

THE UNIVERSITY OF CHICAGO

DISCLOSURE DYNAMICS AND INVESTOR LEARNING

A DISSERTATION SUBMITTED TO  
THE FACULTY OF THE UNIVERSITY OF CHICAGO  
BOOTH SCHOOL OF BUSINESS  
IN CANDIDACY FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

BY

SHUANG ZHOU

CHICAGO, ILLINOIS

MARCH 2017

Copyright © 2017 by Shuang Zhou

All Rights Reserved

I am indebted to my dissertation committee at the University of Chicago, Douglas Skinner (chair), Philip Berger, Richard Hahn, and Haresh Sapra, for their help and suggestions.

# TABLE OF CONTENTS

LIST OF FIGURES . . . . .	vi
LIST OF TABLES . . . . .	vii
ACKNOWLEDGMENTS . . . . .	viii
ABSTRACT . . . . .	ix
<b>1 DISCLOSURE DYNAMICS AND INVESTOR LEARNING . . . . .</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Model . . . . .	7
1.2.1 Investor learning about firm profitability . . . . .	7
1.2.2 Model set up . . . . .	8
1.2.3 Equilibrium strategy . . . . .	11
1.3 Research design . . . . .	13
1.3.1 Empirical quantifications . . . . .	13
1.3.2 Empirical methods . . . . .	14
1.3.3 Estimation procedure . . . . .	15
1.3.4 Identification . . . . .	17
1.4 Data and sample selection . . . . .	18
1.5 Descriptive analysis . . . . .	21
1.5.1 Summary statistics of the variables used for estimation . . . . .	21
1.5.2 How investors respond to (non)disclosures . . . . .	24
1.5.3 Dynamics of management forecast issuance . . . . .	25
1.6 Results . . . . .	29
1.6.1 Parameter estimates . . . . .	29
1.6.2 Model fit . . . . .	31
1.6.3 Investor learning and persistent disclosure incentives . . . . .	36
1.6.4 Investor learning and the average forecast frequency . . . . .	39
1.6.5 Robustness . . . . .	44
1.7 Relation to prior literature . . . . .	46
1.7.1 Reputation accumulation, investor learning, and management forecasts . . . . .	46
1.7.2 Investor learning and analyst earnings forecasts . . . . .	48
1.8 Conclusion . . . . .	51
REFERENCES . . . . .	53
<b>A MODEL . . . . .</b>	<b>59</b>
A.1 Discussions of model assumptions . . . . .	59
A.2 The effects of investor learning . . . . .	61
A.3 The effects of other model parameters . . . . .	63
A.4 Proofs . . . . .	64
A.5 Bayesian updating . . . . .	65

A.5.1	Learning $\rho$ . . . . .	66
A.5.2	Prior predictive distribution of $y_t$ before observing $s_t$ . . . . .	67
A.5.3	Posterior predictive of $y_t$ given $s_t$ . . . . .	68
A.6	Derivation of stock price exclusive of disclosure cost . . . . .	68
A.7	Proof of Proposition 1 . . . . .	69
A.8	The equilibrium condition . . . . .	72
B	CIG AND I/B/E/S GUIDANCE . . . . .	75
C	ESTIMATION PROCEDURE . . . . .	76

## LIST OF FIGURES

1.1	Comparison between CIG and I/B/E/S Guidance . . . . .	20
1.2	Firm performance and management forecast decisions . . . . .	24
1.3	Investors' reactions to management forecast decisions . . . . .	26
1.4	Cross-sectional distribution of management forecast frequencies . . . . .	27
1.5	Management forecast frequency over time . . . . .	28
1.6	Marginal effects of structural parameters . . . . .	42

## LIST OF TABLES

1.1	Types of forecasts . . . . .	19
1.2	Sample selection . . . . .	22
1.3	Summary Statistics of Structural Sample . . . . .	23
1.4	The distribution of parameter estimates . . . . .	29
1.5	Model fit . . . . .	34
1.6	Investor learning and persistent disclosure incentives . . . . .	38
1.7	Marginal effects on future disclosure incentives . . . . .	41
1.8	Investors' beliefs and analyst forecasts . . . . .	50

## ACKNOWLEDGMENTS

I appreciate the comments of Ray Ball, John Barrios, Jeremy Bertomeu, Matt Bloomfield, Jung-Ho Choi, Peter Easton, Pingyang Gao, Rachel Geoffroy, Joseph Gerakos, Brandon Gipper, Ilan Guttman, Anya Kleymenova, Hee-Min Lee, Christian Leuz, Brett Lombardi, Mike Minnis, Valeri Nikolaev, Sorabh Tomar, Katie Wang, Laurence van Lent, Regina Wittenberg-Moerman, Anastasia Zakolyukina, Sarah Zechman, Yuqing Zhou, and seminar participants at the University of Chicago accounting workshop and IO lunch, University of Michigan, Duke University, University of Texas at Dallas, MIT, University of Pennsylvania, Stanford GSB, University of Southern California, London Business School, and University of Toronto. All errors are my own.



## ABSTRACT

This paper examines the effects of investors' learning of unknown underlying firm profitability on the dynamics of managers' voluntary disclosure decisions. I structurally estimate a voluntary disclosure model in which investors learn profitability over multiple years from management earnings forecasts and realized earnings. The results suggest that investor learning induces persistent disclosure incentives. Other things equal, a 10% change in the likelihood of disclosure caused by shocks to investors' beliefs about profitability is associated with a 10% change in the likelihood of disclosure in the next year. In addition, disclosure is more likely when investors' beliefs about profitability are more pessimistic relative to the true profitability or investors are more uncertain about their beliefs. This paper enriches the understanding of the role of investor learning of unknown parameters in driving disclosure decisions. In doing so, it also provides an economic explanation for the observed "stickiness" of managers' disclosure decisions.

# CHAPTER 1

## DISCLOSURE DYNAMICS AND INVESTOR LEARNING

### 1.1 Introduction

Prior studies find that managers' voluntary disclosure decisions are "sticky," that is, a past disclosure is associated with the likelihood of a current disclosure [e.g., 35, 78]. This paper proposes one mechanism, investor learning about unknown underlying firm profitability, that could help explain this phenomenon, and aims to quantify the empirical effects of the mechanism on management's decisions to issue annual earnings forecasts.

The paper starts by incorporating investors' uncertainty about the mean of the earnings process into the disclosure model by Verrecchia [81]. A manager decides whether to withhold or truthfully disclose her private, noisy information about unrealized forthcoming earnings.<sup>1</sup> A disclosure incurs an exogenous fixed cost. The disclosure decision depends on its effect on the current stock price, the sum of discounted future expected earnings net of disclosure costs.<sup>2</sup>

Following Lang [61], I assume that firm earnings follow an AR1 process with a normal random shock and that investors do not know underlying firm profitability, the intercept of firm earnings process. After observing a disclosure, investors' updated beliefs about profitability are a weighted average of their prior beliefs and the disclosure. I refer to the Bayesian updating process as "investor learning" (about the unknown profitability) following seminal work on parameter uncertainty [e.g., 86, 70]. A consequence of investor learning

---

1. Verrecchia [81] examines the disclosure of a noisy signal of the liquidation value. The mean of the earnings distribution of my model can be thought of as the liquidation value of Verrecchia [81] and the management forecasts being the noisy signal. While not theoretically different from a one-period perspective, my model has a natural multi-period flavor in that investors' beliefs about the mean of the earnings process change as they obtain more data over time, which facilitates empirical estimation.

2. Using stock price as the objective function is consistent with many disclosure models [e.g., 81, 30]. The model also produces similar empirical predictions to frameworks that assume explicit management incentives to align expectations or reduce information asymmetry [e.g., 5, 27, 34]. The model is also robust to assuming managers' putting weight on future stock prices as long as they do not ignore the current stock price. The model uses expected earnings instead of dividend for simplicity.

focused on by this paper is that investors' beliefs are serially correlated.

I show that without any commitment device, learning firm profitability leads to serially correlated disclosure probabilities. The mechanism is as follows. The model features a threshold equilibrium, that is, the manager discloses if and only if her private signal exceeds a threshold. The likelihood of disclosure of a given period (year) is then the probability that manager's private information about forthcoming earnings will be larger than the disclosure threshold. Importantly, from the perspective of an empiricist trying to understand disclosure decisions, this probability is based on the true distribution of the manager's private information about forthcoming earnings, not investors' perceived distribution, because the true distribution determines what the manager observes.<sup>3</sup> The role of investors' perceived distribution is that it determines investors' reaction to a (non)disclosure, which affects the disclosure threshold. This perspective means that the paper does not focus on what investors thought of the likelihood of disclosure given investors' information set prior to the disclosure decision. Instead, it aims to explain disclosure decisions in an ex-post sense based on what researchers observe from the data. New information arrival (a draw from the true distribution) does not change the true distribution of the manager's private information but can change investors' beliefs about profitability, which move the disclosure thresholds and consequently disclosure probability. Finally, because investors' beliefs are serially correlated, the likelihood of disclosure is serially correlated.

The main objective of the paper is to quantify the effects of investor learning about unknown firm profitability on the issuance of management annual earnings forecasts. I require firms to have at least five consecutive years of data on actual earnings, management annual earnings forecast decisions, and stock prices between 2003 and 2014. I examine two aspects: the average disclosure frequency and the persistence of disclosure incentives. The average disclosure frequency is the frequency of forecast issuance over the following 12 years, and

---

3. I estimate the true distribution of earnings using earnings data from 2003 to 2014 for each firm. I simulate fake firms using the cross-sectional distribution of firm profitability, which is arguably more stable than the profitability of individual firms.

the persistence of disclosure is the slope coefficient from regressing future disclosure probabilities on current disclosure probabilities, holding constant firm characteristics.<sup>4</sup> Higher persistence means a longer time, following a temporal shock, for the likelihood of disclosure to adjust back to the long run average. The primary focus is on disclosure probabilities, because investor learning directly affects the ex-ante incentive to disclose by changing the disclosure threshold.

The empirical estimation utilizes four outcome variables the model generates: stock prices, actual earnings, disclosure decisions, and the contents of management forecasts. From the four outcome variables, I estimate the cross-sectional distributions of eight model primitives: investors' prior beliefs, investor uncertainty, firm profitability, earnings persistence, the volatility of earnings shocks, the cost of capital, manager's information precision, and the fixed cost of disclosure. These model primitives can be thought of as independent variables. The estimation is structural, because I strictly impose the model structure. Structural estimation is chosen because of the following benefits. It is internally consistent in that it respects the threshold equilibrium and nonlinearities among the model parameters. It estimates investors' beliefs and avoids the issue that empirical measures of investors' beliefs might be correlated with firm fundamentals, which facilitates disentangling the effects of investor learning. It allows isolating the mechanism of investor learning holding other factors constant through simulation.

Using predicted forecast decisions based on the parameter estimates from the model, I demonstrate that the model adequately describes the dynamics of management forecast decisions that are present in the data. Using the predicted forecast decisions, a pooled regression of current disclosure decision on prior year disclosure decision produces a slope

---

4. I obtain frequency of forecast issuance using the simulation described in section 1.6, 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. I generate disclosure probabilities from the model. Measuring persistence requires variation in disclosure probabilities. 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.

coefficient of 0.63, which economically and statistically resembles the slope coefficient of 0.64 using the same regression and the actual data. Because the model identification does not require the model to match the intertemporal correlation of disclosure decisions, a good model fit increases the confidence in the model and the following analysis through model simulation.

For the persistence of disclosure incentives, I examine how changes in the likelihood of disclosure caused by shocks to investors' beliefs propagate into the future. Via simulation, I shock investors' beliefs by drawing a series of shocks to earnings, which in turn change the likelihood of disclosure. I find that a 10% increase in disclosure probability induced by changes in investors' prior beliefs caused by shocks to earnings on average increases the disclosure probability for the next year by around 10.3%. The effect then declines monotonically to about a 7.5% increase in the disclosure probability in five years. The results suggests investor learning is an important channel explaining intertemporally correlated disclosure incentives.<sup>5</sup>

For the average forecast frequency, I find investor learning (of the unknown profitability) has extremely heterogeneous effects on the average forecast frequency. For around 40% of firms, marginal changes in investors' beliefs or uncertainty, the two learning parameters, do not change disclosure frequencies. I obtain similar results for other firm characteristics such as firm profitability and firm risk.<sup>6</sup> Further analysis reveals that these firms always disclose, which explains the negligible marginal effects. The finding is consistent with 57% of firms in the sample issuing annual earnings forecasts every year.

For the remaining 60% of firms, the effects on the average future disclosure frequency

---

5. If investors know the true earnings distribution (that is, no learning), shocks to earnings will not change their beliefs about firm profitability (the mean of earnings distribution). Because the probability of disclosure depends on the difference between investors' beliefs about the mean of earnings and the actual mean of earnings, no learning implies that the two are the same and therefore their difference is zero. Holding other firm characteristics constant, the disclosure probability is a constant that depends on other firm characteristics.

6. I change each parameter by one standard deviation of its cross-sectional distribution recovered from the structural estimation. I change investors' beliefs by one standard deviation of the earnings shock of each firm to reflect the within-firm nature of belief changes.

are much larger. A one standard deviation increase in investor uncertainty (decrease in investors' prior beliefs about profitability) on average increases the future forecast frequency by about 0.32 (0.23) standard deviations of the empirical disclosure frequency. The effects of learning are economically comparable with those of other firm characteristics. For example, the effect of a one standard deviation change in firm profitability is about 0.44 standard deviations, and the effect of a one standard deviation change in earnings persistence is about 0.55 standard deviations.

This paper makes several contributions. First, the paper suggests that a factor potentially explaining the “stickiness” of managers' earnings forecasts is investor learning about unknown firm profitability. The paper both models investor learning about unknown firm profitability as a mechanism through which persistent disclosure incentives occur, and empirically shows that the mechanism is economically important. The paper is related to studies on the dynamics of voluntary disclosure. Prior empirical studies document sticky voluntary disclosure incentives [e.g., 35, 78]. I extend these studies by providing a mechanism. Analytical studies explain disclosure dynamics using managers' reputation concern for being transparent [17], reputation for being uninformed [32], the timing of voluntary disclosures [3, 48], and the likelihood of a negative economic shock [64]. I complement these studies by showing that investors' information processing affects management's voluntary disclosure incentives. My model can be viewed a multi-period replication of Verrecchia [81] that allows for the impact of current disclosure decision on future disclosure decisions.

Second, the paper contributes to the literature on how investors affect managers' disclosure decisions. For example, prior studies show that fundamental uncertainty such as earnings volatility reduces disclosure [e.g., 84], whereas Billings et al. [19] shows that a run-up in volatility increases the likelihood of a bundled forecast. I show that investor uncertainty increases disclosure, consistent with the short window results of Billings et al. [19]. In my model, the result of Waymire [84] on fundamental uncertainty represents a source of disclosure cost, which reduces disclosure. Moreover, I find investor optimism reduces disclosure.

The result is consistent with the finding of Bergman and Roychowdhury [13] that higher investor sentiment reduces the likelihood of disclosure, but appears to be at odds with the finding of Cotter et al. [27] that analyst optimism increases disclosure. The distinction suggests that forces driving analyst optimism are distinct from those driving investor learning. Specifically, investors' beliefs about firm profitability differ from the actual firm profitability, because investors need to form inference based on limited historical information (when they are yet to observe future earnings) but are fully rational. In contrast, analysts' affiliation with brokerage firms could affect their optimism [e.g., 63]. The results highlight the need to better understand the types of information users (analysts versus investors) when examining managers' disclosure incentives.

As only an initial step towards understanding the effect of investor learning of unknown parameters (firm profitability) on managers' disclosure incentives, I acknowledge that investor learning can co-exist with other factors that drive disclosure decisions, and discuss in section 1.7 how they affect the empirical findings.<sup>7</sup> A detailed analysis of these forces is beyond the scope of this paper. In addition, the magnitude of the empirical results are strictly within sample and might not generalize to other types of disclosure other than annual earnings forecast.

The paper proceeds as follows. Section 1.2 presents the model. Section 1.3 discusses the research design. Section 4 describes the data. Section 5 presents descriptive analysis followed by discussion of results and prior literature in sections 1.6 and 1.7. I conclude in section 1.8.

