \right)
\]

\[
= c \sqrt{\frac{\kappa \Sigma_{T-m}^0}{\kappa + \Sigma_{T-m}^0}} \tau,
\]

where \( \Gamma_{T-m} \) does not depend on \( H_{T-m} \).

Again, I use investors’ perceived distribution to compute the disclosure probability rather than the actual distribution. The expected cost of disclosure can be computed as follows:

\[
G_d^{m+1} = cE \left[ Pr\{s_{T-m} \geq s_{T-m}^{**}|H_{T-m}\} \right] + \beta G_d^{m+1}
\]

\[
= c \int \int Pr\{s_{T-m} \geq s_{T-m}^{**}|s_{T-m-1}, y_{T-m-1}, H_{T-m-1}\} f(s_{T-m-1}, y_{T-m-1}|d_{T-m-1}, H_{T-m-1})
\]

\[ dy_{T-m-1} ds_{T-m-1} + \beta G^m \]

\[
= c \eta \left( \frac{\kappa \Sigma_{T-m}^0}{\kappa + \Sigma_{T-m}^0} \tau \right),
\]

(33)

with the final line following from the probability of disclosure being independent of \( s_{T-m-1}, y_{T-m-1} \). It immediately follows that \( G_d^{m+1} \) does not depend on \( d \). Therefore disclosure decision at period \( T - m - 1 \) can be modeled without incorporating the expected future disclosure cost.
A.4 The equilibrium condition

Recall stock price excluding disclosure costs is

\[ P_t(s_t) = A_t s_t + B_t, \]

where

\[
A_t = \frac{1}{1 - \beta \rho_1} \left[ 1 + \frac{\beta}{(1 - \beta)(\lambda_t + 1)} \right],
\]

and

\[
B_t = \frac{1}{1 - \beta \rho_1} \left[ \left[ 1 + \frac{\beta}{(\kappa + 1)} \right] q_t + \left[ 1 - \frac{\beta}{1 - \beta \lambda_t} \right] \rho_1 y_{t-1} \right].
\]

Two things are important for the disclosure strategy. The first is the sensitivity of stock price to disclosure \( A_t \), which is increasing in \( \rho_1, \kappa, \) and \( \beta \) and decreasing in \( \lambda_t \). The second is the prior predictive distribution of \( s_t \), which is

\[ s_t | \mathcal{H}_t \sim \mathcal{N} \left( q_t + \rho_1 y_{t-1}, \left( \frac{\lambda_t + 1}{\lambda_t + 2} \right) \right). \]

The variance of \( s_t \) decreases with \( \lambda_t \), investors' information precision, and \( \kappa \), managers' information precision.

Recall that Lemma 1 says that for a random variable \( x \) has a standard normal distribution, the function \( w(s) = s - \mathbb{E}(x|x \leq s) \) increases with \( s \). Now re-write the equilibrium condition as

\[
A_t \left( \sigma_{st} (\bar{s}^*_t - \mathbb{E}(\bar{s}_t|\bar{s}_t \leq \hat{s}^*_t)) \right) = c \implies \hat{s}^*_t - \mathbb{E}(\bar{s}_t|\bar{s}_t \leq \hat{s}^*_t) = \frac{c}{A_t \sigma_{st}}.
\]

One can now see that given RHS, \( \hat{s}^*_t \) is uniquely defined. Note the RHS does not depend on \( \rho_0 \).
or \( q_{0t} \). Hence, the threshold can be expressed as

\[
s_t^* = \sigma_{st} \cdot \tilde{s}_t^* + q_{0t} + \rho_{1yt-1}.
\]

Then the probability of disclosure is:

\[
\Phi \left( \frac{\sigma_{st} \cdot \tilde{s}_t^* + q_{0t} - \rho_0}{\sigma_s} \right)
\]

I now discuss the effects of model parameters on disclosure probabilities.

1. \( \lambda_t \). When \( \lambda_t \) increases, RHS of (34) increases (because of \( A_t \) and \( \sigma_{st} \)), so \( \tilde{s}_t^* \) increases. Because \( \sigma_{st} \) decreases and \( \tilde{s}_t^* < 0 \), \( s_t^* \) increases. Since \( \sigma_s \) is unchanged, disclosure probability decreases.

2. \( \beta \). When \( \beta \) increases, RHS of (34) increases. With the same logic as \( \lambda_t \), \( s_t^* \) increases, which decreases the disclosure probability.

3. \( \rho_1 \). The same reasoning as \( \beta \).

4. \( c \). When \( c \) increases, RHS of (34) increases. The disclosure probability decreases with the same reasoning as above.

5. \( q_{0t} \). When \( q_{0t} \) increases, \( \tilde{s}_t^* \) does not change, so \( s_t^* \) increases, which decreases the disclosure probability.

6. \( \rho_0 \). When \( \rho_0 \) increases, \( \tilde{s}_t^* \) and \( s_t^* \) both stay the same, but \( \sigma_{st} \cdot \tilde{s}_t^* + q_{0t} - \rho_0 \) decreases, which increases the disclosure probability.

