When Dye meets Verrecchia: a structural estimation of disclosure theory

E. Cheynel M. Liu-Watts*

October 31, 2016

Abstract

We develop structural estimators for disclosure costs (or benefits), and empirical bounds for the probability of information endowment in a generalized disclosure model nesting Verrecchia (1983) and Dye (1985). The baseline estimator of disclosure costs is a closed-form function of (i) the minimum disclosure surprise, (ii) the average disclosure surprise, and (iii) the frequency of disclosure. Furthermore, it can be computed without solving for either the entire equilibrium of the game or knowledge of distributions. We derive the asymptotic properties of the estimator and illustrate, in simulations, its properties in small samples. A new empirical test is derived which can test for samples incompatible with the theory. The probability of information endowment is not point-identified but can be bounded from above in the case of disclosure benefits. Additionally, adapting an argument from conditional choice probabilities, the estimation extends to a multi-period setting with time-varying frictions as a function of observable state variables. As an application, we conduct the estimation using quarterly management earnings forecasts, characterize the magnitude of the disclosure costs, describe covariates associated with disclosure costs and perform various tests to detect the fraction of firms' behavior inconsistent with the theory or whose probability of information endowment is significantly less than one. The framework offers a simple theory-based approach to estimating voluntary disclosure models.

*E. Cheynel is an associate professor at Columbia University and M. Liu-Watts is an associate professor at Hunter College. Contact author: Edwige Cheynel, Columbia Business School, Columbia University, 3022 Broadway, New York, NY 10027. Email address: ec2694@gsb.columbia.edu. We warmly thank Jeremy Bertomeu, Moritz Hieman, Ivan Marinovic, Korok Ray, Paul Fischer, Michael Kirschenheiter, and participants at the Stanford Summer Camp, the Burton Workshop at Columbia Business School and the 2016 Annual American Accounting Association for helpful feedback.

In this paper, we propose a theory-based econometric framework to recover the disclosure costs, or benefits, implied by the classic model of Verrecchia (1983), generalized to incorporate uncertainty about information endowment (Dye 1985). Specifically, we use the theoretical restrictions implied by the model to determine an estimator of disclosure frictions. The estimator is in closed-form, can be easily interpreted, can be computed without solving the equilibrium of the game, and the estimation procedure is non-parametric, that is, it does not require distributional assumptions about privately-observed information. Furthermore, we separately identify disclosure costs (or benefits) from uncertainty about information endowment showing that, if both frictions are simultaneously present, the frequency of disclosure need not be a suitable proxy for each friction.

A novel benefit of the estimation procedure is to point identify disclosure benefits (or negative disclosure costs) provided information endowment is uncertain. Thus, the framework can potentially offer future research with measurements relevant to testing theories in which disclosure provides benefits, ranging from investment efficiency (Stocken and Verrecchia 2004; Liang and Wen 2007), lower threat of litigation (Dye 2013; Caskey 2013) or proprietary gains (Wagenhofer 1990; Darrough and Stoughton 1990; Darrough 1993). Such measurements cannot be achieved without theory because we only observe a single disclosure or no-disclosure decision in any given firm-period. We also provide a proof that, unlike costs or benefits, the probability of not receiving information is only set identified - however, in the case of disclosure benefits, the lower bound can allow researchers to test whether information endowment is uncertain. Lastly, while disclosure theory can, in principle, rationalize any observed frequency of disclosure, we derive a new test to reject the theory, that is, to detect samples that are incompatible with any choice of disclosure costs (or benefits) and information endowment.

Our baseline estimator for disclosure costs is a function of three observables: the minimum disclosure surprise, i.e., the lowest disclosure net of consensus ($min.surprise$), the average disclosure surprise ($avg.surprise$) and the disclosure frequency ($freq$),[1]

$$\hat{c} = min.surprise + \frac{freq}{1 - freq} avg.surprise. \tag{1}$$

We show that $\hat{c}$ is a consistent estimator of the disclosure cost in the generalized theory, as long as the theory does not imply full disclosure or no disclosure with probability one.[2]

[1] Both frequency and disclosure surprises are available, or can be calculated, in natural applications of this model. However, we can make no theoretical predictions about the correct variable to measure and scale surprises. For example, in the context of management earnings forecasts, one might use any variable capturing market change in expectations, such as short-window market response (Kasznik and Lev, 1995), earnings per share (Cheong and Thomas 2011), or earnings surprise scaled by lagged assets or prices.