---

7. In addition to the multi-period incentives mentioned above, others incentives include but are not limited to incentives to bias or obfuscate disclosures [40, 36], litigation risk [37, 75, 64], preference for aligning investors' expectations [5], capital structure [39], corporate governance [4], and signaling management types [e.g., 80].

## 1.2 Model

This section describes a voluntary disclosure model that features investor learning about underlying firm profitability and examines how learning affects the likelihood of disclosure.

### *1.2.1 Investor learning about firm profitability*

In principle, learning can be broadly defined as investors updating their beliefs about any metrics of their interest upon observing new information. For example, Gao and Liang [41] 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 [33]. In these studies, agents are uninformed about some unknown parameters, for example, product quality [33], firm profitability [70], analysts' forecasting ability [22], or managers' forecasting ability [53]. They update their beliefs after observing noisy signals that contain information about the pertinent parameters.

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, and with prior research on investor learning [70, 74]. Uncertainty about the distribution of earnings relates to the literature on parameter uncertainty, which assumes agents do not know the distribution of stock returns.<sup>8</sup>

---

8. See Brown [20], Barry and Brown [11], Coles and Loewenstein [25], Coles et al. [26], and Xia [86].



## 1.2.2 Model set up

### Time line

The model extends the voluntary disclosure model by Verrecchia [81] to multiple periods and incorporates investor learning as an additional element. Specifically, a firm  $i$  operates for an infinite 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  $\eta_{it} \sim \mathcal{N}(0, \frac{1}{\kappa_i \tau_i})$ , 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 [81]. The manager's utility depends on stock price after the disclosure decision, which is the sum of investors' expectations of discounted unrealized earnings net of current and future disclosure costs,

$$P^c(d_{it}, \mathcal{H}_{it}) = \sum_{t'=t}^T \left\{ \beta_i^{t'-1} \mathbb{E}(y_{it'} - \mathbb{1}\{d_{it'} = s_{it'}\} c_i | d_{it}, \mathcal{H}_{it}) \right\}, \quad (1.1)$$

where  $d_{it}$  is the disclosure decision,  $\beta_i$  is the discount rate for firm  $i$  (exogenous to managers' disclosure decisions), and history  $\mathcal{H}_{it}$  includes all past earnings and disclosures,  $\mathcal{H}_{it} \equiv (\mathbf{y}'_{i,t-1}, \mathbf{d}'_{i,t-1})'$ , where  $\mathbf{y}_{i,t-1} \equiv (y_{i1}, y_{i2}, \dots, y_{i,t-1})'$  and  $\mathbf{d}_{i,t-1} \equiv (d_{i1}, d_{i2}, \dots, d_{i,t-1})'$ .

#### 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 [46, 65]. The cost captures factors that make disclosures expensive. Examples 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.

## 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:

$$y_{it} = \rho_{0i} + \rho_{1i}y_{i,t-1} + \epsilon_{it}, \quad \text{where } \epsilon_{it} \sim \mathcal{N}\left(0, \frac{1}{\tau_i}\right), \quad (1.2)$$

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.<sup>9</sup>

The key assumption is that investors do not know the firm’s profitability  $\rho_{0i}$  and learn it through the manager’s disclosures and 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:

$$\rho_{0i} \sim \mathcal{N}\left(q_{0it}, \frac{1}{\lambda_{it}\tau_i}\right),$$

where  $q_{0it}$  denotes investors’ average assessment of  $\rho_{0i}$  at the beginning of period  $t$  and  $\lambda_{it}$  denotes investors’ information precision relative to the precision of earnings shocks. I call  $\lambda_{it}$

---

9. Prior research has found mixed evidence on the time-series properties of the annual earnings process, suggesting its complexity. Ball and Watts [9] describe annual earnings as a submartingale process. Watts and Leftwich [83] find evidence that supports a random walk model. Albrecht et al. [6] find the choice of AR1 or MA process depends on the firm’s industry.

investors' information precision for short and refer to its reciprocal as investor uncertainty. Note  $\frac{1}{\lambda_{it}\tau_i}$ , the variance of investors' beliefs about  $\rho_{0i}$ , does not necessarily indicate their beliefs are 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 optimistic beliefs could be a result of a history of lucky firm performance and do not necessarily reflect actual high profitability.

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

$$\lambda_{it} = \lambda_{i,t-1} + 1; \quad (1.3)$$

$$q_{0it} = (\lambda_{it})^{-1} (\lambda_{i,t-1} q_{i,t-1} + y_{i,t-1} - \rho_{1i} y_{i,t-2}). \quad (1.4)$$

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

$$\lambda_{it}^s = \frac{\kappa}{\kappa+1} + \lambda_{it}; \quad (1.5)$$

$$q_{0it}^s = (\lambda_{it}^s)^{-1} \left( \frac{\kappa}{\kappa+1} (s_{it} - \rho_{1i} y_{i,t-1}) + \lambda_{it} q_{0it} \right), \quad (1.6)$$

where  $\kappa$  measures the manager's information precision (recall  $s_{it} \equiv y_{it} + \eta_{it}$  with  $\eta_{it} \sim \mathcal{N}(0, \frac{1}{\kappa_i \tau_i})$ ).<sup>10</sup>

As demonstrated later, after realizing nondisclosure, the posterior beliefs follow truncated normal distribution. However, the posterior distribution becomes normal again, after

---

10. The higher  $\kappa$  is, the more weight investors put on manager's disclosure. Note that investors will never put 100% weight on  $s_{it}$  unless  $\lambda_{it} = 0$ . 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.

investors observe the realized earnings. The reason is that manager's private information is only a noisy signal about realized earnings and hence does not provide incremental information once investors observe realized earnings. This result ensures model tractability.

### 1.2.3 Equilibrium strategy

To describe the equilibrium, I re-write stock price  $P^c(d_t, \mathcal{H}_t)$  as

$$P^c(d_t, \mathcal{H}_t) = P(d_t, \mathcal{H}_t) - \mathbf{1}\{d_t = s_t\}c - \beta \mathcal{G}_{d_t}c, \quad (1.7)$$

where  $P(d_t, \mathcal{H}_t)$  is the stock price excluding the disclosure costs, and  $\mathcal{G}_{d_t}c$  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.*

*There exists an equilibrium in which the manager follows a Markovian 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:*

$$P(s_t^*, \mathcal{H}_t) - c = P(s_t \leq s_t^*, \mathcal{H}_t). \quad (1.8)$$

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  $\mathcal{G}_{d_t}$  (investors' expectation of future disclosure costs), both of which are proved in the appendix.<sup>11</sup> Specifically,  $P(d_t, \mathcal{H}_t)$

---

11. The intuition for the independence of  $d_t$  and  $\mathcal{G}_{d_t}$  is as follows. Investors do not know the true value of  $\rho_0$ , and therefore use their perceived distribution of future earnings (instead of the true distribution) to compute the expected future disclosure cost. With mandatory disclosure, that is, true earnings are observed at the end of the fiscal year, investors' beliefs about  $\rho_0$  remain normal even following non-disclosure, implying that investors' perceived distribution of future earnings is also normal. Normality implies the probability of disclosure is independent of the mean of the distribution (because although smaller mean reduces the threshold value, the distribution also shifts to the left, resulting in a zero net effect on the probability of

is linear in the manager's private information (when  $d_t = s_t$ ) or investors' expectation of the manager's private information (when  $d_t = \emptyset$ ),

$$P(d_t, \mathcal{H}_t) = A(\kappa, \lambda_t, \beta, \rho_1) s_t(d_t) + B(\kappa, \lambda_t, \beta, \rho_1, p_{0t}, y_{t-1}), \quad (1.9)$$

where the expressions of  $A$  and  $B$  can be found in the appendix.

## Investor learning and disclosure strategy

From the linearity result, the disclosure threshold  $s_t^*$  and the probability of disclosure  $\Pr\{d_t = 1\}$  can be derived as:

$$s_t^* = \sigma_{st} \tilde{s}_t^* + q_{0t} + \rho_1 y_{t-1}, \quad (1.10)$$

$$\Pr\{d_t = 1 | \mathcal{H}_t\} = 1 - \Phi\left(\frac{\sigma_{st} \tilde{s}_t^* + q_{0t} - \rho_0}{\sigma_s}\right), \quad (1.11)$$

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,

$$\sigma_{st} = \sqrt{\left(\frac{\lambda_t + 1}{\lambda_t} + \frac{1}{\kappa}\right) \frac{1}{\tau}}; \quad (1.12)$$

$$\sigma_s = \sqrt{\frac{1}{\kappa\tau}}. \quad (1.13)$$

From (1.11), lower investor beliefs increase disclosure. Intuitively, lower investors' beliefs give the manager more room to convey good news to improve market perception, so the manager is more likely to disclose information. From (1.12), greater investor uncertainty increases disclosure. The reason is that greater implies larger revision of investors' beliefs given a disclosure, which in turn increases the relative benefit to disclose. The derivations

---

disclosure). Therefore, current disclosure decision does not affect investors' expectation of future disclosure costs.

of these effects, including the effects of other determinants of disclosure, can be found in the appendix.

Investor learning induces serial correlation in the likelihood of disclosure as follows. Because they do not know firm profitability, investors put positive weight on their prior beliefs about profitability in forming their posterior beliefs, resulting in serial correlation in investors' beliefs as in (1.4). Because managers account for the impact of a disclosure on stock price as in (1.1), investor learning, by affecting stock price, induces serial correlation in the likelihood of disclosure. The magnitude of the serial correlation is an empirical question.

## 1.3 Research design

### 1.3.1 *Empirical quantifications*

The empirical analysis aims to quantify two effects investor learning has on managers' disclosure decisions. I first examine how changes in investors' beliefs and investor uncertainty affect the likelihood of a disclosure. To gauge the empirical importance of investor learning, I compare the effects of investor learning with those associated with changes in other determinants of disclosure, for example, the precision of managers' information.

I then estimate the serial correlation in the likelihood of disclosure induced by investor learning. Focusing on the likelihood of disclosure is consistent with the model in which investor learning affects the disclosure threshold, which maps into the probability of disclosure. For completeness, I also examine whether the model adequately captures the intertemporal correlation in managers' voluntary disclosure decisions.

The empirical analysis requires a system that describes investors' learning process for each firm, that is, the evolution of investors' beliefs. Disentangling the impact of investor learning needs to hold other firm characteristics constant. I next describe the empirical approach.

### 1.3.2 *Empirical methods*

Instead of assuming a linear relation between disclosure decisions, investors' beliefs, and investor uncertainty, I structurally estimate the model presented in section 1.2, that is, 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 that investor learning affects disclosure decisions requires finding an instrument that is related to investors' beliefs but not firm fundamentals. Finding such an instrument is next to impossible, because belief changes necessarily result from changes in firm fundamentals. Fourth, subsequent empirical analysis requires simulating many samples for a given firm, which a structural model can easily achieve.

Structural estimation 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 instruments. The causal mechanism is transparent from the model structure. 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 gradually learn firm profitability is not an unreasonable starting point for understanding the effects of investor learning and has empirical support from prior literature [e.g., 70, 53]. Nevertheless, economic agents will never behave exactly as the model predicts. What matters is whether the model reasonably describes the behavior of economic agents, which I evaluate in section 1.6.2.

Another drawback is that structural estimation excludes many incentives identified by prior literature on management forecasts, which include but are not limited to equity is-

suance [39], warning investors [55], litigation risk [75], aligning investors' expectations [5], and corporate governance [4]. 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 learning, the current learning model can be viewed as a benchmark; other incentives can modify investors' and/or managers' behavior.

### 1.3.3 Estimation procedure

The first step of the structural estimation specifies the outcome variables. The model in section 1.2 characterizes the data generating process of disclosure decisions, disclosure contents, firm stock prices, and realized earnings, which are summarized below:

1. The manager's disclosure decision  $d_{it}$  trades off the stock price following disclosure with that following nondisclosure. Disclosure incurs an exogenous cost  $c_i$ .
2. 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,  $s_{it} = y_{it} + \eta_{it}$  with  $\eta_{it} \sim \mathcal{N}(0, \frac{1}{\kappa_i \tau_i})$ .
3. Stock price  $P_{it}^c$  is investors' expectation of the sum of discounted future earnings minus disclosure costs. 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.
4. Firm earnings follow an AR1 process,  $y_{it} = \rho_{0i} + \rho_{1i} y_{it-1} + \epsilon_{it}$  with  $\epsilon_{it} \sim \mathcal{N}(0, \frac{1}{\tau_i})$ .

Structural estimation recovers the joint distribution of the eight model primitives from the outcome variables: investors' prior beliefs  $q_{0it}$ , investor uncertainty  $\frac{1}{\lambda_{it}}$ , firm profitability  $\rho_{0i}$ , earnings persistence  $\rho_{1i}$ , the volatility of earnings shocks  $\frac{1}{\tau_i}$ , the cost of capital  $\frac{1}{\beta_i}$ , manager's information precision  $\kappa_i$ , and the fixed cost of disclosure  $c_i$ . Except for investors' beliefs  $q_{0it}$  and investor uncertainty  $\frac{1}{\lambda_{it}}$ , all other parameters are assumed to be exogenous to



the manager’s disclosure decisions and fixed over time, which represents a trade-off between tractability and empirical richness.

The estimation starts with the joint likelihood of the four outcome variables as a function of the eight model primitives. Firms differ in the values of their primitives, which are assumed to follow a pre-specified distribution with unknown parameters (called hyper-parameters by the literature) to be estimated.<sup>12</sup> For example, I assume the natural logarithm of investors’ prior beliefs at the beginning of the sample period follows a normal distribution  $\log(q_{0i0}) \sim \mathcal{N}(\bar{q}_0, \sigma_q)$ , where  $\bar{q}_0$  and  $\sigma_q$  are to be estimated.<sup>13</sup>

I then use Bayesian estimation to find the distribution of the model primitives conditional on the data. Denote the model primitives of firm  $i$  as  $\Theta_i$ . The estimation recovers the distribution of  $\Theta_i$  conditional on the data. The rough idea of Bayesian estimation is to specify a prior distribution for the model primitives, and use the Bayes rule to recover the posterior distribution of the parameters conditional on the data:

$$\pi(\Theta_i|\text{data}) \propto L(\text{data}|\Theta_i)f(\Theta_i), \tag{1.14}$$

where  $\pi(\cdot)$  is the posterior density to be recovered,  $L(\cdot)$  is the likelihood function, and  $f(\Theta_i)$  is the prior belief specified by the researcher. Note that  $f(\Theta_i)$  can be arbitrary. I choose prior beliefs that have minimal influence on the estimation, that is, diffuse priors. Interested readers can find the details in the appendix.

---

12. 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.

13. See the Appendix for more details.

### 1.3.4 Identification

The model primitives are identified through how they affect the joint likelihood of the outcome variables. First, differences between actual earnings and the voluntary disclosure of firm  $i$  identify  $\kappa_i$ , the information precision of firm  $i$ 's manager. Larger differences imply a smaller  $\kappa_i$ , that is, lower precision. The identification of  $\kappa_i$  relies on the linearity assumption between actual earnings and managers' disclosures, which is  $s_{it} = y_{it} + \eta_{it}$ , where  $\eta_t \sim \mathcal{N}(0, \frac{1}{\kappa_i \tau_i})$ .

Second, the disclosure cost  $c_i$  is identified through three channels. The disclosure cost enters stock price directly in case of a disclosure but does not directly enter stock price in case of nondisclosure. Differences between stock prices following a disclosure and nondisclosure identify the disclosure cost. Moreover, a higher disclosure cost is associated with lower stock prices, which serves as another channel for identifying the disclosure cost. Identifying  $c_i$  requires the additive separability of  $c_i$  from other components of stock price.

A third channel to identify the disclosure cost is changes in disclosure decisions. Because disclosure decisions change infrequently, using variations in disclosure decisions to identify the disclosure cost turns out to be difficult. Fortunately, stock prices alone can identify the disclosure cost, as discussed above. Untabulated simulation studies suggest that a low disclosure cost is harder to be identified than a high disclosure cost. So having variations in disclosure decisions can help identify the disclosure cost.