7. \( \kappa \). When \( \kappa \) increases, \( A_t \) increases, so \( \sigma_{st} \cdot \tilde{s}_t^* \) decreases, that is, becomes more negative. In the meantime, \( \sigma_s \) decreases. The overall effect on \( \frac{\sigma_{st} \cdot \tilde{s}_t^* + q_{0t} - \rho_0}{\sigma_s} \) is decreasing, which increases the disclosure probability.
B CIG and I/B/E/S Guidance

To compare CIG with I/B/E/S Guidance, I compute four measures of data coverage in Figure 1: the percentage of EPS forecasts, the total number of forecasts, the total number of firms, and the percentage of firms that have had I/B/E/S analyst coverage. The figure shows that I/B/E/S Guidance improves over CIG along all dimensions, with the difference more apparent after 2003, the effective date of I/B/E/S Guidance. Specifically, I/B/E/S Guidance does not bias toward including EPS forecasts only. The average percentage of EPS forecasts is 36%, comparable to the hand-collected results of Chuk et al. (2013). The total number of forecasts and total number of firms are also much larger. For example, in 2010, I/B/E/S Guidance covers 3,526 firms, 2.6 times as large as the 1,367 firms covered by CIG. Finally, I/B/E/S Guidance firms have more I/B/E/S analyst coverage.

C Estimation procedure

The structural model has two layers. At the lower level, eight firm-level parameters describe the data generating process: the variance of the earnings process $1/\tau_i$, investors’ initial beliefs $q_{0i}$, investors’ belief precision $1/(\lambda_{0i}\tau_i)$, the cost of disclosure $c_i$, discount factor $\beta_i$, the manager’s information precision $1/(\kappa_{i}\tau_i)$, firm long-run earnings $\rho_{0i}$, and firm performance persistence $\rho_{1i}$. At the upper level, I assume that each firm-level parameter is drawn from a cross-sectional distribution, $f(\theta_i|\vartheta)$, where $\theta_i$ is the firm-level parameter and $\vartheta$ is a vector of hyper-parameters that describe the density of $\theta_i$. I call $\vartheta$ population parameters. When estimating the model, I allow for an unobserved random factor that causes the model generated stock price to deviate from the actual stock price. Because the random factor does not enter the decision process, it does not affect the computation of the equilibrium disclosure decision. The mean of the random factor is zero, and the variance is denoted by $\sigma_{ei}$.

I estimate the model via Gibbs sampling. Gibbs sampling iteratively draws from conditional densities to approximate the joint density. For example, if one is interested in characterizing the joint density of $\alpha, \beta$, one can first draw from the density of $\alpha$ conditional on $\beta$ and then from the

\footnote{See Table 1 for comparison between the two databases by forecast types.}

\footnote{Different types of forecasts are treated as distinct observations. For example, if a manager issues an EPS forecast along with a revenue forecast on the same day, the total number of forecasts are two instead of one.}
density of $\beta$ conditional on the new $\alpha$. Repeating this step many times recovers the joint density of $\alpha$ and $\beta$. I would like to learn the joint distribution of the firm-level parameters as well as the population parameters conditional on the data. I perform Gibbs sampling via two steps.

In the first step, I draw the parameters corresponding to each firm conditional on the population parameters $\vartheta$ and the data. The data points for each firm are disclosure decisions $d_{iT}$, stock prices conditional on the disclosure decisions $p_{iT}$, management earnings forecasts $s_{iT}$, and realized earnings $y_{iT}$, where $T$ is the total number of observations (periods) for the firm and bold letters denote vectors. I express the posterior density conditional on the population parameters and the data, $\pi(\theta_i|y_{iT}, s_{iT}, d_{iT}, p_{iT}, \vartheta)$, as follows:

$$
\pi(\theta_i|y_{iT}, s_{iT}, d_{iT}, p_{iT}, \vartheta) \propto \prod_{t=1}^{T} \left[ \frac{N(p_{it} - P(s_{it}|y_{it}, \theta_i) \, d_{it} \, P(s^\ast_t(y_{it}, \theta_i)|y_{it}, \theta_i)^{1-d_{it}}; 0, \sigma_{ei})}{L(p_{iT}|s_{iT}, y_{iT}, d_{iT}, \theta_i)} \right] \prod_{t=1}^{T} \left[ \frac{N(s_{it}|y_{it}, \frac{1}{\kappa_i \tau_i})}{1 - N(s_{it} \leq s^\ast_t(y_{it}, \theta_i); y_{it}, \frac{1}{\kappa_i \tau_i})} \right]^{d_{it}} \left[ \frac{1}{L(s_{iT}|y_{iT}, d_{iT}, \theta_i)} \right] \prod_{t=1}^{T} \left[ 1 - N(s_{it} \leq s^\ast_t(y_{it}, \theta_i); y_{it}, \frac{1}{\kappa_i \tau_i})^{1-d_{it}} \right] \left[ N(s_{it} \leq s^\ast_t(y_{it}, \theta_i); y_{it}, \frac{1}{\kappa_i \tau_i}) \right] \prod_{t=1}^{T} N(y_{it} - \rho_{0i} - \rho_{1i}y_{it-1}; 0, \frac{1}{\tau_i}) f(\theta_i|\vartheta),
$$

(36)

where $P$ is the model-generated stock price, and $N(\cdot; \mu, \sigma)$ is the normal density function (when admitting scalars) or the cumulative distributional function (when admitting intervals) with mean $\mu$ and variance $\sigma$.