[2] We have left aside considerations of risk-aversion modeled in the original Verrecchia (1983) model; that is, we assume

A practical challenge in the estimation is the existence of unobserved heterogeneity. Specifically, unlike in ordinary least squares regressions, estimates cannot be interpreted as averages over the population. In our setting, observations across firms and periods may feature different costs, or some periods may feature disclosures due to reasons that fall outside of the theory. Unfortunately, there is no simple solution for this, and we discuss a partial remedy along a modification of the baseline estimator.

Our main concern is heterogeneity due to disclosure by certain firms for reasons exogenous to the theory, so we assume that disclosures may be caused, by some probability, for reasons outside of the theory. We derive an equation analogous to (1) using the sample frequency and average surprise but the sample minimum must be modified. We offer two solutions that, under certain conditions, do not decrease the asymptotic performance of the estimation. First, if misspecification may occur only for a subset of an observable characteristic, we can replace the minimum by the median of the minimum across the characteristic. For example, if only certain firms may not follow the model, we will calculate the minimum per firm and use the median of the minimum of the subset. Second, for cases in which a sufficiently large set of observations may not follow the model, or characteristics are not observable, we show that the minimum can be replaced by a non-parametric estimate of the disclosure threshold. Both methods can be easily implemented using a standard statistical package.

We conduct Monte-Carlo simulations of the estimator to assess its performance with small samples, since we know in each simulation the true parameters (Gow, Larcker and Reiss, 2015). Reasonable firm-level estimates of the cost can be obtained with samples as low as 40 observations, although longer time-series are preferable if the disclosure cost and uncertainty about information endowment are large, as both reduce variation in the sample. We also find that a modified baseline estimator with a correction for finite-sample bias is highly desirable in small samples. The sample minimum has positive finite-sample bias as an estimator of the minimum possible disclosure, and we construct a modified estimator by substracting the difference between the minimum and the second lowest minimum disclosure. We also compare the performance of these estimators to the recent parametric moment-based estimator of disclosure costs in Bertomeu, Beyer and Taylor (2015a) and show that the estimators perform better in different samples.

Lastly, we show that the theory can be extended to incorporate time-variation in costs and information endowment based on observables. To give an example, disclosure costs may change as a function of the lifecycle of the firm or the behavior of competitors. Or, as modelled by Einhorn and Ziv (2008), disclosure

that the set of potential investors is large and the disclosure is about a diversifiable component such that there is risk-neutral pricing of the disclosed value (Cheynel, 2013). For most likely uses of this procedure, recall that individual investor risk-aversion is divided by the number of traders for the risk-aversion in the pricing function under exponential preferences; thus, for most firms traded in large exchanges, it would take a limited number of traders to be close to risk-neutrality. For a discussion of stocks with very low levels of active traders, we refer the reader to Armstrong, Core, Taylor and Verrecchia (2011) and for which risk-aversion adjustments should be considered. Note that, in our empirical section, we only use shares traded in one of the largest three exchanges.

strategies will be a function of public beliefs which depend on the public history of prior disclosures and prior realized earnings. Using a formulation from conditional choice probabilities based on Hotz and Miller (1993), we show that the repeated model can be estimated given knowledge of the state variables that determine the disclosure strategy in one period and inverting the estimated parameters to recover disclosure costs conditional on each state. A caveat is that per-period estimates depend on the discount factor, which is not identified from disclosure behavior and requires estimation via other data sources (such as a cost of capital model).

While our primary objective is to design a method that can be used across various disclosure samples, we illustrate its use in a simple empirical application, using management earnings quarterly guidance forecasts from 2004 to 2015. We choose this application because management forecasts are crucial to the development of disclosure theory, and they present many of the empirical features predicted by the theory, such as (on average) selection of positive news and periods of withholding that are not entirely predictable (Beyer, Cohen, Lys and Walther 2010). This application is also an opportunity to describe the practical implementation of the framework within a widely-studied sample.