Third, investors' belief precision  $\lambda_{i0}$  and the discount factor  $\beta_i$  are identified from the sensitivity of stock prices to managers' disclosures. Investors' belief uncertainty varies over time according to the functional form of (1.3). 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 (1.9)). Identifying  $q_{i0}$  depends on the assumption that investors follow the Bayes rule. when updating their beliefs about  $\rho_{0i}$ . Finally, the time series of

earnings is informative about  $\rho_{0i}$ ,  $\rho_{1i}$ , and  $\tau_i$ .<sup>14</sup>

## 1.4 Data and sample selection

The empirical analysis examines management annual earnings forecast decisions. The choice is motivated by the way I model stock price as a function of expected future earnings and with the assumption that managers have the relevant information to disclose. 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 after January 2003 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. [24] document that about 41% of their hand-collected forecasts are not covered by the CIG database. I show 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. [24].<sup>15</sup>

Following Houston et al. [51], I drop forecasts for a particular quarter that are issued after the corresponding fiscal year-end date, 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,  $d_{it} = \emptyset$ .

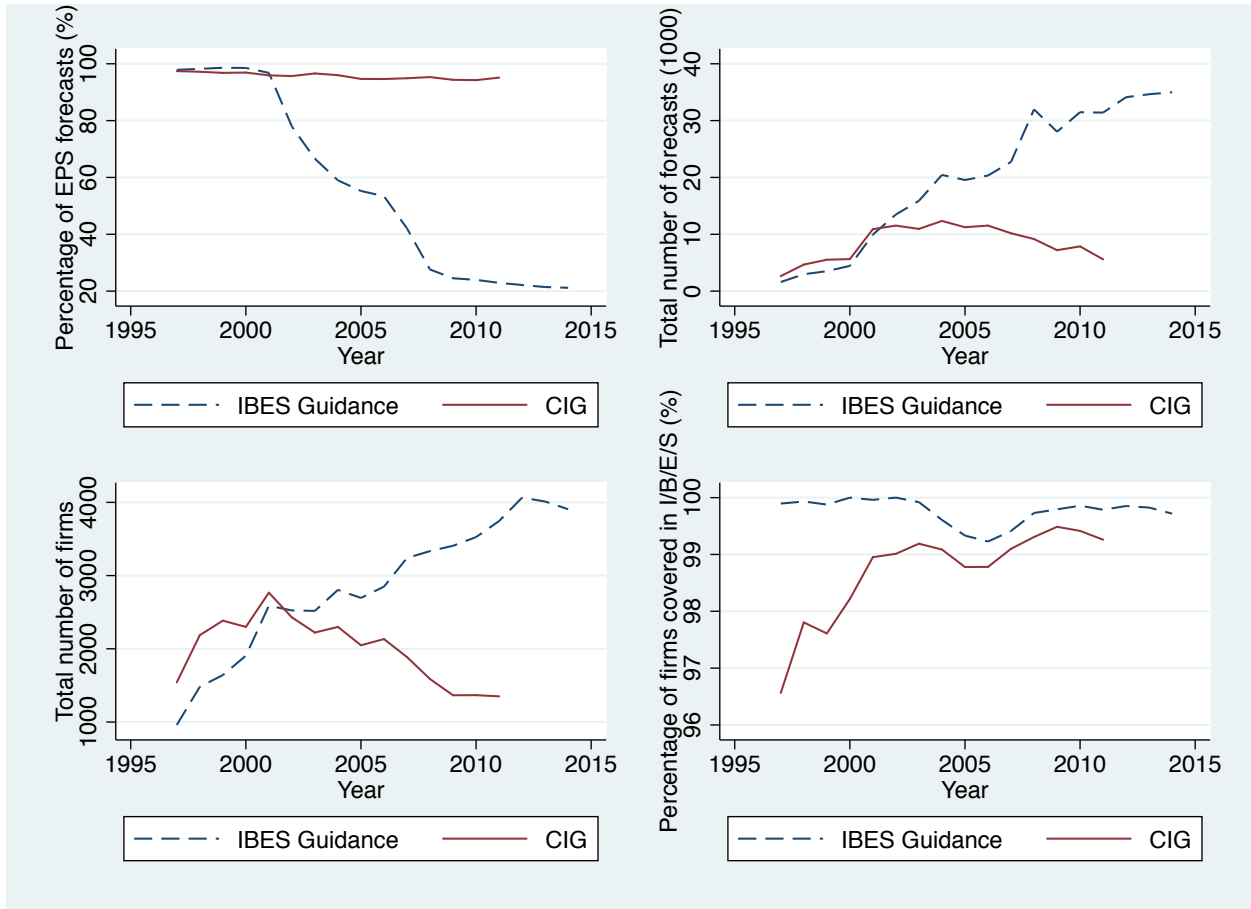
---

14. That identifying these parameters does not depend solely on disclosure frequency alleviates the concern of a mechanical relation 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.

15. See Table 1.1 for the comparison between I/B/E/S Guidance and the hand-collected sample of Chuk et al. [24] and Figure 1.1 for the comparison between I/B/E/S Guidance and CIG.



Figure 1.1: 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.

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 is not pursued.

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 an annual forecast by then. I require at least five consecutive years of nonmissing earnings (from I/B/E/S) and stock prices (from CRSP). In the end, I have 6,111 firm-year observations. Table 1.2 describes the sample construction in detail.

## 1.5 Descriptive analysis

In this section, I describe how investors respond to a disclosure and the absence of a disclosure, the pattern of disclosure dynamics, as well as other outcome variables used for the empirical analysis.

### *1.5.1 Summary statistics of the variables used for estimation*

This section describes the data used for the empirical analysis. Each firm has four time series: annual earnings, management forecast decisions, management forecasts, and stock prices. To increase comparability across firms, I scale all variables by the stock price at the beginning of the sample period except for the forecast decision dummies. Such scaling is necessary, because the way investor learn as shown in equation (1.4) suggests that the time series of earnings can only admit a single scaler across all time periods, that is, scaling both the left hand side and right hand side by the same number. Moreover, the scaling ensures a tractable model of disclosure decision. Scaling earnings by concurrent stock price requires

Table 1.2: Sample selection

This table describes the sample selection procedure for management forecasts.

Sample	Observations
IBES Guidance 1992-2014	363,754
Non-US firms, firms not matched to Compustat, firms with no I/B/E/S tickers and forecasts issued before 2002	292,594
Drop firms that issue zero or one forecast and firms that issue FFO (funds from operations) forecasts <sup>a</sup>	280,535
Drop forecasts issued after the corresponding fiscal year-end	263,837
Keep firms that issue EPS forecasts each year of forecast issuance; Drop non-EPS forecasts	74,999
Keep firms that issue annual forecasts each year of forecast issuance; Drop quarterly and semi-annual forecasts	39,141
Keep firms with more than five consecutive years of actual EPS from I/B/E/S	36,583
Drop firms that do not issue forecast for the subsequent fiscal period;	31,890
Drop duplicate forecasts for each firm-year <sup>b</sup>	7,451
Drop firms that issue forecasts only in one year	7,239
Drop firms with fewer than five consecutive years of nonmissing CRSP price and earnings forecast decision data.	6,111

<sup>a</sup>. These firms are mostly REITS. The relation between FFO and price is likely to be different from that between EPS and price.

<sup>b</sup>. I keep the earliest forecast for each firm-year.

Table 1.3: 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 1.2). 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.

	Mean	SD	25%	50%	75%	Obs
Earnings $y_{it}$	12.262	17.754	5.938	9.141	15.356	6,111
Price $p_{it}$	214.091	272.694	100.000	145.613	232.368	6,111
Disclosure decision $d_{it}$	0.873	0.333	1.000	1.000	1.000	6,111
Management forecast $s_{it}$	13.329	15.928	6.513	9.528	15.561	6,111

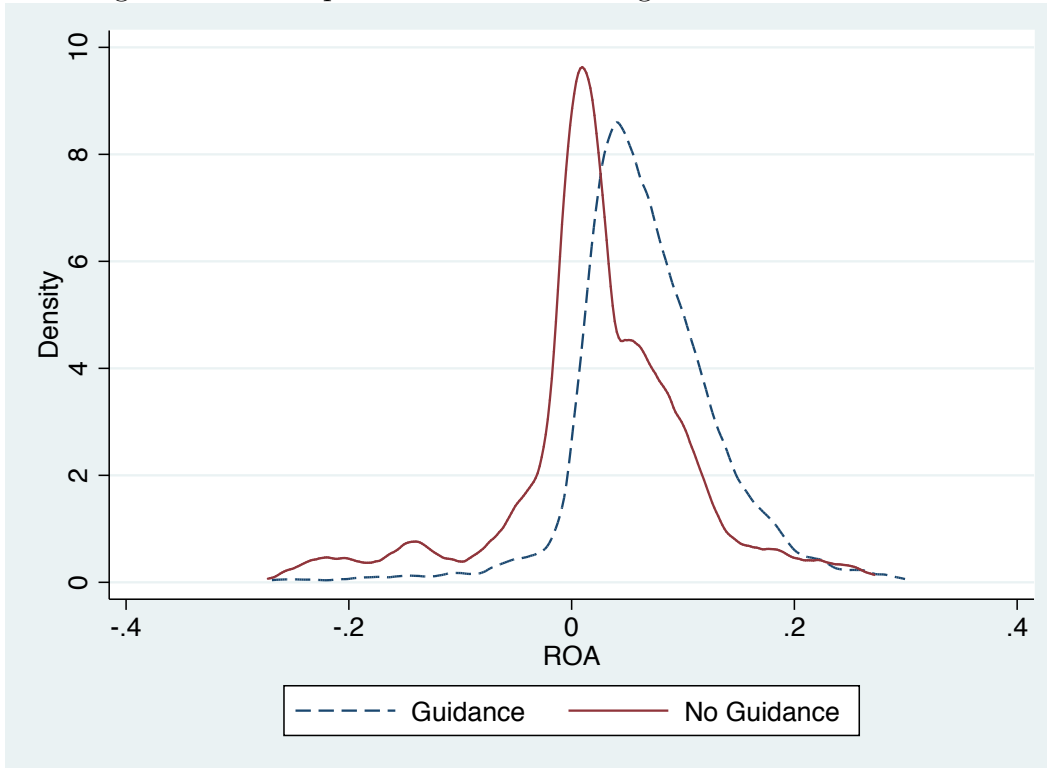
investors to form beliefs about not only future earnings but also future stock prices, which significantly complicates the model.

Table 1.3 shows 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., 23].

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. Although economic 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 assume a hidden long-run earnings process that evolves over time [e.g., 69]. Doing so requires identifying the hidden process, which is hard given the length of the time series (12 years), and is not further pursued.



Figure 1.2: Firm performance and management forecast decisions



This figure presents the density plots of firm performance separately for firm years with and without a management forecast. The horizontal axis represents firm performance, measured as net income divided by lagged total assets. The vertical axis is the density. The dashed (solid) line represents firm years with (without) a management forecast.

The identified earnings process pooling across years can be thought of as a “steady-state” earnings process.

### 1.5.2 How investors respond to (non)disclosures

This section focuses on investors’ response to a management forecast decision. The model predicts that investors’ reactions to a management forecast should be more favorable than their reactions to the lack of a management forecast. Figure 1.2 and 1.3 provide evidence consistent with this prediction.

Figure 1.2 shows that forecast provision is associated with better firm performance, measured as return on asset (net income divided by total asset of the previous year). In fact,

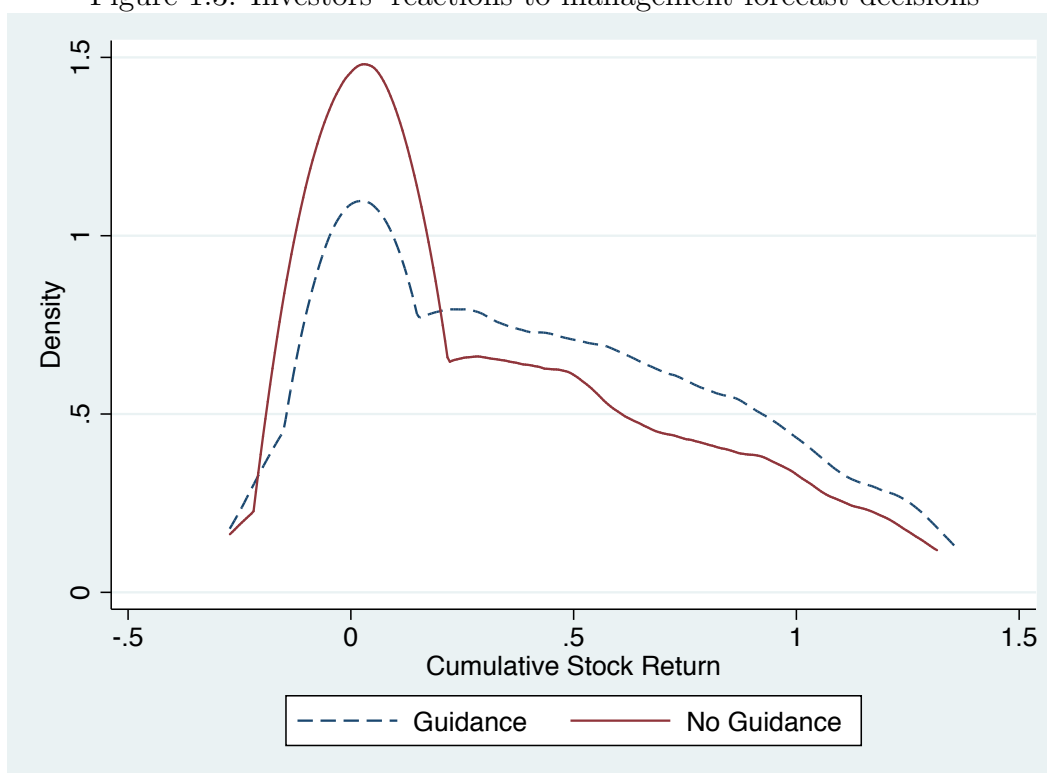
the distribution of return on asset for firm years with a management forecast (the dashed line) first order stochastic dominates that for firms years without a management forecast (the hard line). The pattern is consistent with numerous studies that use firm performance to explain managers' earnings forecast decisions [e.g., 23, 51].

Figure 1.3 depicts stock prices following a forecast decision. To facilitate comparison across firms, I use the log ratio of the stock price following a forecast decision to the stock price at the beginning of the sample period, that is, log cumulative stock returns. As explained before, if a manager issues a forecast, 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. Figure 1.3 shows that the lack of a forecast is associated with lower stock prices than providing a forecast. This result confirms the performance results in Figure 1.2. Chen et al. [23] find that investors' reaction to the public announcement of stopping earnings guidance is on average negative. The evidence here supports their finding with a larger sample that possibly contains firms that do not publicly announce their decision to withhold management forecasts.