The first line of (36) follows because the posterior density is proportional to the joint distribution of the data and the individual firm-level parameters. Applying the Bayes rule, the joint distribution can be written as the likelihood function of the data conditional on the firm-level parameters times the density of these parameters. The second to fourth lines further separate the likelihood of data into four parts: the likelihood of prices, the likelihood of management forecast contents, the likelihood of disclosure decisions, and the likelihood of realized earnings. The posterior density has no closed form and involves solving nonlinear equations that characterize the disclosure decision for
each firm-year. To draw from this density, I use the Metropolis Hastings (MH) algorithm, which does not require knowing the entire functional form of the posterior density.

In the second step, I draw the population parameters conditional on the individual firm-level parameters. Because all the firm-level parameters are assumed to be positive, I transform them to the real line by taking the natural logarithm, and assume the transformed variables follow normal distribution. Following convention, if a parameter takes value from the positive real line, I assume its natural logarithm follows a normal distribution. These parameters include $q_0, \lambda_0, c, \kappa, \rho_0$, and $\tau$. If a parameter is between 0 and 1, I use a logit kernel, so $\ln \left( \frac{x}{1-x} \right)$ follows a normal distribution. These parameters include $\beta$ and $\rho_1$.

Finally, researchers need to pick their own prior beliefs over these parameters. First, specifying prior beliefs for the individual firm-level parameters is unnecessary. Given the population parameters, neither the likelihood of the data conditional on the firm-level parameters nor the prior density of the firm level parameters admits researcher’s prior beliefs. For the population parameters, I choose prior beliefs with very large variances, which ensures a minimal influence of the econometrician’s prior beliefs on results.

I pick feasible starting values for each firm and repeat the two steps above 20,000 times. I discard the first 10,000 times to minimize the influence of starting values (called “burn-in period” in the literature).

1. Impose necessary parametric assumptions, determine the set of parameters to estimate, and specify prior beliefs (of the econometricians) about their value.

- Parametric assumptions.

Recall investors’ prior beliefs about firm $i$’s profitability $\rho_{i0}$ have a normal distribution $\mathcal{N}(\bar{\rho}_0, \frac{1}{\lambda_0 \tau_0})$. It is public knowledge that, for firm $i$, the disclosure cost is $c_i$, the precision of managers’ information is $\kappa_i \tau_i$, the persistence of profit is $\rho_{1i}$, and the variance of profit is $\frac{1}{\tau_i}$.

- The set of parameters to estimate.

- Population parameters: $\theta \equiv (\bar{\rho}_0, \bar{\rho}_0, \bar{\rho}_1, \bar{\kappa}, \bar{c}, \bar{\tau})'$ and $w \equiv (w_{\rho_0}, w_{\rho_0}, w_{\rho_1}, w_\lambda, w_\kappa, w_c, w_\tau)'$, where $\theta$ is the mean vector and $w$ is the variance vector.

---

25Computation intensity precludes more complex disclosure strategies that have multiple nondisclosure regions.
Individual parameters: \( \vartheta_i \equiv (p_{i0}, \rho_{i0}, \rho_{i1}, \kappa_i, \lambda_i, \tau_i, c_i) \).

- Prior beliefs (of the econometrician)

The prior belief of \( \theta \) is \( \mathcal{N}(0, W_1) \) and that of \( w \) is \( IG(a_1, b_1) \). The distributions of these prior beliefs are chosen to ensure tractable posterior distribution. The parameters of the prior distributions are chosen by the econometricians and might not be innocuous when the number of observations is small. I choose different starting points to ensure stability.

2. Initializing the Markov chain.

I first draw a single vector of population parameters \( \theta \) and \( w \) from its prior distribution. Conditional on the population parameters, I draw firm-level parameters \( \vartheta_i \). These parameters serve as the starting values of the Markov chain.

3. Specify the likelihood function given the data and iterate the Markov chain.

The posterior likelihood of parameters conditional on data is:

\[
\pi(\vartheta | y_t, s_t, d_t, P_t) \propto L(y_t, s_t, d_t, P_t | \vartheta) f(\vartheta | \theta) = \prod_{t=1}^{T} \left[ g(P(s_t | y_T, d_t, \vartheta)) 1^{d_t} \right]^{d_t} \prod_{t=1}^{T} \left[ f(s_t | y_T, \vartheta) 1 - F(s_t^* | y_t | \Theta) \right]^{d_t} \prod_{t=1}^{T} \left[ 1 - F(s_t^* | y_t | \Theta) \right]^{d_t} \prod_{t=1}^{T} \left[ l(y_t | \vartheta) \right]^{d_t}, \tag{37}
\]

where \( g, f \) and \( l \) represent the respective density functions.