First, we estimate the model assuming homogeneity only at the firm level. Out of a sample of $1,283$ firms, we find that 27% have significantly positive disclosure costs, and 40% have a probability of information endowment significantly below one. For 24% of the sample, neither positive costs nor perfect information endowment can be rejected, and about 9% of the sample rejects the theory. In summary, the firm-level estimates suggest that both disclosure costs and incomplete information endowment play an important role in explaining disclosure behavior. The median disclosure cost in the sample is negative and equal to -0.3% of lagged assets (Chart 1D). We conduct exploratory cross-sectional analyses with a regression of disclosure costs on several firm-level characteristics, averaged over the sample period. We observe that characteristics appear to be a function of whether there are costs or benefits of disclosure. For firms with positive costs, costs are increasing in growth opportunities (proxied by the market-to-book ratio), consistent with a proprietary component. For firms with disclosure benefits, benefits are decreasing in market concentration, capital expenditures, volatility of earnings, leverage and size. Industries that have been documented in the literature to bear more litigation risk tend to have more disclosure benefits, which is weak evidence that some of the disclosure benefits are related to litigation risk (Skinner 1997).

Second, we estimate the model by pooling over observations that may share homogenous costs and information endowment, pooling over past history of disclosure, number of segments and litigation risk. We find that disclosure costs are small (or a benefit) in the quarter after a disclosure, but then they may be positive after two or more quarters of withholding. This suggests that, as expected, the proprietary content of a forecast may be low if the firm has a history of frequent disclosure. Firms with less than two segments tend to exhibit disclosure benefits, which is unexpected based on the proprietary costs hypothesis but in line with potential agency costs in opaque diversified conglomerates. As in firm-level regressions, we find weak evidence that industries exposed to higher litigation risk tend to exhibit disclosure benefits. Lastly, we estimate a simple version of the multi-period estimator using past history of disclosure as a state variable. For simplicity, we use the number of periods since the last disclosure as a state, which is one of the state of variables in Einhorn and Ziv (2008). This choice of states is also motivated by the fact that most of the disclosure behavior is predicted by past disclosure. We find some evidence that costs may be greater after one period of withholding, especially in the three to four quarters that follow one period of withholding.

Literature Review

Our paper develops a micro-founded estimation of the disclosure costs and, in doing so, is a continuation of an extensive body of empirical studies testing implications of voluntary disclosure theory. Unfortunately, this body of research is far too large for us to faithfully review, and complete surveys of this literature can be found in Beyer, Cohen, Lys and Walther (2010). For our purpose, let us note two early studies to the extent that they pose the fundamental tests of theory. In an early study, Penman (1980) documents that firms voluntarily disclose not only "good" news but also relatively "bad" news because their disclosure leads to a downward re-evaluation. Lev and Penman (1990) test whether strategic disclosure choices are determined by higher earnings relatively to peers and whether such disclosures have a positive impact on the stock price. Their empirical insight is critical to show that the joint model is falsifiable, in that we show that the test can be used to reject the model if expected disclosure surprises are significantly negative. Note that the existing literature does not focus on the measurement of disclosure costs. Furthermore, we use the theory to develop a new test of the theory that relies on the relationships implied by the model.

Identifying the implied disclosure cost falls into the general category of approaches that unravel an unobservable cost from observed data. This approach is not unfamiliar to empirical research as a whole and, indeed, is recognized in the related area of implied cost of capital (Botosan, 1997; Hughes, Liu and Liu, 2009; Caskey, Hughes and Liu, 2012). As in our approach, the implied cost of capital assumes a model which is a function of an unobserved primitive, the cost of capital, as well as inputs, which may range from analyst forecasts to past growth. By inverting a pricing equation, it is then possible to infer the true cost of capital and, hence, map observables into the unobserved cost. Similar to this literature, we rely on assumptions from a model to obtain an estimate of an unobserved opportunity cost.

Our formal model builds on a disclosure threshold derived from price-maximization behavior by an informed owner-manager who optimally discloses to maximize current market prices. In this respect, our model fits closely with current structural estimation approaches in which the utility-maximizing problem is modeled, of which we give a few examples below. Four recent studies specifically focus on structurally estimating misreporting of earnings. Marinovic (2013) takes a cheap talk model to data, under the assumption of uncertainty about managerial willingness to misreport, and examines misreporting propensity prior to versus after the Sarbanes-Oxley Act of 2002. Within the same vein, Gayle, Li and Miller (2015) estimate an agency model in which managers are induced by an optimal contract to report earnings and compare the cost of the agency problem before versus after the Sarbanes-Oxley Act. Beyer, Guttman and Marinovic (2014) structurally estimate the equilibrium model of earnings management given investor uncertainty about accounting biases, which implies a firm-level measure of manipulation. Zakolyukina (2012) offers a different approach under the assumption that managers trade off the benefits of manipulation with a probability of detection. In her model, therefore, actual detections (restatements) can be used to identify the level of manipulation for firms that have not been detected. To our knowledge, the only study that applies structural estimation outside of reporting problems in accounting is Gerakos and Syverson (2015), who structurally estimate a model of choice by preparers over auditors.