### *1.5.3 Dynamics of management forecast issuance*

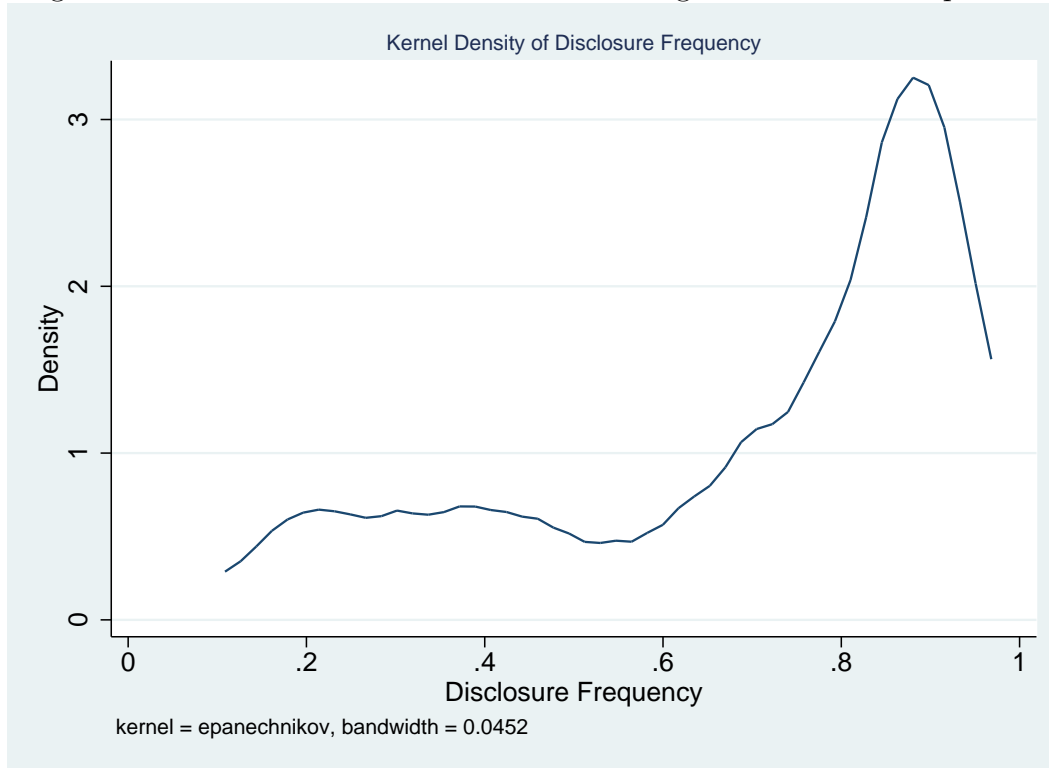
This section examines the time series dynamics of management earnings forecast decisions. Table 1.3 shows managers issue forecasts frequently, about 87% of all firm-years. About 57% of the firms in my sample always issue forecasts in the sample period (untabulated). This finding is not surprising, because I require a firm to issue at least one annual earnings forecast within a fiscal year to classify the firm year as disclosure. I keep firms that always issue forecasts in the sample for the empirical estimation, because they remain useful in understanding the average tendency to issue forecasts. For example, consistent disclosures could indicate a low disclosure cost other things equal. However, the question of how past and future disclosure incentives are related is not meaningful for these firms, because the likelihood of disclosure for them is close to one. So in this case, it is impossible to run a

Figure 1.3: Investors' reactions to management forecast decisions



This figure presents the density plots of cumulative stock returns separately for firm years with and without a management forecast. The horizontal axis represents firm cumulative stock return, measured as the log of the ratio of the stock price following a management forecast decision to the stock price of that firm at the beginning of the sample period. The timing of stock price is chosen as follows. When a manager issues a forecast in a year, I use the stock price three days following the forecast to compute the stock return. When a manager does not issue a forecast in a year, I use the stock price at the end of the final trading date of the fiscal year to compute the stock return. The vertical axis is the density. The dashed (solid) line represents firm years with (without) a management forecast.

Figure 1.4: Cross-sectional distribution of management forecast frequencies



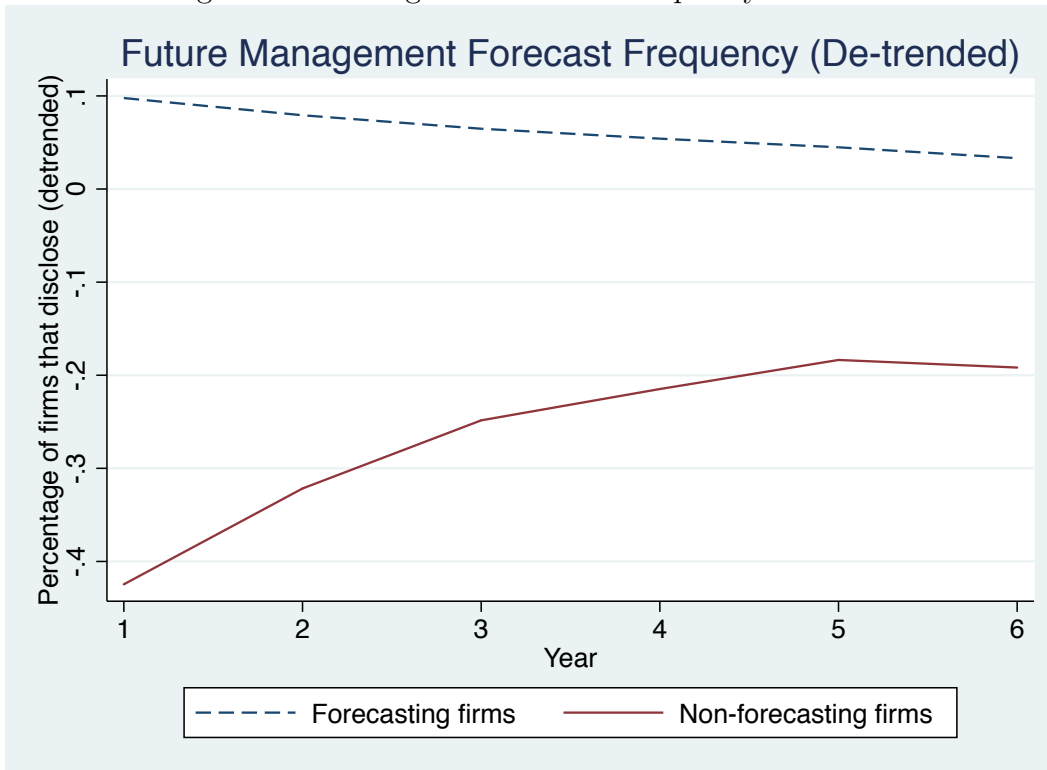
This figure presents the cross-sectional density of management forecast frequency for firms that do not always disclose over the sample period from 2003 to 2014. Disclosure frequency is computed as the percentage of years with forecast issuance. The horizontal axis represents the management forecast frequency. The vertical axis is the density.

regression of current disclosure decisions on past disclosure decisions for these firms, because the disclosure decisions have zero variance over time.

Figure 1.4 plots the cross-sectional density of forecast frequencies between 2003 and 2014 for firms that do not always issue forecasts. 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. [74] that management forecasts can be regular or sporadic.

Next, I examine the times-series properties of forecast decisions. Figure 1.5 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.

Figure 1.5: Management forecast frequency over time



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.

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 disclosure decisions being sticky. On the other hand, disclosure frequencies tend to go down (up) for forecasting (non-forecasting) firms, which suggests disclosure decisions are mean-reverting.

Table 1.4: 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.

	Mean	SD	25%	50%	75%
Investors' prior belief $q_{0i}$	12.032	28.150	2.226	4.207	9.160
Investors' belief precision $\lambda_i$	5.572	3.057	3.990	5.183	6.674
Disclosure cost $c_i$	5.692	4.924	3.338	4.559	6.155
Expected return $r_i = \frac{1-\beta_i}{\beta_i}$	0.111	0.090	0.061	0.080	0.129
Manager's information precision $\kappa_i$	4.032	5.333	1.867	2.918	4.584
Intercept of earnings process $\rho_{0i}$	6.637	3.472	4.546	5.855	7.425
Earnings persistence $\rho_{1i}$	0.630	0.171	0.516	0.640	0.771
Variance of earnings process $1/\tau_i$	40.359	84.709	2.776	7.766	24.001

## 1.6 Results

### 1.6.1 Parameter estimates

Structural estimation recovers the joint distribution of model parameters for each firm. For each firm, I compute the median estimate of each model parameter (see section 1.3.3 for the list of all parameters) and report in Table 1.4 the cross-sectional distribution of these medians.

Table 1.4 shows that 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. The fact that investors' beliefs are different from actual firm profitability does not necessarily imply that investors are irrational. Such beliefs 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 stock prices. In addition, because I only have access to 12 years of data for individual firms, it may well be true that investors' beliefs about profitability are eventually right but the estimated actual profitability using earnings data is wrong. Empirical inferences do not depend on this, because I simulate firms based on the cross-sectional distribution of profitability. It is less likely that the distribution of profitability is wrongly estimated.

Another finding from Table 1.4 is that investors are confident about their prior beliefs. The average (median) belief precision is 5.572 (5.183) at the beginning of the sample period. Investors put significant weight on their prior beliefs when forming their posterior beliefs. They behave as if they had access to 5.572 years of historical data (Bayesian investors equally weight all historical data when forming their beliefs and precision increases by one after each year by construction). Exploring the data suggests high belief precision is associated with low volatility of firm earnings, consistent with the intuition that investor uncertainty is low 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., 73, 29, 38]. The average (median) expected return, computed using  $\frac{1-\beta_i}{\beta_i}$ , is 11% (8%) with a standard deviation of 9%. The magnitude is comparable to the expected return estimated by Easton [31], 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. One can parametrize the disclosure cost as a function of the determinants of disclosure identified by prior research such as earnings volatility [84], analyst optimism [27] and litigation risk [75]. Doing so assumes that these determinants are associated with frictions that change managers' disclosure incentive and at the same time affect firm stock price, which could be ad hoc and adds to the estimation burden (which already takes 4 days). Another reason for the high disclosure cost could be that part of the cost of disclosure could actually be non-monetary, which do not affect stock prices. 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). The latter, possibly due to the more complex structure, produces very large variance in the disclosure cost estimates. Subsequent analysis relies on results from the original model that does not differentiate between monetary and nonmonetary disclosure costs, because the question here is not to examine other sources outside the model that affect disclosure decisions.

### *1.6.2 Model fit*

This section evaluates the within sample model fit. A good model fit increases the confidence that the imposed structure approximates the reality. For forecast decisions, I first examine how well the model captures the forecast frequency for each firm, where forecast frequency refers to the percentage of earnings forecasts out of the entire sample period. I then evaluate



the intertemporal dynamics of forecast decisions, measured as the autoregressive coefficient of current disclosure decisions on previous year disclosure decisions. I finally examine the model fit of firm stock prices.

## Model fit of forecast frequency

For each firm in the sample, I compare forecast frequency predicted from the model with the actual forecast frequency and report a  $p$  value that indicates whether the predicted forecast frequency is significantly different from the actual forecast frequency. I follow the procedure below.

1. Bayesian estimation recovers the distribution of model parameters for each individual firm. I draw one set of model parameters for each firm from their estimated distribution (see section 1.3.3 for the list of all parameters).

2. For each firm year, I generate a forecast decision by comparing management forecast with the threshold computed from the equilibrium condition specified in Proposition 1.

In case of disclosures, I compare the actual management forecasts of each firm with the disclosure thresholds computed from the model. Computing disclosure thresholds require investors' beliefs as the input. Investors' subsequent beliefs are updated according to (1.4), using actual earnings of each firm.

In case of nondisclosures, because I do not observe management forecasts that managers withheld, I draw an imputed management forecast using  $s_{it} = y_{it} + \eta_{it}$ ,  $\eta_{it} \sim \mathcal{N}(0, \frac{1}{\kappa_i \tau_i})$ , where  $s_{it}$  is the management forecast,  $y_{it}$  is actual earnings,  $\eta_{it}$  is a mean zero random noise,  $\tau_i$  is the reciprocal of earnings noise, and  $\kappa_i$  is the relative precision of management forecast. I then compare the imputed management forecast with the disclosure threshold.

3. I then compute the percentage of management forecasts using simulated forecast decisions.

4. I repeat the steps above for 100 set of parameter draws for each firm, to generate a distribution of forecast frequency, from which I compute  $p$  value.

Table 1.5 Panel A shows that the structural estimation captures the forecast frequency reasonably well. The mean of the forecast frequency in the data is 87%, very close to the simulated forecast frequency of 85%.

The remainder of Table 1.5 Panel A tests the (alternative) hypothesis that the predicted forecast frequency is different from the actual forecast frequency. I compute  $p$  value in two ways. Among firms that do not always issue management forecasts (which constitute about 43% of all firms), I conduct a two-sided test by comparing the empirical distribution of the predicted forecast frequency with the actual forecast frequency. I cannot reject the hypothesis that the simulated forecast frequency is the same as the actual forecast frequency at a 5% level for about 84% of firms that always issue management forecasts. The median  $p$  value is 0.37. For firms that always issue forecasts (which constitute about 57% of all firms), a  $p$  value cannot be computed in the same way, because the frequency of predicted disclosure is truncated above at one. I instead compute the percentage of times that the simulated forecast frequency is one. I cannot reject the hypothesis that the simulated disclosure frequency is one for all the firms that always issue management forecasts. The median  $p$  value is 0.95.

## Model fit of the intertemporal dynamics of management forecast decisions

Evaluating the intertemporal dynamics of management forecast decisions serves two purposes. First, the precision of investors beliefs and their beliefs evolve mechanically over time according to equation (1.3) and (1.4). Imposing the mechanical structure of (1.3) and (1.4) might be too stringent. A good fit for the intertemporal dynamics of management forecast decisions provides some support for imposing the structure of (1.3) and (1.4), that is, the assumption of simple Bayesian learning. Second and most crucially, the fact that the model well captures the intertemporal dynamics of actual forecast decisions increases the confidence in the later counterfactual analysis of the intertemporal dynamics of the likelihood of

Table 1.5: Model fit

This table presents the model fit of disclosure decisions (Panel A and B) and stock prices (Panel C). Panel A compares simulated forecast frequency with the actual forecast frequency and report a  $p$  value that indicates whether the difference between the two is significant. Panel B evaluates the model fit for the intertemporal dynamics in disclosure decisions. I obtain the coefficient  $\beta_1$  from regressing the current year forecast decision on previous year forecast decision using actual data.  $R^2$  is the variance of model residuals to the variance of stock prices. For each observation, I draw 100 sets of parameters, which are used to generate 100 simulated stock prices. I then compute the fitted stock price by averaging over the 100 simulated stock prices.

	Mean	SD	25%	50%	75%
Panel A: Forecast frequency across firms					
Actual forecast frequency in the data: $\sum_t d_{it}/T_i$	0.869	0.217	0.833	1.000	1.000
Simulated forecast frequency: $\sum_t \hat{d}_{it}/T_i$	0.852	0.224	0.814	0.967	0.996
p value (excluding firms that always issue forecasts)	0.414	0.310	0.120	0.370	0.680
p value (firms that always issue forecasts)	0.882	0.164	0.850	0.950	0.990
Panel B: Pooled regression $d_{it} = \beta_0 + \beta_1 d_{it-1} + \epsilon_{it}$					
Regression coefficient $\beta_1$ (Data):	0.640				
Regression coefficient $\beta_1$ (Simulated):	0.630	0.023			
p value	0.570				
Panel C: Stock prices					
$1 - \frac{\sum_t (p_{it} - \hat{p}_{it})^2}{\sum_t (p_{it} - \bar{p}_{it})^2}$	0.205	0.887	0.077	0.379	0.680

a forecast.

I measure the intertemporal dynamics of management forecasts from the sample using the first order autoregressive coefficient of the management forecast decision time series, that is, the slope coefficient from a pooled cross-sectional and time series regression of current year forecast decision on the previous year forecast decision. I compare the slope coefficient computed from the actual forecast decisions with those from simulated forecast decisions.

I follow a similar simulation procedure to that in section 1.6.2. The first two steps of the simulation procedure are identical to the first two steps of the procedure explained in section 1.6.2. For the third step, I pool the predicted disclosure decisions across all firms and time periods and compute one slope coefficient. Then I repeat the steps above 100 times to generate a distribution of the simulated slope coefficients. Note that it is not necessarily true that a good model fit for the average forecast issuance frequency means an equally good model fit for the intertemporal dynamics. For example, there might be temporal deviation of the likelihood of disclosure from the corresponding long run mean that the model fails to capture.

Table 1.5 Panel B shows the slope coefficient using simulated forecast decisions. The average is 0.63 with a standard deviation of 0.023, not statistically different from the slope coefficient using actual forecast decisions, which is 0.64. The result suggests that the structural model captures the intertemporal dynamics of management forecast decisions at least in terms of the first order autoregressive coefficient of the management forecast decision time series.

## Model fit of stock prices

To measure the model fit of stock prices, I compute R squared of stock prices for each firm. R squared is one minus the ratio of the variance of the model residual to the variance of stock prices. R squared can be negative, because the variance of the model residual can be larger than the variance of stock prices. In this case, using the average stock price to explain

stock prices produces better model fit than using the structural model. A higher R squared means a better model fit for stock prices. Table 1.5 Panel C shows that for most firms, the structural model has a better fit than using the average stock price. The average (median) R squared is 0.205 (0.379).