MCMC is performed in two steps. First, I draw each individual firm-level parameter from its corresponding posterior density conditional on other individual firm-level parameters, the population parameters, and the data. Second, I draw the population parameters conditional on the individual parameters.

(a) Individual parameters \( \vartheta_1 \equiv (\lambda_0, p_0, c) \) conditional on other parameters and data.

Because only \( \kappa_i \) and \( \tau_i \) enter \( f(s_t | y_T) \) and only \( \rho_{i0} \) and \( \rho_{i1} \) enter \( f(y_t | \Theta) \), the posterior
The individual parameter $\tau$ does not involve $f(s_t|y_T)$:

$$
\pi(\theta|y_t, s_t * d_t, d_t, P_t, \theta^-) \propto \prod_{t=1}^{T} \left[ g(P(s_t|y_T, d_t, \vartheta))^{d_t} g(P(s_t \leq s_t^*(\vartheta, y_t)|y_T, d_t, \vartheta))^{1-d_t} \right] ^{\prod_{t=1}^{T} \left[ F(s_t^*(\vartheta, y_t|y_T, \vartheta)) \right]^{1-d_t}} f(\theta|\theta), \quad (38)
$$

where $\theta \in \vartheta_1$.

(b) The individual parameter $\rho_1$ conditional on other parameters and data.

Individual parameter $\rho_{1t}$ does not enter $f(s_t|y_T)$. The posterior density is

$$
\pi(\rho_1|y_t, s_t * d_t, d_t, P_t, \rho_1^-) \propto \prod_{t=1}^{T} \left[ g(P(s_t|y_T, d_t, \vartheta))^{d_t} g(P(s_t \leq s_t^*(\vartheta, y_t)|y_T, d_t, \vartheta))^{1-d_t} \right] ^{\prod_{t=1}^{T} \left[ F(s_t^*(\vartheta, y_t|y_T, \vartheta)) \right]^{1-d_t}} \prod_{t=1}^{T} f(y_t) f(\rho_1|\theta). \quad (39)
$$

(c) The individual parameter $\kappa$ conditional on other parameters and data.

The individual parameter $\kappa$ does not enter $f(y_t|\vartheta)$. The posterior density is

$$
\pi(\kappa|y_t, s_t * d_t, d_t, P_t, \kappa^-) \propto \prod_{t=1}^{T} \left[ g(P(s_t|y_T, d_t, \vartheta))^{d_t} g(P(s_t \leq s_t^*(\vartheta, y_t)|y_T, d_t, \vartheta))^{1-d_t} \right] ^{\prod_{t=1}^{T} \left[ F(s_t^*(\vartheta, y_t|y_T, \vartheta)) \right]^{1-d_t}} \prod_{t=1}^{T} f(s_t|y_T, \vartheta) \prod_{t=1}^{T} \left( f(y_t|\vartheta) f(\kappa|\theta) \right). \quad (40)
$$

(d) The individual parameter $\tau$ conditional on other parameters and data.

Because $\tau$ enters every component, its posterior density is

$$
\pi(\tau|y_t, s_t * d_t, d_t, P_t, \tau^-) \propto \prod_{t=1}^{T} \left[ g(P(s_t|y_T, d_t, \vartheta))^{d_t} g(P(s_t \leq s_t^*(\vartheta, y_t)|y_T, d_t, \vartheta))^{1-d_t} \right] ^{\prod_{t=1}^{T} \left[ F(s_t^*(\vartheta, y_t|y_T, \vartheta)) \right]^{1-d_t}} ^{\prod_{t=1}^{T} \left[ 1 - F(s_t^*(\vartheta, y_t|y_T, \vartheta)) \right]^{d_t}} \prod_{t=1}^{T} f(y_t) f(\tau|\theta). \quad (41)
$$

(e) The individual parameter $\beta$ conditional on other parameters and data.
Because $\beta$ only affects stock price, its posterior depends only on the part of stock price,

$$
\pi(\beta|y_t, s_t, d_t, P_t, \tau_i^-) \propto \prod_{t=1}^{T} \left[ g\left(P(s_t|y_T, d_t, \vartheta)\right)^{d_t} g\left(P(s_t \leq s_t^*(\vartheta, y_t)|y_T, d_t, \vartheta)\right)^{1-d_t} \right] f(\beta|\theta). 
$$

(42)

(f) The individual parameter $\rho_0$ conditional on other parameters and data.

The individual parameter $\rho_0$ only enters $f(y_t)$. Its posterior density is

$$
\pi(\rho_0|y_t, s_t, d_t, P_t, \rho_0^-) \propto \prod_{t=1}^{T} \left(f(y_t|\vartheta)) f(\rho_0|\theta)\right).
$$

(43)

(g) Population parameters conditional on individual parameters.