Within this literature, our approach is closest to Bertomeu, Ma and Marinovic (2015c), hereafter BMM, since they also estimate a strategic disclosure model. On some dimensions, the non-parametric method used here is partly assumption-lite, in that we do not necessarily impose normally-distributed signals, and our estimation procedure will be robust to uncertainty about information endowment; in BMM, on the other hand, the cost is assumed to be zero. Another advantage that has guided our empirical construct is to offer a simple closed-form estimation procedure that does not require us to solve a formal equilibrium model to be used. On the other hand, BMM allow for a richer setting with time-series correlation in the latent friction, while our estimation procedure requires the friction to be a function of observables.

Two recent papers present alternative structural models that speak to the measurement of disclosure costs. Bertomeu, Beyer and Taylor (2015a) derive a closed-form estimator for the Verrecchia (1983) under risk-neutral investors with normally distributed information. The estimator is different (i.e., their estimator is not a special case of ours without normality) and imposes by assumption positive costs. Interestingly, they also find that, on its own, the threshold equilibrium is inconsistent with a large portion of observed disclosures (about 14% of management forecasts are lower than the disclosure threshold implied by the estimates of disclosure costs from Verrecchia 1983). Note that our procedure specifically imposes the threshold as a tool for estimation and provides additional flexibility by allowing for uncertainty about information endowment. In further simulations, we find that their estimator performs similarly to our baseline estimator if the distribution is normal, but one benefit of their approach is that it tends to be less sensitive to heterogeneity due to exogenous disclosures. Zhou (2016) estimates a Bayesian model where firms may have different costs in the cross-section, and investors learn about the persistent components of earnings. His model has the capacity to accommodate heterogeneity about the cost parameter. Unfortunately, we do not know whether his methods can be easily applied to the joint model of costs and information uncertainty and, to be fair, his analysis of disclosure dynamics goes far beyond our more modest objective of creating a simple measure with some embedded theoretical content.[3]

1 Theoretical framework

We closely follow the disclosure cost model of Verrecchia (1983) and, furthermore, allow for the possibility of uncertainty about information endowment as in Dye (1985). For expositional purposes, we restrict the disclosure cost c and the probability to be uninformed q to the following conditions: either (i) $c > 0$ and $q \in [0, 1)$ or (ii) $c \in \mathbb{R}$ and $q \in (0, 1)$. If neither conditions (i) or (ii) hold, we cannot extract much information about the structural parameters since there will not be any variation in disclosure. For example, if $q = 0$ and $c \leq 0$, the model predicts unravelling to full-disclosure, and we will only be able to set identify $c \leq 0$ or, if $q = 1$, the model will predict no-disclosure regardless of c.[4]

Hereafter, we refer to a forecast about earnings, although the theory applies more generally to any price-relevant information. The manager observes a private signal s with probability $1 - q$. We define $x \equiv \mathbb{E}(e|s)$, where the private information x is drawn from a continuous distribution with p.d.f. $g(.)$ and c.d.f. $G(.)$, and such that $\mathbb{E}(.)$ reflects the expectation conditional on all relevant predetermined public information and e is the future earnings. All random variables are continuously distributed with support over $\mathbb{R}$ and have finite moments. For expositional purposes, we focus here on disclosure in a single period setting, but we show later on that the procedure extends to time-varying disclosure costs or information endowment.

Assume that the pre-forecast market expectation about earnings $v = \mathbb{E}(e)$ is common-knowledge

[3]Our approach also relates to a growing literature that draws explicit connections between a formal model and its empirical tests. While reviewing this literature in its entirety would take us away from our research question, note that several recent papers offers combined approaches to formal theory and testing. Bertomeu, Evans III, Feng and Tseng (2015b) develop a model of collusive competition and test its predictions in the automotive industry; Chen and Jiang (2006) test whether analysts use information in a Bayesian manner; Indjejikian and Matějka (2009) and Matějka and Ray (2014) test theoretical properties of optimal incentive contracts; Gerakos and Kovrijnykh (2013) offer a measure of misreporting based on theory-predicted reversals.