### 1.6.3 *Investor learning and persistent disclosure incentives*

In this section, I examine the effects of investor learning on the persistence of managers' disclosure incentives, that is, the likelihood of a management annual earnings forecast. Section 1.6.2 demonstrates that the model adequately captures the intertemporal dynamics of actual disclosure decisions. A disclosure decision is affected by news unobserved to researchers. All researchers can observe and examine is the likelihood of disclosure. An important question this paper aims to address is how a learning-induced change in the likelihood of disclosure propagates into the future. That is, to what extent changes in the current disclosure incentive are associated changes in future disclosure incentives through the channel of investor learning?

To answer this question, I first generate 1,000 firms using the distribution of parameter estimates. To test statistical significance, I simulate 500 samples each with 7 years for each simulated firm, totaling 3,500 observations. I compute the probability of a management forecast from Proposition 1.

Second, for each simulated firm, I regress the forecast probabilities of years 3-7 each on the forecast probabilities of year 2.<sup>16</sup> The slope coefficients of the five regressions capture how changes in the forecast probability in year 2 are associated with the forecast probabilities in years 3-7 respectively. Specifically, I use the following model:

$$\Pr\{d_{i,t+2} = s_{i,t+2} | \mathcal{I}_{i,1}\} = \alpha_{i,t} + \beta_{i,t} \Pr\{d_{i,2} = s_{i,2} | \mathcal{I}_{i,1}\} + \varepsilon_{i,t+2}, \quad t \in \{1, 2, 3, 4, 5\},$$

---

16. Beliefs only start to vary starting from the second year. For a given firm, the initial beliefs are assumed to be the median of the estimated distribution.

where  $\mathcal{I}_{i,1}$  includes investors' posterior beliefs after observing earnings of year 1.

It is important to realize that shocks to earnings generate variations in investors' beliefs in year 2 and onwards, which change the likelihood of disclosure in subsequent years. Because the 500 samples of each firm share the same set of model parameters such as firm risk, earnings persistence, etc., the simulation procedure ensures that variation in investors' beliefs and uncertainty is solely responsible for the variation and intertemporal correlation in management forecast probabilities.<sup>17</sup>

Table 1.6 shows the cross-sectional distribution of the regression slopes,  $\beta_{i,t}$ . The first row suggests that an increase of 10% in the current forecast 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 is almost perfectly associated with that in the following year. Table 1.6 also suggests that the 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. Combined with the finding that the model well fits the intertemporal dynamics of the forecast decisions in the sample (see section 1.6.2), the evidence in this section further suggests that investor learning is an economically important factor driving the intertemporal dynamics of managers' voluntary disclosure decisions.

---

17. One might question whether shocks to earnings by themselves could change the likelihood of disclosure. I prove that without investors' uncertainty about profitability, the likelihood of disclosure in the model does not change with shocks to earnings. The reason is that a shock to current earnings changes the actual mean of next period earnings as well as investors' beliefs about the mean of next period earnings. The difference between the two is always zero, when investors know the true mean of earnings, i.e., investors do not make any mistakes. The likelihood of disclosure stays constant, because it depends on the difference between the two. In contrast, when investors are uncertainty about the true mean of earnings, investors are sluggish in updating their beliefs, which introduces the intertemporal dependence in the likelihood of disclosure.

Table 1.6: 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 (1.11) 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.

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

#### 1.6.4 *Investor learning and the average forecast frequency*

I use the frequency of forecast issuance over the subsequent 12 years to measure the average future disclosure incentive. I choose 12 years as the measurement window to be consistent with the length of the sample period for most firms (from 2003 to 2014). The purpose is to quantify whether investors' prior beliefs and uncertainty have economically significant effects on the likelihood of future management forecasts. Their effects could be trivial, given the presence of other model primitives that also affect management forecast decisions.

I again compute forecast frequencies via simulation. 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 future forecast frequency for that firm. Simulation allows me to control for such random variation by fixing  $\rho_1$  for particular firms but still allow for cross-sectional distribution in  $\rho_1$ . The simulation procedure works as follows.

1. I generate 2,000 firms using the cross-sectional distributions of firm level parameters (see section 1.3.3 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 and assuming the model is correct.
2. For each sample of a simulated firm, I compute the percentage of the total number of forecast issuances over the 12 years. I obtain the average forecast frequency for each firm by averaging over the 1,000 sample forecast frequencies.
3. The first two steps are similar to those for computing model fit, except that actual earnings and management forecasts are simulated instead of using the actual data. The final step computes the marginal effects associated with each model parameter.



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 forecast frequency as a proportion of the standard deviation of the forecast frequency in the data, which is 22 percentage points. For example, I vary  $\lambda_i$  of each firm by 3.057 (the standard deviation of  $\lambda_i$ ) holding other parameters constant, compute the change in the average disclosure frequency, and divide the change by 0.22. The economic interpretation is the change in the average forecast frequency over a 12-year window that is caused by one standard deviation change in the precision of investors' prior beliefs at the beginning of the 12-year window and measured as the number of standard deviations of the empirical forecast frequency. For investors' beliefs, I additionally 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 percentage points—too large to be considered “marginal,” especially given the standard deviation of earnings of 17 percentage points. Using earnings shock is consistent with the idea that changes in investors' beliefs about profitability should be closely related to the size of firm earnings shocks.

Figure 1.6 and Table 1.7 present the distribution of the simulated disclosure frequencies (the percentage of disclosure in twelve years) and the marginal effects for other determinants of disclosure. 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%). The 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.

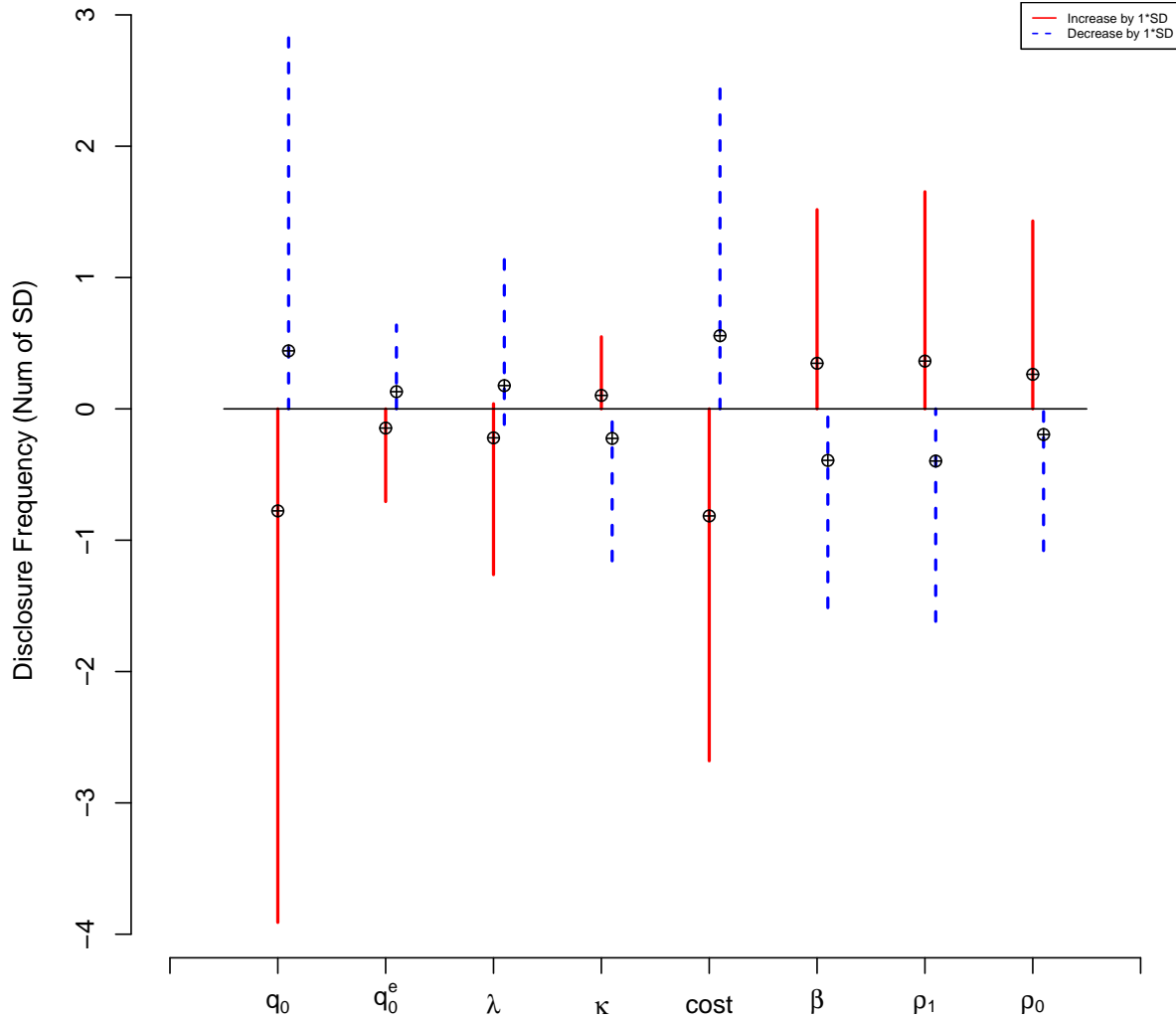
Table 1.7 demonstrates that the effects of investor learning, that is, the effects of investors' prior beliefs and uncertainty, are extremely heterogeneous across firms. The marginal effects are zero for about 40% of firms (the final column). These firms always disclose, so a marginal change in one parameter is not sufficient to change forecast incentives of these firms. The

Table 1.7: Marginal effects on future disclosure incentives

This table presents the marginal effects of the following structural parameters: investors' prior belief  $q_{0i}$ , 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.

	Mean	Median	SD	5%	25%	75%	95%	Zero(%)
Original	0.736	0.986	0.377	0.000	0.483	1.000	1.000	
$q_0^+$	-0.258	-0.121	0.305	-0.916	-0.429	-0.011	0.000	34.70%
$q_0^-$	0.154	0.041	0.238	0.000	0.005	0.188	0.760	37.80%
$q_0^{e+}$	-0.053	-0.031	0.059	-0.172	-0.087	-0.004	0.000	40.35%
$q_0^{e-}$	0.049	0.026	0.056	0.000	0.004	0.079	0.174	41.70%
$\lambda^+$	-0.077	-0.022	0.142	-0.347	-0.111	-0.001	0.034	37.50%
$\lambda^-$	0.063	0.014	0.137	-0.065	0.001	0.085	0.343	39.00%
$\kappa^+$	0.039	0.015	0.058	0.000	0.002	0.052	0.160	43.55%
$\kappa^-$	-0.082	-0.038	0.103	-0.296	-0.129	-0.004	0.000	40.65%
$c^+$	-0.237	-0.195	0.213	-0.614	-0.400	-0.034	0.000	25.30%
$c^-$	0.194	0.098	0.220	0.000	0.007	0.357	0.598	37.65%
$\beta^+$	0.123	0.067	0.144	0.000	0.006	0.211	0.377	38.65%
$\beta^-$	-0.123	-0.084	0.132	-0.367	-0.202	-0.010	0.000	30.90%
$\rho_1^+$	0.129	0.071	0.151	0.000	0.006	0.211	0.417	38.70%
$\rho_1^-$	-0.123	-0.078	0.138	-0.395	-0.191	-0.010	0.000	29.85%
$\rho_0^+$	0.096	0.038	0.137	0.000	0.005	0.134	0.372	40.25%
$\rho_0^-$	-0.072	-0.029	0.107	-0.285	-0.097	-0.004	0.000	41.00%

Figure 1.6: Marginal effects of structural parameters



This figure plots the marginal effects of selected model parameters (see section ?? for the definition of the parameters). I simulate 2,000 firms and 12 years for 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 forecast 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 percentage points. The black dots represent the mean effects.

finding is consistent with 57% of firms in the sample always issue management forecasts from 2003 to 2014.

Next, conditional on the marginal effects of investor learning being non-zero, Table 1.7 shows that investor learning has economically significant effects on the average forecast frequency both in absolute terms and relative to other determinants of forecast issuance specified by the model. A one standard deviation change in investors' prior beliefs changes future forecast frequency by an average of 15-25 percentage points, about one standard deviation of the empirical forecast frequency (22 percentage points). As discussed before, a one standard deviation change in investors' prior beliefs might be too large (30 percentage points) relative to the standard deviation of earnings (17 percentage points). The third row presents the change in forecast frequency associated with 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 forecast frequency (22 percentage points). Finally, about 5% of firms have effects larger than 17 percentage points, about 0.77 standard deviations of the empirical disclosure frequency (22 percentage points).

The precision of investors' beliefs affect the average forecast frequency with a similar magnitude. The average marginal effect for a one standard deviation change in investors' belief precision is about 6-8 percentage points or about 0.32 standard deviations of the empirical forecast frequency. For 5% of firms, the marginal effects are larger than 35 percentage points or about 1.60 standard deviations of the empirical forecast frequency.

To benchmark the magnitude of the effects of investor learning, I summarize the marginal effects of firm risk, earnings persistence, firm profitability, managers' information precision, and disclosure cost in Table 1.7. 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, earnings persistence and firm risk are associated with an average marginal effect of 12 percentage points—larger than

those of investor learning.<sup>18</sup> 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.

The estimation is the first attempt to estimate how investor learning affects managers' disclosure decisions in an internally consistent framework that incorporates other determinants of disclosure. Increases in investors' beliefs lower the likelihood of disclosure, because there is less room for the manager to convey news better than investors' expectation. Increases in investors' belief precision also lower the likelihood of disclosure, because investors put smaller weight to a given disclosure, which reduces the “importance” of disclosure.<sup>19</sup> 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. So the estimated marginal effects using a 12-year window are conservative. The results illustrate that investors' initial perception about the firm has a meaningful impact on managers' disclosure decisions over a reasonably long time (12 years), holding constant other firm characteristics specified in the model.

### *1.6.5 Robustness*

#### News not specified by the model

So far I focus on firm annual financial reports and management forecasts of annual earnings as the only information sources investors rely on. However, investors also learn from management quarterly forecasts, quarterly financial reports, analyst forecasts, media, and other information sources, which the model does not specify. I now discuss how the existence of

---

18. 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 a given disclosure, which decreases the manager's disclosure incentive.

19. See the Appendix for the model derivation and comparative statistics.

news not specified by the model affects the empirical results.

I assume that unspecified news  $\varepsilon$  satisfies two properties. First, unspecified news does not invalidate an AR1 earnings process, which ensures the model tractability. Investors' expectation of the forthcoming earnings  $y_t$  at time  $t - \Delta$  ( $\Delta \in (0, 1)$ ) can be expressed as:

$$E(y_t | \mathcal{I}_{t-\Delta}) = q_{0t-\Delta} + \rho_1 y_{t-1},$$

where  $\mathcal{I}_{t-\Delta}$  is investors' information set at  $t - \Delta$  and  $q_{0t-\Delta}$  is investors' belief about  $\rho_0$ . Second, I assume that unspecified news is unexpected by investors, and therefore has a zero mean.

With the properties above, assuming for now that investors' information precision  $\lambda_t$  does not change with unspecified news, the new disclosure threshold denoted by  $s_{t1}^*$  can be shown to be:

$$s_{t1}^* = s_t^* + \varepsilon_t, \tag{1.15}$$

where  $s_t^*$  is the original disclosure threshold assuming no unspecified news and  $\varepsilon_t$  is the unspecified news.<sup>20</sup> Equation (1.15) shows that the expected bias in estimating the disclosure threshold is zero.

I now discuss the possibility that the unspecified news changes investors' information precision  $\lambda_t$ . To the extent that having access to the unspecified news increases the precision of investors' beliefs, the precision parameter is likely to be over-estimated relative to its true value, that is, the  $\lambda_t$  that can explain the pattern of stock prices in the data is higher than the true  $\lambda_t$  that the model is meant to capture. To see this, based on the model, the precision of investors' beliefs after observing management forecast evolves according to:

$$\lambda_{it}^s = \frac{\kappa}{\kappa + 1} + \lambda_{it},$$

---

20. The additive separability of  $\varepsilon_t$  follows from the linearity of the disclosure threshold with respect to investors' beliefs as shown in equation (1.10). That is, changes in investors' beliefs linearly alter the disclosure threshold.

where  $\lambda_{it}$  measures the precision of investors' beliefs at the beginning of year  $t$  (right after the previous year earnings announcement) and  $\kappa$  measures the precision of firm  $i$  manager's earnings forecast. Manager's precision is estimated from the difference between management forecasts and actual earnings, and is therefore not affected by the existence of additional news. The additional precision as a result of other information sources will show up in the estimate of  $\lambda_{it}$ .