I first convert the individual firm-level parameters to the real line. Then the population parameters can be computed iteratively by

- $\theta|\vartheta, w$, which has a normal distribution;
- $w|\vartheta, \theta$, which has an inverse gamma distribution.
Figure 1: Comparison between CIG and I/B/E/S Guidance

This figure compares I/B/E/S Guidance data with First Call’s Company Issued Guidance (CIG) data. The upper left panel is the percentage of EPS forecasts out of all types of forecasts. The upper right panel is the total number of forecasts by thousands. In these two panels, each forecast is treated as a distinct observation. For example, if a manager issues an EPS forecast along with a revenue forecast on the same day, the total number of forecasts is two instead of one. The lower left panel is the total number of firms. The lower right panel is the percentage of firms that have had I/B/E/S analyst coverage. The horizontal line represents year. The dashed line represents I/B/E/S Guidance data. The solid line represents CIG data.
Figure 2: Cross-sectional distribution of management forecast frequencies

This figure presents the cross-sectional density of management forecast frequency for firms that do not always disclose over the sample period from 2003 to 2014. Disclosure frequency is computed as the percentage of years with forecast issuance. The horizontal axis represents the management forecast frequency. The vertical axis is the density.
This figure tracks firms’ future forecast issuances for up to six years conditional on their current forecast decisions. The upper dashed line presents firms that issue management forecasts and the lower solid line presents firms that do not. The vertical axis is the percentage of firms that issue forecasts each year after the current forecast decision, demeaned by the yearly trend in the management forecast issuance. The horizontal axis represents each subsequent year.
Figure 4: ROA over time

This figure tracks firms’ future ROA (return on assets) based on their current forecast decisions. The upper dashed line presents firms that issue management forecasts, and the lower solid line presents firms that do not. The vertical axis is the median ROA for each year after the current forecast decision. The horizontal axis represents each subsequent year.
This figure plots the cross-sectional distribution of model fit of disclosure decisions. For each firm, I compute the predictive probabilities of (non)disclosure decisions. A higher predictive probability implies a better model fit. The horizontal axis is the model fit, and the vertical axis is the density. The vertical dotted line represents the model fit of a naive model of disclosure decisions with past earnings as the only explanatory variable.
This figure plots the marginal effects of model parameters (see section 2.2.4 for the definition of the parameters). I simulate 2,000 firms and 12 periods each firm with parameter values drawn from their cross-sectional distributions. I vary each parameter by one standard deviation of its cross-sectional distribution and compute the average change in disclosure frequency over the following 12 years. The red solid (blue dashed) line corresponds to the 5th percentile and 95th percentile of the change in disclosure frequency when I increase (decrease) the parameter at the horizontal axis by one standard deviation of its cross-sectional distribution. I also vary investors’ prior beliefs by one standard deviation of earnings shock with the results denoted by $q_0^e$. Disclosure frequency is expressed in terms of the number of standard deviations of the empirical disclosure frequency, which is about 22%. The black dots represent the mean.
The table presents types of forecasts I/B/E/S Guidance covers from year 1997. I classify forecasts according to Table 1 of Chuk et al. (2013).

<table>
<thead>
<tr>
<th>Types of Forecasts</th>
<th>1997</th>
<th>1999</th>
<th>2001</th>
<th>2003</th>
<th>2005</th>
<th>2007</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>EBITDA only</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.20</td>
</tr>
<tr>
<td>EPS + EBITDA</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.13</td>
</tr>
<tr>
<td>EPS only</td>
<td>97.89</td>
<td>743</td>
<td>98.50</td>
<td>1711</td>
<td>96.88</td>
<td>5210</td>
<td>35.09</td>
</tr>
<tr>
<td>FFO only</td>
<td>2.11</td>
<td>16</td>
<td>1.50</td>
<td>26</td>
<td>2.34</td>
<td>126</td>
<td>2.19</td>
</tr>
<tr>
<td>Other</td>
<td>0.00</td>
<td>0.00</td>
<td>0.78</td>
<td>42</td>
<td>44.71</td>
<td>3082</td>
<td>73.53</td>
</tr>
<tr>
<td>Total</td>
<td>100.00</td>
<td>759</td>
<td>100.00</td>
<td>1737</td>
<td>100.00</td>
<td>5378</td>
<td>100.00</td>
</tr>
</tbody>
</table>
Table 2: Sample selection