[4]$c \leq 0$ and $q = 0$ are observationally equivalent. In the empirical analysis, if we observe full disclosure, we will drop these observations. Similarly, if $q = 1$, we cannot identify the disclose cost and we will drop these observations.

as in, say, a market consensus. Without loss of generality, we normalize the consensus to $\mathbb{E}(e) = 0$.[5] This assumption implies that the manager's private information represents the manager's belief about the deviation from a consensus, that is, a forecast surprise. The manager maximizes market perception about future earnings — this is equivalent to maximizing current stock price if the expected stock price is a linear function of earnings. If the manager receives the information with probability $1 - q$, he can disclose but then bears a proprietary cost that reduces firm value by c, measured in units of earnings, or withhold information at a cost normalized to zero.

Given any information set $\mathcal{I}$, investors are risk-neutral and value the firm linearly as a function of expected earnings. Without loss of generality, assume that[6]

$$P(\mathcal{I}) = \mathbb{E}(e|\mathcal{I}). \tag{2}$$

It is not important for the model if not disclosing may involve a payoff as long as we interpret c as the incremental cost (positive or negative) of disclosure. In what follows, since our purpose is to develop the theoretical background for the estimator, we assume that the researcher has constructed a sample that is facing the same information environment.[7]

If the manager receives information, he can truthfully disclose x. We define $P(\{x\})$ as the market price if the firm's private information x is observed and disclosed. The price $P(ND)$ is offered if the firm withholds. If the manager is informed, he chooses to disclose to maximize current price if

$$P(\{x\}) - c \geq P(ND) \tag{3}$$

and not disclose otherwise.[8] Substituting the pricing equation (3), the disclosure decision in (4) reduces to

$$x - c \geq \mathbb{E}(x|ND). \tag{4}$$

From this equation, in a Bayesian-Nash equilibrium, the manager discloses if and only if the manager's information is above a threshold, $x \geq \tau$, where the disclosure threshold τ is given by[9]

[5] We could have then equivalently used the demeaned variables $e^s = e - v$ and $x^s = x - v$, which have mean zero, as the earning surprise and the manager's belief about the deviation from consensus respectively.

[6] In other words, if the market price is $P(\mathcal{I}) = \alpha\mathbb{E}(e|\mathcal{I})$, estimated variables can be rescaled by α, which can be estimated empirically. We omit α since it plays no further role in the estimation.

[7] Note that the estimation is flexible to a true data generating process with changing public information. For example, for frequent (sticky) disclosers, the estimator can be computed separately for periods after a disclosure and periods after no disclosure. If disclosures are seasonal, the estimator can be computed depending on the quarter; if disclosures have changed after a major regulatory change, the estimation can be run pre versus post change. Vice-versa, if some firms are reasonably similar (in the sense of their information and disclosure cost), the estimation can be run using disclosures for multiple firms, such as firms with similar size and assets or those in the same industry.

[8] The reader may note that we have laid out the model as if the manager paid the disclosure cost directly, mainly to avoid having to redefine the price equation to include the costs and burdening the exposition.

[9] In fact, this equation is exactly the same as in Jung and Kwon (1988), subtracting the cost from the disclosing firm

$$\tau - c = \frac{(1-q)G(\tau)\mathbb{E}(x|x<\tau)}{q+(1-q)G(\tau)}. \tag{5}$$

The above indifference condition can admit multiple threshold equilibria, which can cause technical problems when solving the model theoretically. One solution is to assume that x is logconcave with sub-exponential lower tail or select the highest threshold in the set of threshold equilibria (see Dye 1986) because it is Pareto preferred. For our empirical analysis, we may alternatively assume that the same threshold equilibrium is played in the entire sample.

One problem when estimating equation (6) is that the withheld information is not observable, so that it is not possible to directly observe $\mathbb{E}(x|x<\tau)$ to recover the cost. This turns out to be a relatively minor issue given some elementary arithmetic. Recall that, by definition,

$$0 = \mathbb{E}(x) = p\mathbb{E}(x|x \geq \tau) + (1-p)\mathbb{E}(x|x<\tau), \tag{6}$$

where $p = (1-q)Prob(x \geq \tau)$ is the equilibrium probability of disclosure.

Substituting in $\mathbb{E}(x|x<\tau)$ from equation (6) into equation (7), the cost is then given by

$$c = \tau + \frac{p}{1-p}\mathbb{E}(x|x \geq \tau). \tag{7}$$

The left-hand side of this equation is the cost to be estimated, and the right-hand size includes three components that can be directly estimated from the sample as suggested by equation (1).