Given the unspecified news, the estimated parameter should be interpreted as the precision prior to the release of a management forecast or prior to investors realizing the lack of a disclosure. The problem with this interpretation is that it involves the timing of management forecasts, which can differ across firms and over time and could vary with the quantity and quality of the unspecified news. Because I do not model the timing of management forecasts, I leave it as a future research possibility to examine this issue.

## 1.7 Relation to prior literature

### *1.7.1 Reputation accumulation, investor learning, and management forecasts*

Prior research suggests that management accumulates reputation through voluntary disclosure. The definitions of reputation include the accuracy of management forecasts [85, 50, 53], being transparent [45], being forthcoming [17], being uninformed [32], and the likelihood of bad news [64]. Note that investors' beliefs about firm profitability can be viewed as a type of reputation by itself.

I capture reputation for being accurate in two ways. First, I capture the accuracy of management forecasts through  $\kappa_i$ , the ratio of the variance of actual earnings to the variance of management forecast. Second, I capture the effects of the number of past disclosures through the precision of investors' beliefs, where a higher number of past disclosures is associated with a higher precision of investors' beliefs. The accuracy of management forecasts and the

precision of investors' beliefs both have economically meaningful effects on the likelihood of a forecast provision, as shown by Figure 1.6 and Table 1.7. The results are consistent with several prior studies. For example, Feng and Koch [34] show that management is more likely to omit a forecast issuance, if management forecast of the prior quarter was not able to influence analyst forecasts. Williams [85] and Hutton and Stocken [53] find that analyst revisions and market reactions to management forecasts increase with the precision of management forecast.

I capture reputation for being transparent through its effect on cost of capital. Graham et al. [45] argue that voluntary disclosure helps management to establish a reputation for being transparent. Empirical research associates the benefit of transparency with a lower cost of capital (see Beyer et al. [16] for a review of recent literature). I capture this effect using the discount factor  $\beta_i$ . Figure 1.6 and Table 1.7 suggests that a lower cost of capital is associated with more disclosure, which is consistent with findings of prior research.

A few theory models examine disclosure decisions featuring dynamic effects of reputation [e.g., 32, 17, 64]. Investor learning occurs because of their uncertainty regarding profitability and is distinct from other dynamic effects. However, it is possible that investor learning interacts with these dynamic effects. For example, investors might be more responsive to a disclosure if they deem management to be more forthcoming. Incorporating other dynamic effects requires modeling the interactions between these effects and investor learning, which is beyond the scope of this paper.

A final note is that the paper deliberately stays away from modeling commitment to disclosure. Because once commitment is assumed, it automatically generates persistent disclosure incentives. The purpose of the paper is to show that correlated investors' beliefs over time can result in persistent disclosure incentives without commitment to disclosure.



### 1.7.2 *Investor learning and analyst earnings forecasts*

This section discusses the relation between analyst forecasts and investors' beliefs. Analysts are important market participants and learn from managers' disclosures. Prior research finds that analyst forecasts affect management earnings forecasts in two ways. First, optimistic analyst forecasts are associated with fewer management earnings forecast provision [e.g., 27, 34]. Second, managers tend to issue more bad news forecasts when they have consistently met or beaten analyst forecasts than when they have not [60]. In both cases, earnings forecast provision is a tool to guide analyst forecasts downwards to avoid the cost of disappointing the market at earnings announcements [56].

Investors' beliefs as defined in this paper differ from analyst forecasts in two important ways. First, investor beliefs about profitability can be different from the actual profitability, because investors have access to limited historical information that is relevant for learning the underlying firm profitability. The difference can persist over multiple years, as learning occurs gradually [70].<sup>21</sup> In contrast, although analysts also form their beliefs about firm profitability, their reported forecasts can be affected by their affiliation with brokerage firm [63] or their trading incentives [28].

Second, investors and analysts affect managers' disclosure decisions through different channels, resulting in different directional predictions. Higher investor beliefs increase the threshold of disclosure (see section 1.2), reducing managers' disclosure incentive, because managers have less room to convey good news. In contrast, management facing optimistic analyst forecasts are more likely to disclose, because managers want to lower analyst forecasts towards a beatable target. Moreover, the disclosure incentive arising from optimistic analyst forecasts focuses on how investors respond to subsequent announced earnings (that is, meet or beat analyst forecasts at the subsequent earnings announcement) but does not address how investors are likely to respond to the current voluntary disclosure, which is the channel

---

21. Seminal work on parameter uncertainty builds on the same idea. For example, Xia [86] shows that investors' initial assessment of return predictability differs from the actual return predictability, leading to changes in their portfolio allocation as they learn the actual return predictability.

of interest in this paper.

To provide empirical evidence on the relation between analyst forecasts and investors' beliefs, I regress investors' beliefs at the beginning of the sample period on analyst earnings forecasts up until two years after the measurement of investors' beliefs. I use investors' beliefs at the beginning of the sample period as the dependent variable instead of investors' subsequent beliefs, because investors' initial beliefs are estimated as parameters, whereas realized earnings "contaminate" subsequent investors' beliefs via equation (1.4), introducing a mechanical relation between analyst earnings forecasts and investors' subsequent beliefs. I scale analyst earnings forecasts, the independent variable of interest, by the stock price at the beginning of the sample period for two reasons. First, investors' beliefs are also scaled by the stock price at the beginning of the sample period. Second, because investors' beliefs are forward-looking, future stock prices will reflect investors' current beliefs and future earnings. Scaling future analyst earnings forecasts by future prices induces a mechanical relation between the scaled analyst earnings forecasts and investors' current beliefs via future prices.

Table 1.8 presents the results. I find that investors' beliefs and analyst forecasts are positively related. The finding is robust to controlling for size, performance (ROE), investor uncertainty (using implied option volatility), and market to book ratio, and is robust to using analyst earnings forecasts from the entire sample rather than the first two years. The positive but imperfect relation between analyst forecasts and investors' beliefs suggests that analyst forecasts affect disclosure decisions in two ways. Suppose analyst forecasts contain both their assessment of future profitability and their strategic incentive. This paper suggests that analysts' assessment of firm long run profitability (the learning part) negatively affects the likelihood of a disclosure, whereas prior research shows that incentives to provide optimistic forecasts (the strategic part) positively affect the likelihood of a disclosure. Future research can attempt to separately quantify the importance of the two channels by imposing structure on analyst forecasts.

Table 1.8: Investors' beliefs and analyst forecasts

This table examines the relation between investors' beliefs estimated from the structural model and analyst earnings forecasts. The dependent variable is the median investor beliefs of each firm at the beginning of the sample period. The independent variables include: analyst forecasts, measured as median analyst forecasts for the earnings of the subsequent fiscal year issued within 30 days after the earnings announcement of the current fiscal year, scaled by the stock price at the beginning of the sample period; ROE, measured as earnings per share divided by the stock price at the beginning of the sample period; the natural logarithm of firm total assets; investor uncertainty, measured as the average implied option volatility 30 days following the earnings announcement, divided by the stock price at the beginning of the sample period; market to book ratio. The first three columns use observations from the first two years. The last three columns use the entire sample. All standard errors are clustered at the firm level and reported in the brackets.

VARIABLES	(1) $t \leq 2$	(2) $t \leq 2$	(3) $t \leq 2$	(4) All	(5) All	(6) All
Analyst forecasts	3.088 [2.233]	4.675* [2.510]	6.079* [3.557]	1.161*** [0.414]	2.629*** [0.427]	2.349*** [0.685]
ROE		-1.269 [1.088]	-3.597*** [1.202]		-1.483*** [0.393]	-2.017*** [0.510]
Log(Assets)	-0.195*** [0.042]	-0.206*** [0.040]	-0.090* [0.054]	-0.171*** [0.029]	-0.178*** [0.029]	-0.117*** [0.038]
Investor uncertainty			7.552*** [2.285]			7.207*** [1.901]
Market to book			-0.005 [0.016]			-0.011 [0.012]
Observations	348	348	175	2,883	2,699	1,531
$R^2$	0.100	0.111	0.174	0.086	0.114	0.168

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

## 1.8 Conclusion

I model and empirically quantify the effects of investor learning on the intertemporal dynamics of managers' voluntary disclosure decisions. More pessimistic investor prior beliefs and higher investor uncertainty increase disclosure and that investor learning induces persistence in the likelihood of disclosure. For 60% of firms, 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 7.5%.

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 [69] 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. [74] 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 is incorporating other reporting incentives. For instance, managers could have incentives to bias disclosures [e.g., 36]. 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., 34]. How this incentive arises as an equilibrium outcome and how it affects the dynamics of

disclosures are both interesting questions.

Future research can also 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 can rely on observable data to identify them by imposing various model structures, that is, by structurally estimating the models.

## REFERENCES

- [1] Abarbanell, J. S., Lanen, W. N., Verrecchia, R. E., 1995. Analysts' forecasts as proxies for investor beliefs in empirical research. *Journal of Accounting and Economics* 20, 31–60.
- [2] Aboody, D., Kasznik, R., 2000. CEO stock option awards and the timing of corporate voluntary disclosures. *Journal of Accounting and Economics* 29, 73–100.
- [3] Acharya, V. V., DeMarzo, P., Kremer, I., 2011. Endogenous information flows and the clustering of announcements. *The American Economic Review* 101, 2955–2979.
- [4] Ajinkya, B., Bhojraj, S., Sengupta, P., 2005. The association between outside directors, institutional investors and the properties of management earnings forecasts. *Journal of Accounting Research* 43, 343–376.
- [5] Ajinkya, B. B., Gift, M. J., 1984. Corporate managers' earnings forecasts and symmetrical adjustments of market expectations. *Journal of Accounting Research* 22, 425–444.
- [6] Albrecht, W. S., Lookabill, L. L., McKeown, J. C., 1977. The time-series properties of annual earnings. *Journal of Accounting Research* 15, 226–244.
- [7] Anilowski, C., Feng, M., Skinner, D. J., 2007. Does earnings guidance affect market returns? The nature and information content of aggregate earnings guidance. *Journal of Accounting and Economics* 44, 36–63.
- [8] Ball, R., Shivakumar, L., 2008. How much new information is there in earnings? *Journal of Accounting Research* 46, 975–1016.
- [9] Ball, R., Watts, R., 1972. Some time series properties of accounting income. *The Journal of Finance* 27, 663–681.
- [10] Barron, O. E., Kim, O., Lim, S. C., Stevens, D. E., 1998. Using analysts' forecasts to measure properties of analysts' information environment. *The Accounting Review* 73, 421–433.
- [11] Barry, C. B., Brown, S. J., 1985. Differential information and security market equilibrium. *Journal of Financial and Quantitative Analysis* 20, 407–422.
- [12] Berger, P. G., Hann, R., 2003. The impact of SFAS No. 131 on information and monitoring. *Journal of Accounting Research* 41, 163–223.
- [13] Bergman, N. K., Roychowdhury, S., 2008. Investor sentiment and corporate disclosure. *Journal of Accounting Research* 46, 1057–1083.
- [14] Bertomeu, J., Beyer, A., Taylor, D. J., 2015. From Casual to Causal Inference in Accounting Research: The Need for Theoretical Foundations. Working paper, available at SSRN 2694105.

- [15] Bertomeu, J., Ma, P., Marinovic, I., 2015. How often do managers withhold information? Working paper, Baruch College, City University of New York.
- [16] Beyer, A., Cohen, D. A., Lys, T. Z., Walther, B. R., 2010. The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics* 50, 296–343.
- [17] Beyer, A., Dye, R. A., 2012. Reputation management and the disclosure of earnings forecasts. *Review of Accounting Studies* 17, 877–912.
- [18] Bhojraj, S., Libby, R., Yang, H., 2011. Guidance frequency and guidance properties: The effect of reputation-building and learning-by-doing. Working paper, Johnson School Research Paper Series.
- [19] Billings, M. B., Jennings, R. H., Lev, B., 2014. On guidance and volatility. Working paper, Kelley School of Business, Indiana University.
- [20] Brown, S., 1979. The effect of estimation risk on capital market equilibrium. *Journal of Financial and Quantitative Analysis* 14, 215–220.
- [21] Bushee, B. J., Noe, C. F., 2000. Corporate disclosure practices, institutional investors, and stock return volatility. *Journal of Accounting Research* 38, 171–202.
- [22] Chen, Q., Francis, J., Jiang, W., 2005. Investor learning about analyst predictive ability. *Journal of Accounting and Economics* 39, 3–24.
- [23] Chen, S., Matsumoto, D., Rajgopal, S., 2011. Is silence golden? An empirical analysis of firms that stop giving quarterly earnings guidance. *Journal of Accounting and Economics* 51, 134–150.
- [24] Chuk, E., Matsumoto, D., Miller, G. S., 2013. Assessing methods of identifying management forecasts: CIG vs. researcher collected. *Journal of Accounting and Economics* 55, 23–42.
- [25] Coles, J. L., Loewenstein, U., 1988. Equilibrium pricing and portfolio composition in the presence of uncertain parameters. *Journal of Financial Economics* 22, 279–303.
- [26] Coles, J. L., Loewenstein, U., Suay, J., 1995. On equilibrium pricing under parameter uncertainty. *Journal of Financial and Quantitative Analysis* 30, 347–364.
- [27] Cotter, J., Tuna, A., Wysocki, P. D., 2006. Expectations management and beatable targets: How do analysts react to explicit earnings guidance? *Contemporary Accounting Research* 23.
- [28] Cowen, A., Groyberg, B., Healy, P., 2006. Which types of analyst firms are more optimistic? *Journal of Accounting and Economics* 41, 119–146.
- [29] Dichev, I. D., Tang, V. W., 2009. Earnings volatility and earnings predictability. *Journal of Accounting and Economics* 47, 160–181.

- [30] Dye, R. A., 1985. Disclosure of nonproprietary information. *Journal of Accounting Research* 23, 123–145.
- [31] Easton, P. D., 2004. PE ratios, PEG ratios, and estimating the implied expected rate of return on equity capital. *The Accounting Review* 79, 73–95.
- [32] Einhorn, E., Ziv, A., 2008. Intertemporal dynamics of corporate voluntary disclosures. *Journal of Accounting Research* 46, 567–589.
- [33] Erdem, T., Keane, M. P., 1996. Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science* 15, 1–20.
- [34] Feng, M., Koch, A. S., 2010. Once bitten, twice shy: The relation between outcomes of earnings guidance and management guidance strategy. *The Accounting Review* 85, 1951–1984.
- [35] Field, L., Lowry, M., Shu, S., 2005. Does disclosure deter or trigger litigation? *Journal of Accounting and Economics* 39, 487–507.
- [36] Fischer, P. E., Verrecchia, R. E., 2000. Reporting bias. *The Accounting Review* 75, 229–245.
- [37] Francis, J., Philbrick, D., Schipper, K., 1994. Shareholder litigation and corporate disclosures. *Journal of Accounting Research* 32, 137–164.
- [38] Frankel, R., Litov, L., 2009. Earnings persistence. *Journal of Accounting and Economics* 47, 182–190.
- [39] Frankel, R., McNichols, M., Wilson, G. P., 1995. Discretionary disclosure and external financing. *The Accounting Review* 70, 135–150.
- [40] Gao, P., 2013. A measurement approach to conservatism and earnings management. *Journal of Accounting and Economics* 55, 251–268.
- [41] Gao, P., Liang, P. J., 2013. Informational feedback, adverse selection, and optimal disclosure policy. *Journal of Accounting Research* 51, 1133–1158.
- [42] Gigler, F., 1994. Self-enforcing voluntary disclosures. *Journal of Accounting Research* 86, 224–240.
- [43] Gigler, F., Hemmer, T., 1998. On the frequency, quality, and informational role of mandatory financial reports. *Journal of Accounting Research* 36, 117–147.
- [44] Gow, I., Larcker, D., Reiss, P., 2015. Causal inferences in accounting research. Working paper, Stanford University.
- [45] Graham, J. R., Harvey, C. R., Rajgopal, S., 2005. The economic implications of corporate financial reporting. *Journal of Accounting and Economics* 40, 3–73.