This table describes the sample selection procedure for management forecasts.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBES Guidance 1992-2014</td>
<td>363,754</td>
</tr>
<tr>
<td>Non-US firms, firms not matched to Compustat, firms with no I/B/E/S</td>
<td>(71,160)</td>
</tr>
<tr>
<td>ticks and forecasts issued before 2002</td>
<td>292,594</td>
</tr>
<tr>
<td>Drop firms that issue zero or one forecast</td>
<td>(12,059)</td>
</tr>
<tr>
<td>and firms that issue FFO (funds from operations) forecasts(^a)</td>
<td>280,535</td>
</tr>
<tr>
<td>Drop forecasts issued after the corresponding fiscal year-end</td>
<td>(16,698)</td>
</tr>
<tr>
<td>Keep firms that issue EPS forecasts each year of forecast issuance;</td>
<td>(188,838)</td>
</tr>
<tr>
<td>Drop non-EPS forecasts</td>
<td>74,999</td>
</tr>
<tr>
<td>Keep firms that issue annual forecasts each year of forecast issuance;</td>
<td>(35,858)</td>
</tr>
<tr>
<td>Drop quarterly and semi-annual forecasts</td>
<td>39,141</td>
</tr>
<tr>
<td>Keep firms with more than five consecutive years of actual EPS from</td>
<td>(2,558)</td>
</tr>
<tr>
<td>I/B/E/S</td>
<td>36,583</td>
</tr>
<tr>
<td>Drop firms that do not issue forecast for the subsequent fiscal period;</td>
<td>(4,693)</td>
</tr>
<tr>
<td>Drop duplicate forecasts for each firm-year(^b)</td>
<td>(24,439)</td>
</tr>
<tr>
<td>Drop firms that issue forecasts only in one year</td>
<td>(212)</td>
</tr>
<tr>
<td>Drop firms with fewer than five consecutive years of nonmissing</td>
<td>(1,128)</td>
</tr>
<tr>
<td>CRSP price and earnings forecast decision data.</td>
<td>6,111</td>
</tr>
</tbody>
</table>

\(^a\)These firms are mostly REITS. The relation between FFO and price is likely to be different from that between EPS and price.

\(^b\)I keep the earliest forecast for each firm-year.
Table 3: Summary Statistics of Structural Sample

This table presents the summary statistics. The sample covers from 2003 until 2014 and retains only firms that issue EPS forecasts (see Table 2). Structural estimation utilizes four variables: earnings $y_{it}$, stock prices $p_{it}$, disclosure decisions $d_{it}$, and management forecasts $s_{it}$. I scale all variables of each firm by the stock price at the beginning of the sample period and then multiply them by 100.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>5%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>95%</th>
<th>Obs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earnings $y_{it}$</td>
<td>12.262</td>
<td>17.754</td>
<td>0.070</td>
<td>5.938</td>
<td>9.141</td>
<td>15.356</td>
<td>35.373</td>
<td>6,111</td>
</tr>
<tr>
<td>Price $p_{it}$</td>
<td>214.091</td>
<td>272.694</td>
<td>56.625</td>
<td>100.000</td>
<td>145.613</td>
<td>232.368</td>
<td>570.153</td>
<td>6,111</td>
</tr>
<tr>
<td>Disclosure decision $d_{it}$</td>
<td>0.873</td>
<td>0.333</td>
<td>0.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>6,111</td>
</tr>
<tr>
<td>Management forecast $s_{it}$</td>
<td>13.329</td>
<td>15.928</td>
<td>2.942</td>
<td>6.513</td>
<td>9.528</td>
<td>15.561</td>
<td>36.175</td>
<td>6,111</td>
</tr>
</tbody>
</table>
Table 4: The distribution of parameter estimates

This table presents the cross-sectional distribution of parameter estimates. Each firm has eight parameters: investors’ prior belief $q_{0i}$, investors’ belief precision $\lambda_i$, cost of disclosure $c_i$, expected return $r_i$, manager’s information precision $\kappa_i$, the intercept of earnings process $\rho_{0i}$, earnings persistence $\rho_{1i}$, and variance of earnings process $1/\tau_i$, where $i$ refers to the $i$’s firm. With the exception of $\lambda_i$, $\kappa_i$, $r_i$, and $\rho_{1i}$, which are scale invariant, all parameters are scaled by the initial stock price of firm $i$ and multiplied by 100. For each firm, structural estimation recovers the distribution of these parameters. I obtain their medians and present the cross-sectional distribution (across firms) of these medians.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>SD</th>
<th>5%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investors’ prior belief $q_{0i}$</td>
<td>12.032</td>
<td>28.150</td>
<td>0.959</td>
<td>2.226</td>
<td>4.207</td>
<td>9.160</td>
<td>43.725</td>
</tr>
<tr>
<td>Investors’ belief precision $\lambda_i$</td>
<td>5.572</td>
<td>3.057</td>
<td>1.059</td>
<td>3.990</td>
<td>5.183</td>
<td>6.674</td>
<td>10.612</td>
</tr>
<tr>
<td>Expected return $r_i = \frac{1-\beta_i}{\beta_i}$</td>
<td>0.111</td>
<td>0.090</td>
<td>0.040</td>
<td>0.061</td>
<td>0.080</td>
<td>0.129</td>
<td>0.287</td>
</tr>
<tr>
<td>Manager’s information precision $\kappa_i$</td>
<td>4.032</td>
<td>5.333</td>
<td>0.300</td>
<td>1.867</td>
<td>2.918</td>
<td>4.584</td>
<td>10.970</td>
</tr>
<tr>
<td>Intercept of earnings process $\rho_{0i}$</td>
<td>6.637</td>
<td>3.472</td>
<td>3.441</td>
<td>4.546</td>
<td>5.855</td>
<td>7.425</td>
<td>12.430</td>
</tr>
<tr>
<td>Earnings persistence $\rho_{1i}$</td>
<td>0.630</td>
<td>0.171</td>
<td>0.339</td>
<td>0.516</td>
<td>0.640</td>
<td>0.771</td>
<td>0.883</td>
</tr>
<tr>
<td>Variance of earnings process $1/\tau_i$</td>
<td>40.359</td>
<td>84.709</td>
<td>0.797</td>
<td>2.776</td>
<td>7.766</td>
<td>24.001</td>
<td>273.437</td>
</tr>
</tbody>
</table>
Table 5: Model fit