Let us conclude with a word of modesty about what we can learn about c from observed disclosure behavior, by trying to distinguish between (possibly useful) extrapolations from additional assumptions and information obtained from data. Our main purpose here is the measurement of a disclosure cost as implied by the revealed preference of the marginal discloser. To make this idea clear, the model can be generalized costs that are a function of the news $c(x)$ (Dye 1986) in which case equation (8) becomes

$$\tau - c(\tau) = \frac{p}{1-p}\mathbb{E}(x|x \geq \tau). \tag{8}$$

Hence, the model can only identify the cost for the marginal discloser which may or may not (depending on the setting) be generalizable to other $x \neq \tau$ that were not at the threshold.[10]

Equation (9) also illustrates that the structure of the problem does not make assumptions about the cost being of one particular nature. As a useful analogy, a model of expected return may not always be complete enough to identify the reasons behind the existence of this return (even if doing so remains the

price, i.e., changing τ into $\tau - c$.

[10] To be interpreted as the cost for the marginal discloser, $x - c(x)$ needs to be increasing in the forecast disclosed.

ultimate objective). By writing the threshold equation with variable cost

$$\tau - c(\tau) = \mathbb{E}(x|x \leq \tau),$$

we could think about $c(\tau) < 0$ as being a benefit in the form of a reduction in future litigation when disclosing, or a loss due to bad investments being made when not disclosing. Or, within product market theories, $c(\tau)$ can be positive or negative depending on the effect of disclosure of the aggressiveness of competitors. That is, the model is agnostic about the nature of factors that cause a cost or benefit, in that the net effect of these factors will be captured in $c(\tau)$. Nevertheless, obtaining an estimate of the net effect may help, in a second step, detect variables that correlate with the estimate in order to build a complete economic model that clarifies the determinants of the variable.

2 Econometric model

2.1 Notations

We present next the econometric model following the treatment and notation in Matzkin (2007) in order to establish the identification of the disclosure cost and the empirical content of the theoretical model. The econometric model is defined to make explicit and formal the mapping from observables to unobserved variables of interest.

Let us define three variables that are exogenously determined, the *observable* earnings variable $\tilde{e}$ and the *unobserved* vector $\mathcal{E} = (\tilde{x}, \tilde{r})$ where $\tilde{x}$ is the potential private information of the manager and have support over $[\underline{l}, \bar{l}]^2$ and $\tilde{r}$ is a binary variable that equals one if the manager is informed. Then, we denote $Y = (\tilde{z} \in \{NI\} \cup \mathbb{R}, \tilde{d})$, the vector of endogenous observable variables, where $\tilde{z}$ is equal to the manager's private information if disclosed or NI otherwise, and $\tilde{d}$ is a binary variable equal to one if the manager discloses.

The econometric model S is the set of parameters and distributions that satisfy the restrictions imposed by the model. Specifically, let us define an element of S as $\xi = (h(e, x), q, c)$ where $h(.)$ is the joint p.d.f. of $(\tilde{x}, \tilde{r})$, q is the probability to be uninformed and c is the disclosure cost. As before, we assume that either $q \in (0, 1)$ or $c > 0$. For any $\xi \in S$, let $F_{Y,e}(.; \xi)$ be the distribution of the endogenous observable variables Y and we refer to $F_{Y,e} = F_{Y,e}(.; \xi^*)$ as the distribution conditional on the true value ξ^*. In the analysis, it will be simpler to work with the p.d.f. $f_{z,e}(z|e)$ when $z \in \mathbb{R}$, which can be derived from $F_{Y,e}$.[11] Then, we define the minimum τ, the probability to disclose for a given earning $p(e)$, and

[11] To be more precise, z is a mixture of a continuous random variable (when z is disclosed) and a discrete random variable (when $z = NI$). We define $f_{z,e}(z|e)/(1 - p(e))$ as the density of $z|e \neq NI$ where $p(e) = Prob(z \neq NI)$.

the marginal distribution of earnings $f(e)$ by

$$\begin{aligned}\tau &= Inf\{z : f(z|e) > 0\}, \\ p(e) &= \int_{\tau}^{+\infty} f(z|e)dz, \\ f(e) &= \int h(x,e)dx.\end{aligned}$$