- [46] Grossman, S. J., 1981. The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics* 24, 461–483.
- [47] Guay, W. R., Samuels, D., Taylor, D. J., 2015. Guiding through the fog: Financial statement complexity and voluntary disclosure. Working paper, available at SSRN 2564350.
- [48] Guttman, I., Kremer, I., Skrzypacz, A., 2014. Not only what but also when: A theory of dynamic voluntary disclosure. *The American Economic Review* 104, 2400–2420.
- [49] Heckman, J. J., Urzua, S., 2010. Comparing IV with structural models: What simple IV can and cannot identify. *Journal of Econometrics* 156, 27–37.
- [50] Hirst, D. E., Koonce, L., Venkataraman, S., 2008. Management earnings forecasts: A review and framework. *Accounting Horizons* 22, 315–338.
- [51] Houston, J. F., Lev, B., Tucker, J. W., 2010. To guide or not to guide? Causes and consequences of stopping quarterly earnings guidance. *Contemporary Accounting Research* 27, 143–185.
- [52] Hutton, A. P., Miller, G. S., Skinner, D. J., 2003. The role of supplementary statements with management earnings forecasts. *Journal of Accounting Research* 41, 867–890.
- [53] Hutton, A. P., Stocken, P. C., 2009. Prior forecasting accuracy and investor reaction to management earnings forecasts. Working paper, available at SSRN 817108.
- [54] Jung, W. O., Kwon, Y. K., 1988. Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting Research* 26, 146–153.
- [55] Kasznik, R., Lev, B., 1995. To warn or not to warn: Management disclosures in the face of an earnings surprise. *The Accounting Review* 70, 113–134.
- [56] Kasznik, R., McNichols, M. F., 2002. Does meeting earnings expectations matter? Evidence from analyst forecast revisions and share prices. *Journal of Accounting Research* 40, 727–759.
- [57] Kim, O., Verrecchia, R. E., 1991. Market reaction to anticipated announcements. *Journal of Financial Economics* 30, 273–309.
- [58] Kim, O., Verrecchia, R. E., 1991. Trading volume and price reactions to public announcements. *Journal of Accounting Research* 29, 302–321.
- [59] Kothari, S. P., Shu, S., Wysocki, P. D., 2009. Do managers withhold bad news? *Journal of Accounting Research* 47, 241–276.
- [60] Kross, W. J., Ro, B. T., Suk, I., 2011. Consistency in meeting or beating earnings expectations and management earnings forecasts. *Journal of Accounting and Economics* 51, 37–57.
- [61] Lang, M., 1991. Time-varying stock price response to earnings induced by uncertainty about the time-series process of earnings. *Journal of Accounting Research* 29, 229–257.

- [62] Lennox, C. S., Park, C. W., 2006. The informativeness of earnings and management's issuance of earnings forecasts. *Journal of Accounting and Economics* 42, 439–458.
- [63] Lim, T., 2001. Rationality and analysts' forecast bias. *The Journal of Finance* 56, 369–385.
- [64] Marinovic, I., Varas, F., 2014. No news is good news: Voluntary disclosure in the face of litigation. Rock Center for Corporate Governance at Stanford University Working Paper 169, 13–19.
- [65] Milgrom, P. R., 1981. Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics* 12, 380–391.
- [66] Miller, G. S., 2002. Earnings performance and discretionary disclosure. *Journal of Accounting Research* 40, 173–204.
- [67] Nagar, V., 1999. The role of the manager's human capital in discretionary disclosure. *Journal of Accounting Research* 37, 167–181.
- [68] Ng, J., Tuna, I., Verdi, R., 2013. Management forecast credibility and underreaction to news. *Review of Accounting Studies* 18, 956–986.
- [69] Pástor, L., Stambaugh, R. F., 2012. Are stocks really less volatile in the long run? *The Journal of Finance* 67, 431–478.
- [70] Pástor, L., Veronesi, P., 2003. Stock valuation and learning about profitability. *The Journal of Finance* 58, 1749–1789.
- [71] Patton, A. J., Verardo, M., 2012. Does beta move with news? Firm-specific information flows and learning about profitability. *Review of Financial Studies* 25, 2789–2839.
- [72] Pownall, G., Waymire, G., 1989. Voluntary disclosure credibility and securities prices: Evidence from management earnings forecasts, 1969–73. *Journal of Accounting Research* 27, 227–245.
- [73] Richardson, S. A., Sloan, R. G., Soliman, M. T., Tuna, I., 2005. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics* 39, 437–485.
- [74] Rogers, J. L., Skinner, D. J., Van Buskirk, A., 2009. Earnings guidance and market uncertainty. *Journal of Accounting and Economics* 48, 90–109.
- [75] Skinner, D. J., 1994. Why firms voluntarily disclose bad news. *Journal of Accounting Research* 32, 38–60.
- [76] Sletten, E., 2012. The effect of stock price on discretionary disclosure. *Review of Accounting Studies* 17, 96–133.
- [77] Stein, J. C., 1989. Efficient capital markets, inefficient firms: A model of myopic corporate behavior. *The Quarterly Journal of Economics* 104, 655–669.

- [78] Tang, M. M., 2012. What guides the guidance? An empirical examination of the dynamic disclosure theory. Working paper, University of Rochester.
- [79] Taylor, L. A., 2010. Why are CEOs rarely fired? Evidence from structural estimation. *The Journal of Finance* 65, 2051–2087.
- [80] Trueman, B., 1986. Why do managers voluntarily release earnings forecasts? *Journal of Accounting and Economics* 8, 53–71.
- [81] Verrecchia, R. E., 1983. Discretionary disclosure. *Journal of Accounting and Economics* 5, 179–194.
- [82] Verrecchia, R. E., 2001. Essays on disclosure. *Journal of Accounting and Economics* 32, 97–180.
- [83] Watts, R. L., Leftwich, R. W., 1977. The time series of annual accounting earnings. *Journal of Accounting Research* 15, 253–271.
- [84] Waymire, G., 1985. Earnings volatility and voluntary management forecast disclosure. *Journal of Accounting Research* 23, 268–295.
- [85] Williams, P. A., 1996. The relation between a prior earnings forecast by management and analyst response to a current management forecast. *Accounting Review* 71, 103–115.
- [86] Xia, Y., 2001. Learning about predictability: The effects of parameter uncertainty on dynamic asset allocation. *The Journal of Finance* 56, 205–246.

# APPENDIX A

## MODEL

### A.1 Discussions of model assumptions

Prior literature identifies other factors that prevent full disclosure, such as entry concerns [42] and probabilistic information endowment [30, 54].<sup>1</sup> In multi-period models, incentives such as building a reputation for being forthcoming [17] or being uninformed [32] can also arise. Abstracting away from these allows me to focus on how investor learning affects the trade-off between disclosure and nondisclosure.

Another important assumption is truth telling, which is consistent with prior research [81, 30]. 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., 43]. Stein [77] 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.

Recall that the manager's current-period utility derived from disclosure decision  $d_{it}$  is

$$U_{it}(d_{it}) = P^c(d_{it}, \mathcal{H}_{it}), \tag{A.1}$$

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,  $\mathcal{H}_{it} \equiv (\mathbf{y}'_{i,t-1}, \mathbf{d}'_{i,t-1})'$ , where

---

1. See Beyer et al. [16] for a review of voluntary disclosure models.

$\mathbf{y}_{i,t-1} \equiv (y_{i1}, y_{i2}, \dots, y_{i,t-1})'$  and  $\mathbf{d}_{i,t-1} \equiv (d_{i1}, d_{i2}, \dots, d_{i,t-1})'$ .

This utility specification says the manager is rewarded for a higher stock price, which many voluntary disclosure models use [e.g., 81, 30, 3]. Stock price is the sum of investors' expectations of discounted unrealized earnings net of current and future disclosure costs,

$$P^c(d_{it}, \mathcal{H}_{it}) = \sum_{t'=t}^T \left\{ \beta_i^{t'-t} \mathbb{E}(y_{it'} - \mathbb{1}\{d_{it'} = s_{it'}\}c_i | d_{it}, \mathcal{H}_{it}) \right\}, \quad (\text{A.2})$$

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.<sup>2</sup> It follows that the disclosure decision of period  $t$  will not affect investors' beliefs or stock prices beyond period  $t$ .<sup>3</sup> 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., 77]. It simply refers to the fact that the manager maximizes her current utility when making the disclosure decision.

Modeling investor learning assumes disclosure is only informative about firm earnings. Trueman [80] 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

---

2. This result follows from  $s_{it} = y_{it} + \eta_{it}$ , where  $\eta_{it}$  is random noise. See equation (1.4) and the appendix for derivation.

3. As in Verrecchia [81], 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.

both firm fundamentals and the manager’s ability. The choice is consistent with management forecasts being mainly informative about firm fundamentals [8].

Voluntary disclosure accelerates the release of earnings news. Investors interpret nondisclosure 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.<sup>4</sup> Consistent with this assumption, Chen et al. [23] 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.<sup>5</sup> Finally, disclosures impose costs on investors, and are inefficient from an ex-ante perspective [82]. 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.

## A.2 The effects of investor learning

The effects of investors’ beliefs and uncertainty on the probability of disclosure directly follow from equation (1.11), and are summarized in Corollary 1.

**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 in-

---

4. 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.

5. Acharya et al. [3] and Guttman et al. [48] 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.

crease 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 (1.10), increase in  $y_{t-1}$  also enhances the disclosure threshold. The net effect is that they cancel each other. Thus the disclosure probability in (1.11) does not depend on  $\rho_1 y_{t-1}$ . I discuss the intuition in greater detail in the next section.

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 (1.11)), which in turn depend on their beliefs in period  $t$  (see (1.4)). 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 section.

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

*Investor learning induces persistent disclosure incentives, that is, the ex-ante correlation between the likelihood of disclosure in period  $t$  and  $t + m$  is positive, where the ex-ante correlation is defined as the correlation between  $Pr\{d_t = s_t | \mathcal{H}_{t-1}\}$  and  $Pr\{d_{t+m} = s_{t+m} | \mathcal{H}_{t-1}\}$  and  $\mathcal{H}_{t-1}$  is the history of information observed by investors at the beginning of  $t - 1$ .*

*Proof.* Equation (1.11) gives the probability of disclosure conditional on history  $\mathcal{H}_t$ ,

$$\Pr\{d_t = s_t | \mathcal{H}_t\} = 1 - \Phi\left(\frac{\sigma_{st}\tilde{s}_t^* + q_{0t} - \rho_0}{\sigma_s}\right).$$

The probability  $\Pr\{d_t = s_t | \mathcal{H}_{t-1}\}$  is random because of unexpected performance shocks that could change investors' future beliefs  $q_{0t}$ .<sup>6</sup> Specifically,  $q_{0t}$  can be expressed as a function of  $q_{0t-1}$  and random future performance shock  $\epsilon_{t-1}$  (see equation (1.4)):

$$q_{0t} = (\lambda_t)^{-1}(\lambda_{t-1}q_{0t-1} + y_{t-1} - \rho_1 y_{t-2}) = (\lambda_t)^{-1}(\lambda_{t-1}q_{0t-1} + \rho_0 + \epsilon_{t-1}).$$

Likewise,  $q_{0t+m}$  can be iteratively expressed as a function of  $q_{0t-1}$  and random future performance shocks  $\epsilon_{t-1}, \epsilon_t, \dots, \epsilon_{t+m-1}$ .

Although performance shocks are independent over time, because both  $q_{0t}$  and  $q_{0t+m}$  contain  $\epsilon_{t-1}$ , investors' beliefs are positively intertemporally correlated. Therefore,  $\Pr\{d_t = s_t | \mathcal{H}_{t-1}\}$  and  $\Pr\{d_{t+m} = s_{t+m} | \mathcal{H}_{t-1}\}$  is positively correlated (because the normal CDF is increasing over its support).

□

### A.3 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 for later and only discuss the intuition for why these parameters affect

---

6. The Appendix proves that  $\tilde{s}_t^*$  does not depend on  $q_{0t}$  and is a function of non-random firm characteristics and investor uncertainty  $\frac{1}{\lambda_t}$ .



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 (??)), with higher sensitivity increasing disclosure. This is analogous to higher earnings response coefficient (ERC). For example, Lennox and Park [62] 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.

## A.4 Proofs

**Lemma 1.** *Call  $h(s) \equiv \mathbb{E}(\tilde{s}|\tilde{s} \leq s)$ , 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.*

$$\begin{aligned} h'(s) &= -\frac{\phi'(s)\Phi(s) - \phi^2(s)}{\Phi^2(s)} \\ &= \frac{s\phi(s)\Phi(s) + \phi^2(s)}{\Phi^2(s)}. \end{aligned}$$

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

$$\tilde{s}_t^* - \mathbb{E}(\tilde{s}_t | \tilde{s}_t \leq \tilde{s}_t^*) = c\sqrt{\tau}. \quad (\text{A.3})$$

## A.5 Bayesian updating

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

$$y_{t+1} = \rho_0 + \rho_1 y_t + \epsilon_{t+1}, \quad \epsilon_{t+1} \sim \mathcal{N}\left(0, \frac{1}{\tau}\right),$$

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  $s_t \sim \mathcal{N}(y_t, \frac{1}{\kappa \tau})$  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  $\mathcal{H}_t \equiv (\mathbf{s}'_{t-1}, \mathbf{x}'_{t-1}, \mathbf{y}'_{t-1})$ ,  $\mathbf{y}_{t-1} \equiv (y_{t-1}, y_{t-2}, \dots, y_1)'$ ,  $\mathbf{x}_{t-1} \equiv (y_{t-2}, y_{t-2}, \dots, y_0)'$  and  $\mathbf{s}_{t-1} \equiv (s_{t-1}, s_{t-2}, \dots, s_1)'$ . 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.5.1 Learning $\rho$

The joint density of  $\mathbf{y}_{t-1}, \mathbf{x}_{t-1}, \mathbf{s}_{t-1}, \rho$  is

$$f(\mathbf{y}_{t-1}, \mathbf{x}_{t-1}, \mathbf{s}_{t-1}, \rho) \propto e^{-\frac{\tau}{2} \left\{ (\mathbf{y}_{t-1} - \mathbf{X}_{t-1} \boldsymbol{\rho})' (\mathbf{y}_{t-1} - \mathbf{X}_{t-1} \boldsymbol{\rho}) \right\}} \\ * e^{\left\{ \kappa (\mathbf{s}_{t-1} - \mathbf{y}_{t-1})' (\mathbf{s}_{t-1} - \mathbf{y}_{t-1}) + \lambda_0 (\rho_0 - q_0)^2 \right\}},$$

where  $\mathbf{X}_t \equiv (\mathbf{1}, \mathbf{x}_t)$  and  $\boldsymbol{\rho} \equiv (\rho_0, \rho_1)'$ . The prior distribution of  $\rho_0$  is normal:

$$\rho_0 | \mathcal{H}_t \sim \mathcal{N} \left( \frac{(t-1)(\bar{y}_{t-1} - \rho_1 \bar{x}_{t-1}) + \lambda_0 q_0}{t-1 + \lambda_0}, \frac{1}{(t-1 + \lambda_0)\tau} \right). \quad (\text{A.4})$$

I denote the mean and precision (exclusive of  $\tau$ ) of the prior distribution of  $\rho_0$  after  $\mathcal{H}_t$  as:

$$\lambda_t = t - 1 + \lambda_0 = \lambda_{t-1} + 1; \quad (\text{A.5})$$

$$q_t = (\lambda_t)^{-1} \left( (t-1)(\bar{y}_{t-1} - \rho_1 \bar{x}_{t-1}) + \lambda_0 q_0 \right) = (\lambda_t)^{-1} (\lambda_{t-1} q_{t-1} + y_{t-1} - \rho_1 y_{t-2}). \quad (\text{A.6})$$