This table presents the model fit of disclosure decisions (the first row) and stock prices (the second row). I measure the model fit of disclosure decisions as follows. For each observation, I use the model to generate predictive disclosure probability or nondisclosure probability depending on the actual disclosure decision. A higher probability implies a better model fit. I then estimate a linear regression model using past earnings as the only explanatory variable and generate predictive (non-)disclosure probabilities in the same way. Finally, I take the ratio of the two probabilities and present the summary statistics in the first row. I measure the model fit of stock prices by comparing the standard deviation of model residual to the standard deviation of stock prices. This is similar to the $R^2$. The difference is that the standard deviation of the model residual can be larger than the standard deviation of actual stock prices. The measure essentially compares the structural model to a model that uses average stock price. A smaller value implies a better model fit.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>5%</th>
<th>25%</th>
<th>75%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disclosure</td>
<td>1.060</td>
<td>1.143</td>
<td>0.200</td>
<td>0.649</td>
<td>0.949</td>
<td>1.222</td>
<td>1.236</td>
</tr>
<tr>
<td>Stock Price</td>
<td>0.704</td>
<td>0.687</td>
<td>0.327</td>
<td>0.284</td>
<td>0.483</td>
<td>0.864</td>
<td>1.166</td>
</tr>
</tbody>
</table>
Table 6: Marginal effects on future disclosure incentives

This table presents the marginal effects of the following structural parameters: investors’ prior belief \( q_0 \), investors’ belief precision \( \lambda_i \), manager’s information precision \( \kappa_i \), disclosure cost \( c_i \), discount factor \( \beta_i \), firm profitability \( \rho_{0i} \), and earnings persistence \( \rho_{1i} \). I generate 2,000 firms using the cross-sectional distributions of the parameter estimates. For each firm, I simulate 1,000 samples each with 12 periods and compute the average disclosure frequency over these samples. I report in the first row the cross-sectional distribution of the average disclosure frequency over the 12 periods. To evaluate the marginal effects, I increase (labeled as “+”) or decrease (labeled as “−”) each parameter by one standard deviation (computed from its cross-sectional distribution) while holding other parameters constant, and compute the average disclosure frequency following the same simulation procedure and report the changes in the disclosure frequency. I also vary investors’ prior beliefs by one standard deviation of firm earnings shock, \( q_0^e \). The results are reported in the second row and onwards.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>5%</th>
<th>25%</th>
<th>75%</th>
<th>95%</th>
<th>Zero(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>0.736</td>
<td>0.986</td>
<td>0.377</td>
<td>0.000</td>
<td>0.483</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>( q_0^+ )</td>
<td>−0.258</td>
<td>−0.121</td>
<td>0.305</td>
<td>−0.916</td>
<td>−0.429</td>
<td>−0.011</td>
<td>0.000</td>
<td>34.70%</td>
</tr>
<tr>
<td>( q_0^- )</td>
<td>0.154</td>
<td>0.041</td>
<td>0.238</td>
<td>0.000</td>
<td>0.005</td>
<td>0.188</td>
<td>0.760</td>
<td>37.80%</td>
</tr>
<tr>
<td>( q_0^e^+ )</td>
<td>−0.053</td>
<td>−0.031</td>
<td>0.059</td>
<td>−0.172</td>
<td>−0.087</td>
<td>−0.004</td>
<td>0.000</td>
<td>40.35%</td>
</tr>
<tr>
<td>( q_0^e^- )</td>
<td>0.049</td>
<td>0.026</td>
<td>0.056</td>
<td>0.000</td>
<td>0.004</td>
<td>0.079</td>
<td>0.174</td>
<td>41.70%</td>
</tr>
<tr>
<td>( \lambda^+ )</td>
<td>−0.077</td>
<td>−0.022</td>
<td>0.142</td>
<td>−0.347</td>
<td>−0.111</td>
<td>−0.001</td>
<td>0.034</td>
<td>37.50%</td>
</tr>
<tr>
<td>( \lambda^- )</td>
<td>0.063</td>
<td>0.014</td>
<td>0.137</td>
<td>−0.065</td>
<td>0.001</td>
<td>0.085</td>
<td>0.343</td>
<td>39.00%</td>
</tr>
<tr>
<td>( \kappa^+ )</td>
<td>0.039</td>
<td>0.015</td>
<td>0.058</td>
<td>0.000</td>
<td>0.002</td>
<td>0.052</td>
<td>0.160</td>
<td>43.55%</td>
</tr>
<tr>
<td>( \kappa^- )</td>
<td>−0.082</td>
<td>−0.038</td>
<td>0.103</td>
<td>−0.296</td>
<td>−0.129</td>
<td>−0.004</td>
<td>0.000</td>
<td>40.65%</td>
</tr>
<tr>
<td>( c^+ )</td>
<td>−0.237</td>
<td>−0.195</td>
<td>0.213</td>
<td>−0.614</td>
<td>−0.400</td>
<td>−0.034</td>
<td>0.000</td>
<td>25.30%</td>
</tr>
<tr>
<td>( c^- )</td>
<td>0.194</td>
<td>0.098</td>
<td>0.220</td>
<td>0.000</td>
<td>0.007</td>
<td>0.357</td>
<td>0.598</td>
<td>37.65%</td>
</tr>
<tr>
<td>( \beta^+ )</td>
<td>0.123</td>
<td>0.067</td>
<td>0.144</td>
<td>0.000</td>
<td>0.006</td>
<td>0.211</td>
<td>0.377</td>
<td>38.65%</td>
</tr>
<tr>
<td>( \beta^- )</td>
<td>−0.123</td>
<td>−0.084</td>
<td>0.132</td>
<td>−0.367</td>
<td>−0.202</td>
<td>−0.010</td>
<td>0.000</td>
<td>30.90%</td>
</tr>
<tr>
<td>( \rho_{1i}^+ )</td>
<td>0.129</td>
<td>0.071</td>
<td>0.151</td>
<td>0.000</td>
<td>0.006</td>
<td>0.211</td>
<td>0.417</td>
<td>38.70%</td>
</tr>
<tr>
<td>( \rho_{1i}^- )</td>
<td>−0.123</td>
<td>−0.078</td>
<td>0.138</td>
<td>−0.395</td>
<td>−0.191</td>
<td>−0.010</td>
<td>0.000</td>
<td>29.85%</td>
</tr>
<tr>
<td>( \rho_{0i}^+ )</td>
<td>0.096</td>
<td>0.038</td>
<td>0.137</td>
<td>0.000</td>
<td>0.005</td>
<td>0.134</td>
<td>0.372</td>
<td>40.25%</td>
</tr>
<tr>
<td>( \rho_{0i}^- )</td>
<td>−0.072</td>
<td>−0.029</td>
<td>0.107</td>
<td>−0.285</td>
<td>−0.097</td>
<td>−0.004</td>
<td>0.000</td>
<td>41.00%</td>
</tr>
</tbody>
</table>
Table 7: Investor learning and persistent disclosure incentives