We say that $\xi \in S$ if it satisfies the following theoretical restrictions

$$\forall x, \quad x = \frac{\int e h(x,e)de}{\int h(x,e)de}, \tag{9}$$

$$\forall z \geq \tau, \forall e, \quad f(z|e) = \frac{(1-q)h(z,e)}{\int h(x,e)dx}, \tag{10}$$

$$c = \tau + \frac{\int \int_{\tau}^{+\infty} z f(z|e)dz f(e)de}{1 - \int \int_{\tau}^{+\infty} f(z|e)dz f(e)de}. \tag{11}$$

Equation (10) corresponds to the posterior expectation of the manager if he receives a private signal and is implied by the (definitional) assumption made in equation (2) of the theoretical model. Equation (11) is the density of observed disclosures given that information is received with probability $1-q$ and the manager plays a threshold equilibrium. The equilibrium disclosure cost in equation (12) is expressed as the solution to the equilibrium equation derived in equation (8).

2.2 Identification

Suppose that we are interested in a component of ξ, which we define as $\phi(\xi)$ - for example, we could be interested in recovering $\phi(\xi) = q$ or $\phi(\xi) = c$. We say that this component is identified if, given a set of observables, there is a single possible value of $\phi(\xi)$ that can rationalize the observables. Formally, we say that the component is identified if, conditional on observing $F_{Y,e}$, any two $\xi, \xi' \in S$ such that $F_{Y,e}(.|\xi) = F_{Y,e}(.|\xi') = F_{Y,e}$ must also satisfy $\phi(\xi) = \phi(\xi')$.[12]

If a component is not identified, we may place the parameter of interest within a set. Specifically, we say that a component is set identified if, conditional on observing $F_{Y,e}$, any two $\xi, \xi' \in S$ such that $F_{Y,e}(.|\xi) = F_{Y,e}(.|\xi') = F_{Y,e}$ must also satisfy $\phi(\xi), \phi(\xi') \in \Lambda$ where Λ is a set that may depend on $F_{Y,e}$. Below, we show that c is identified and q is set identified.

Proposition 1 *c is identified but q is not identified. Further, q is set identified and in the set* $\Lambda = [0, 1 - \int p(e)f(e)de)$ *if* $c > 0$ *and in the set* $\Lambda = [(1 - \int p(e)f(e)de)(1 - \frac{\tau - c}{\tau}), 1 - \int p(e)f(e)de)$ *if* $c \leq 0$.

[12]Of course, ideally, we would identify $\phi(\xi) = \xi$; however, doing so is impossible in most applications (including ours) since we observe fewer variables than those in ξ — in fact, we shall see that ξ is not identified in our model because q is not identified.

The disclosure cost is identified given some variation in disclosure behavior, that is, if there are some periods with and without disclosure. The proof of identification follows from the theoretical equation (8), since the cost function can be written in terms of variables that can be estimated directly from the data. A similar point identification of the probability of information endowment is not possible, however. To see why, consider the simpler setting in which we only observe disclosure but we do not observe earnings (this argument is shown to extend to observable earnings). Recall that a manager with sufficiently unfavorable earnings will not disclose regardless of the information endowment, so this manager behaves as a manager with no information. We also do not observe this event, because it is withheld. Hence, we cannot distinguish between cases in which the manager does not have information or is informed with very unfavorable information.

2.3 Empirical content

We say that the theory has empirical content if there exist observables that cannot be rationalized. Specifically, the joint theory has empirical content if there exists (at least) one observable $F_{Y,e}$ such that $F_{Y,e} \neq F_{Y,e}(.|\xi)$ for any $\xi \in S$.

Proposition 2 *The joint theory with disclosure costs and uncertainty about information endowment has empirical content.*

(i) Let us denote $ND =$ "$\tilde{d} = 0$" *as a non-disclosure event. If* $\mathbb{E}(e|ND) \geq 0$*, there exists no data generating process (DGP) that satisfies the econometric model* S*.*

(ii) Let us define $\underline{q} = \max(0, (1-p)(1 - \frac{\tau - c}{\tau}))$*,* $(1-p) - \underline{q} < 0$*, there exists no data generating process (DGP) that satisfies the econometric model* S*.*

The proof for empirical content (i) builds on an earlier testing procedure designed by Lev and Penman (1990). They note that, in a threshold equilibrium, market prices should respond (on average) positively to disclosure. In our setting, this observation can be restated as disclosures predicting positive earnings surprises, relative to the pre-disclosure market consensus. By the law of total expectations, this implies that a non-disclosure event should predict a negative earnings surprise. Hence, a data set that features significantly negative earnings surprises following a non-disclosure, ie., $\mathbb{E}(e|\tilde{d} = 0) \geq 0$ can be rejected, is inconsistent with the theory. The second test for empirical content (ii) is related to the fact that we can derive a lower bound for the probability to be uninformed greater than zero when $c < 0$ and if the non-zero lower bound is statistically greater than $(1-p)$, the firms reject the model.