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  $\mathbf{y}_t, \mathbf{x}_t, \mathbf{s}_t, \rho_0$ :

$$\rho_0 | \mathcal{H}_t, s_t \sim \mathcal{N} \left( \left( \frac{\kappa}{\kappa+1} + \lambda_t \right)^{-1} \left( \frac{\kappa}{\kappa+1} (s_t - \rho_1 y_{t-1}) + \lambda_t q_t \right), \frac{1}{\left( \frac{\kappa}{\kappa+1} + \lambda_t \right) \tau} \right). \quad (\text{A.7})$$

I denote the mean and precision (exclusive of  $\tau$ ) of the posterior distribution of  $\rho_0$  after  $\mathcal{H}_t$  as

$$\lambda_t^s = \frac{\kappa}{\kappa+1} + \lambda_t; \quad (\text{A.8})$$

$$q_t^s = (\lambda_t^s)^{-1} \left( \frac{\kappa}{\kappa+1} (s_t - \rho_1 y_{t-1}) + \lambda_t q_t \right). \quad (\text{A.9})$$

### A.5.2 Prior predictive distribution of $y_t$ before observing $s_t$

In this section, I derive the prior predictive distribution of  $y_t$  given  $\mathcal{H}_t$  and before observing  $s_t$ . We have:

$$y_t | \mathcal{H}_t \sim \mathcal{N} \left( q_t + \rho_1 y_{t-1}, \frac{1}{\left( \frac{\lambda_t}{\lambda_t+1} \right) \tau} \right). \quad (\text{A.10})$$

I denote the mean and precision (exclusive of  $\tau$ ) as

$$\Lambda_t^y = \frac{\lambda_t}{\lambda_t + 1}; \quad (\text{A.11})$$

$$\mu_t^y = q_t + \rho_1 y_{t-1}. \quad (\text{A.12})$$

### A.5.3 Posterior predictive of $y_t$ given $s_t$

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

$$\begin{aligned} f(y_t|s_t, \mathcal{H}_t) &\propto f(s_t, y_t, \mathcal{H}_t) = f(s_t|y_t)f(y_t|\mathcal{H}_t) \\ &\propto e^{-\frac{\tau}{2} \left\{ \kappa(s_t - y_t)^2 + \Lambda_t^y (y_t - \mu_t^y)^2 \right\}}, \end{aligned}$$

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  $s_t$ :

$$y_t|s_t, \mathcal{H}_t \sim \mathcal{N} \left( \frac{\kappa s_t + \Lambda_t^y (q_t + \rho_1 y_{t-1})}{\kappa + \frac{\lambda_t}{\lambda_t + 1}}, \frac{1}{\tau \left( \kappa + \frac{\lambda_t}{\lambda_t + 1} \right)} \right). \quad (\text{A.13})$$

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

## A.6 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  $\mathcal{H}_t$ .

First, the expected profit  $m$  period after observing  $s_t$  in period  $t$  is

$$\begin{aligned} \mathbb{E}(y_{t+m}|s_t) &= \mathbb{E} \left( \mathbb{E}(y_{t+m} | \boldsymbol{\rho}, y_t) | s_t \right) \\ &= q_t^s \left( 1 + \rho_1 + \rho_1^2 + \dots + \rho_1^{m-1} \right) + \rho_1^m \frac{\kappa s_t + \Lambda_t^y \mu_t^y}{\kappa + \Lambda_t^y}. \end{aligned} \quad (\text{A.14})$$

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

$$\begin{aligned} & \sum_{m=1}^{\infty} \beta^m \left\{ q_t^s \left( 1 + \rho_1 + \rho_1^2 + \dots + \rho_1^{m-1} \right) + \rho_1^m \frac{\kappa s_t + \Lambda_t^y \mu_t^y}{\kappa + \Lambda_t^y} \right\} \\ &= \frac{\beta}{1 - \beta \rho_1} \left\{ \frac{\frac{\kappa}{\kappa+1} (s_t - \rho_1 y_{t-1}) + \lambda_t q_t}{\left( \frac{\kappa}{\kappa+1} + \lambda_t \right) (1 - \beta)} + \rho_1 \frac{\kappa s_t + \frac{\lambda_t}{\lambda_t+1} (q_t + \rho_1 y_{t-1})}{\kappa + \frac{\lambda_t}{\lambda_t+1}} \right\}, \end{aligned} \quad (\text{A.15})$$

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

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

$$P_t(s_t) = A_t s_t + B_t,$$

where

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

and

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

## A.7 Proof of Proposition 1

The optimal disclosure strategy would be a threshold strategy if the stock price did not include future expected disclosure cost  $\mathcal{G}_{d_t c}$ , because  $P(d_t, \mathcal{H}_t)$  is linear and increasing in the signal  $s_t$  (see equation (1.9)). 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,  $\mathcal{G}_{d_t=s_t} = \mathcal{G}_{d_t=\emptyset}$ .*

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.* Suppose there are  $T$  periods. Let  $\mathcal{G}_d^m$  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  $\mathcal{H}_T = (\mathcal{H}_{T-1}, s_{T-1}, y_{T-1})'$  is obtained from

$$\begin{aligned} \mathbb{E}(y_T | s_T^{**}, \mathcal{H}_T) - c &= \mathbb{E}(y_T | s_T \leq s_T^{**}, \mathcal{H}_T) \\ \Rightarrow \frac{\Sigma_T^0}{\kappa + \Sigma_T^0} &\left( (s_T^{**} - \mu_T^0) \sqrt{\frac{\kappa \Sigma_T^0}{\kappa + \Sigma_T^0}} \tau - \mathbb{E} \left( (s_T - \mu_T^0) \sqrt{\frac{\kappa \Sigma_T^0}{\kappa + \Sigma_T^0}} \tau \mid s_T \leq s_T^{**} \right) \right) \\ &= c \sqrt{\frac{\kappa \Sigma_T^0}{\kappa + \Sigma_T^0}} \tau. \end{aligned}$$

Note that I use investors' perceived distribution 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

$$\begin{aligned} \mathcal{G}_d^1 &= c \mathbb{E} \left[ \Pr \{ s_T \geq s_T^{**} | \mathcal{H}_T \} \right] \\ &= c \int \int \Pr \{ s_T \geq s_T^{**} | s_{T-1}, y_{T-1} \} f(s_{T-1}, y_{T-1} | d_{T-1}) dy_{T-1} ds_{T-1} \\ &= c \eta \left( \frac{\kappa \Sigma_T^0}{\kappa + \Sigma_T^0} \tau \right), \end{aligned} \tag{A.16}$$

with the final line following from the perceived probability of disclosure depending only on the precision. It immediately follows that  $\mathcal{G}_d^1$  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,  $\mathcal{G}_d^m$ , does not depend on  $d_{T-m}$ . I prove below that the expected disclosure cost  $m+1$  periods ahead,  $\mathcal{G}_d^{m+1}$ , does not depend on  $d_{T-m-1}$ .

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

$$\begin{aligned}
& \mathbb{E}(x_{T-m} | s_{T-m}^{**}, \mathcal{H}_{T-m}) + \sum_{t=T-m+1}^T \left\{ \beta^{t-(T-m)} x_t | s_{T-m}^{**}, \mathcal{H}_{T-m} \right\} - c - \beta \mathcal{G}_1^m \\
&= \mathbb{E}(x_{T-m} | s_{T-m} \leq s_{T-m}^{**}, \mathcal{H}_{T-m}) + \sum_{t=T-m+1}^T \left\{ \beta^{t-(T-m)} x_t | s_{T-m} \leq s_{T-m}^{**}, \mathcal{H}_{T-m} \right\} - \beta \mathcal{G}_0^m \\
&\Rightarrow \Gamma_{T-m} \left( (s_{T-m}^{**} - \mu_{T-m}^0) \sqrt{\frac{\kappa \Sigma_{T-m}^0}{\kappa + \Sigma_{T-m}^0}} \tau - \mathbb{E} \left( (s_{T-m} - \mu_{T-m}^0) \right. \right. \\
&\quad \left. \left. \sqrt{\frac{\kappa \Sigma_{T-m}^0}{\kappa + \Sigma_{T-m}^0}} \tau | s_{T-m} \leq s_{T-m}^{**} \right) \right) \\
&= c \sqrt{\frac{\kappa \Sigma_{T-m}^0}{\kappa + \Sigma_{T-m}^0}} \tau,
\end{aligned}$$

where  $\Gamma_{T-m}$  does not depend on  $\mathcal{H}_{T-m}$ .

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:

$$\begin{aligned}
\mathcal{G}_d^{m+1} &= cE \left[ Pr\{s_{T-m} \geq s_{T-m}^{**} | \mathcal{H}_{T-m}\} \right] + \beta \mathcal{G}^{m+1} \\
&= c \int \int Pr\{s_{T-m} \geq s_{T-m}^{**} | s_{T-m-1}, y_{T-m-1}, \mathcal{H}_{T-m-1}\} \\
&\quad f(s_{T-m-1}, y_{T-m-1} | d_{T-m-1}, \mathcal{H}_{T-m-1}) \\
&\quad dy_{T-m-1} ds_{T-m-1} + \beta \mathcal{G}^m \\
&= c\eta \left( \frac{\kappa \Sigma_{T-m}^0}{\kappa + \Sigma_{T-m}^0} \tau \right), \tag{A.17}
\end{aligned}$$

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  $\mathcal{G}_d^{m+1}$  does not depend on  $d$ . Therefore disclosure decision at period  $T - m - 1$  can be modeled without incorporating the expected future disclosure cost.

□

## A.8 The equilibrium condition

Recall stock price excluding disclosure costs is

$$P_t(s_t) = A_t s_t + B_t.$$

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

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

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  $w(s) = s - E(x|x \leq s)$  increases with  $s$ . Now re-write the equilibrium condition as

$$A_t \left( \sigma_{st} (\tilde{s}_t^* - E(\tilde{s}_t | \tilde{s}_t \leq \tilde{s}_t^*)) \right) = c \implies$$

$$\tilde{s}_t^* - E(\tilde{s}_t | \tilde{s}_t \leq \tilde{s}_t^*) = \frac{c}{A_t \sigma_{st}}. \quad (\text{A.18})$$

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

$$s_t^* = \sigma_{st} * \tilde{s}_t^* + q_{0t} + \rho_1 y_{t-1}. \quad (\text{A.19})$$

Then the probability of disclosure is:

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

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

1.  $\lambda_t$ . When  $\lambda_t$  increases, RHS of (A.18) 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 (A.18) increases. With the same logic as  $\lambda_t$ ,  $\tilde{s}_t^*$  increases, which decreases the disclosure probability.
3.  $\rho_1$ . The same reasoning as  $\beta$ .
4.  $c$ . When  $c$  increases, RHS of (A.18) 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  $\sigma_{st} * \tilde{s}_t^* + q_{0t} - \rho_0$  decreases, which increases the disclosure probability.
  
7.  $\kappa$ . When  $\kappa$  increases,  $A_t$  increases, so  $\sigma_{st} \tilde{s}_t^*$  decreases, that is, becomes more negative. In the meantime,  $\sigma_s$  decreases. The overall effect on  $\frac{\sigma_{st} * \tilde{s}_t^* + q_{0t} - \rho_0}{\sigma_s}$  is decreasing, which increases the disclosure probability.

## APPENDIX 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 1.1: 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. [24]. 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.

## APPENDIX 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  $q_{0i0}$ , investor uncertainty  $1/(\lambda_{0i})$ , the cost of disclosure  $c_i$ , discount factor  $\beta_i$ , the manager's information precision  $1/(\kappa_i)$ , firm profitability  $\rho_{0i}$ , 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  $\sigma_{ei}$ .

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 density of  $\beta$  conditional on the new  $\alpha$ . Repeating this step many times recovers the joint density of  $\alpha$  and  $\beta$ . The goal here is to learn the joint distribution of the firm-level parameters as well as the population parameters both conditional on the data. Gibbs sampling works as follows.

1. Impose necessary parametric assumptions, determine the set of parameters to estimate, and specify prior beliefs (of the econometricians) about their value.
  - The set of parameters to estimate.
    - Individual parameters:  $\theta_i \equiv (q_{0i0}, \rho_{0i}, \rho_{1i}, \kappa_i, \lambda_{i0}, \tau_i, \beta_i, c_i)'$  and  $\sigma_{ei}$ .
    - Population parameters:

The natural logarithm of  $q_{i0}, \rho_{i0}, \kappa_i, \lambda_{i0}, \tau_i, c_i$  follow normal distribution. The natural logarithm of the logit transformation of  $\rho_{1i}, \beta_i$ , that is,  $\ln(\frac{x}{1-x})$ , follow normal distribution. The mean and variance of these distributions are:  $\vartheta \equiv (\bar{q}_0, \bar{\rho}_0, \bar{\rho}_1, \bar{\kappa}, \bar{\lambda}, \bar{\tau}, \bar{\beta}, \bar{c})'$  and  $\mathbf{w} \equiv (w_{p_0}, w_{\rho_0}, w_{\rho_1}, w_{\lambda}, w_{\kappa}, w_c, w_{\tau})'$ , where  $\vartheta$  is the mean vector and  $\mathbf{w}$  is the variance vector. I am not interested in the cross-sectional distribution of  $\sigma_{ei}$ . To reduce computational burden, I do not impose a hierarchy on  $\sigma_{ei}$ . The parameters governing the distribution of the population parameters are called hyper-parameters.

- Prior beliefs of the econometrician for the hyper-parameters.

The prior belief of the econometrician for  $\vartheta$  is  $\mathcal{N}(-1, 1000^2)$  and that for  $\mathbf{w}$  is  $IG(1, 1)$ . Setting a large variance for the prior belief of  $\vartheta$  ensures minimal influence on the posterior distribution.

## 2. Initializing the Markov chain for Gibbs sampling.

I first draw a single vector of population parameters  $\boldsymbol{\theta}$  and  $\mathbf{w}$  from its prior distribution. Then 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:

$$\begin{aligned}
\pi(\theta|\mathbf{y}_t, \mathbf{s}_t, \mathbf{d}_t, \mathbf{P}_t) &\propto L(\mathbf{y}_t, \mathbf{s}_t, \mathbf{d}_t, \mathcal{P}_t|\theta)f(\theta|\vartheta) \\
&= \underbrace{\prod_{t=1}^T \left[ g(P(s_t|\mathbf{y}_T, d_t, \theta))^{d_t} g(P(s_t \leq s_t^*(\theta, \mathbf{y}_t)|\mathbf{y}_T, d_t, \theta))^{1-d_t} \right]}_{L(\mathbf{P}_t|\mathbf{s}_t, \mathbf{y}_t, \mathbf{d}_t, \theta)} \\
&\quad * \underbrace{\prod_{t=1}^T \left[ \frac{f(s_t|\mathbf{y}_T, \theta)}{1 - F(s_t^*(\theta, \mathbf{y}_t|\mathbf{y}_T, \theta))} \right]^{d_t}}_{L(\mathbf{s}_t|\mathbf{y}_t, \mathbf{d}_t, \theta)} \\
&\quad * \prod_{t=1}^T \left\{ [1 - F(s_t^*(\theta, \mathbf{y}_t|\mathbf{y}_T, \theta))]^{d_t} [F(s_t^*(\theta, \mathbf{y}_t|\mathbf{y}_T, \theta))]^{1-d_t} \right\} \prod_{t=1}^T (l(y_t|\theta))f(\theta|\vartheta), \quad (\text{C.1})
\end{aligned}$$

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  $\theta_i$  conditional on other parameters and data.

To draw  $\theta_i$ , I use Metropolis Hastings algorithm which involves a truncated  $t$  proposal density.

- (b) Individual parameters  $\sigma_{ei}$  conditional on other parameters and data.

The posterior of  $\sigma_{ei}$  is inverse-Gamma, which can be drawn directly.

- (c) Population parameters conditional on individual parameters.

I first convert the individual firm-level parameters to their corresponding natural logarithm form and then draw the population parameters according to:

- $\vartheta|\theta, \mathbf{w}$ , which has a normal distribution;
- $\mathbf{w}|\vartheta, \theta$ , which has an inverse gamma distribution.