This table presents the intertemporal correlation of disclosure probabilities implied by the structural model. I generate 1,000 firms using the distributions of parameter estimates. For each firm, I simulate 500 samples each with seven periods of data, a total of 3,500 observations. In the simulation, I ensure that all the samples for a given firm start with the same investor prior belief. Changes in investors’ beliefs due to random performance shocks induce variations in the disclosure probabilities of period 2 and onwards. I compute the probability of disclosure as in (14) for each period. To measure persistence in disclosure incentives, for each firm, I regress the disclosure probabilities of periods 3-7 each on the disclosure probability of period 2. The slope of the regression coefficients captures how much change in the disclosure probability in period 2, which is induced by investor learning, affects disclosure probabilities in periods 3-7. I report the cross-sectional distribution of the regression slope coefficients. For example, \( \beta_1 \) represents the slope coefficient from regressing disclosure probabilities in period 3 on disclosure probabilities in period 2; \( \beta_2 \) represents the slope coefficient from regressing disclosure probabilities in period 4 on disclosure probabilities in period 2.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>5%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_1 )</td>
<td>1.036</td>
<td>0.461</td>
<td>0.549</td>
<td>0.825</td>
<td>0.940</td>
<td>1.117</td>
<td>1.909</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>0.986</td>
<td>0.619</td>
<td>0.292</td>
<td>0.677</td>
<td>0.849</td>
<td>1.117</td>
<td>2.206</td>
</tr>
<tr>
<td>( \beta_3 )</td>
<td>0.908</td>
<td>0.659</td>
<td>0.163</td>
<td>0.545</td>
<td>0.762</td>
<td>1.065</td>
<td>2.337</td>
</tr>
<tr>
<td>( \beta_4 )</td>
<td>0.823</td>
<td>0.642</td>
<td>0.083</td>
<td>0.443</td>
<td>0.673</td>
<td>0.971</td>
<td>2.176</td>
</tr>
<tr>
<td>( \beta_5 )</td>
<td>0.745</td>
<td>0.614</td>
<td>0.042</td>
<td>0.367</td>
<td>0.589</td>
<td>0.924</td>
<td>2.041</td>
</tr>
</tbody>
</table>