We may also be interested in testing pure costly disclosure against pure uncertainty information endowment, to assess whether one of the two theories can explain disclosure on its own. With a slight abuse in language, we say that the pure costly disclosure model (resp. uncertainty about information endowment model) has empirical content if there exists $F_{Y,e} \neq F_{Y,e}(.|\xi)$ for any $\xi \in S$ that also satisfies $q = 0$ (resp., $c = 0$).

Corollary 3 *The models with pure costly disclosure and uncertainty about information endowment have empirical content.*

A direct proof of Corollary 3 can be obtained from earlier proofs. First, because c is identified, a zero disclosure cost can be ruled out when observing any $F_{Y,e}(.|\xi)$ for a ξ with $c \neq 0$. Second, while we cannot identify q, there exists $F_{Y,e}(.|\xi)$ such that ξ satisfies $c < 0$ and $q > 0$. Intuitively, absent uncertainty information endowment, any disclosure benefit would imply full-disclosure - so that any sample where disclosure benefits are estimated but some periods of withholding occur, must be feature some uncertainty about information endowment.

2.4 Estimation

Equation (9) suggests the estimator

$$\hat{c} = \hat{\tau} + \frac{\hat{p}}{1-\hat{p}}\hat{m}, \tag{12}$$

where $\hat{\tau}$, $\hat{p}$ and $\hat{m}$ are, respectively, the sample minimum disclosure surprise, the disclosure frequency and the average disclosure surprise.

In the next proposition, we show that this estimator is $\sqrt{n}-$consistent estimator of the disclosure cost.

Proposition 4 *The estimator $\hat{c}$ is a consistent estimator of c (satisfies plim $\hat{c} = c$) with asymptotic variance given by*

$$\sqrt{N}(\hat{c} - c) \rightarrow_d N(0, \sigma_c^2), \tag{13}$$

where $\sigma_c^2 = \frac{p(pm+m(1-p))^2+(1-p)pv_x}{(1-p)^3}$ and $v_x = Var(\tilde{x}|\tilde{x} \geq \tau)$.

The measure relies on a simple intuition. One may think about the estimator $\hat{c}$ as a synthetic summary of the information contained in a sequence of forecasts. For example, if the manager discloses extreme negative surprises, then $\hat{\tau}$ will be very negative, and the estimator will predict negative disclosure costs. Vice-versa, consider a sample in which one observes very high average forecasts $\hat{m}$. In general, this observation will suggest that firms are actively selecting which forecasts to disclose and suggests positive disclosure costs.

The asymptotic variance of the estimator can be easily estimated using sample moments, that is, replacing all elements of the asymptotic variance by their sample estimates, i.e.,

$$\hat{\sigma}_c^2 = \frac{\hat{p}(\hat{p}\hat{m}^s + \hat{m}^s(1-\hat{p}))^2 + (1-\hat{p})\hat{p}\hat{v}_x}{(1-\hat{p})^3},$$

where $\hat{v}_x$ is the sample variance of forecasts, then $\hat{\sigma}_{c^n}$ is a consistent estimator of $\sigma^2_{c^n}$.

Remark that the standard-error of the estimator becomes unboundedly large in the special case of $q, c \to 0$, that is, if the true parameters are very close to unravelling. This is to be expected as, when unravelling occurs, we become unable to identify any c in the range $(-\infty, c)$ as there is no longer cross-sectional variation in disclosure. For samples that are very close to full disclosure, we thus recommend to set identify "$c \leq 0$" over trying to recover the actual cost.[13] Note that equation (14) allows researchers to test whether the cost is significantly positive or negative.

When $\hat{c} < 0$, we can estimate $\underline{q}$ as follows:

$$\hat{\underline{q}} = 1 - \hat{p} + \frac{\hat{p}}{\hat{\tau}}\hat{m}, \tag{14}$$

where $\hat{\tau}$, $\hat{p}$ and $\hat{m}$ are, respectively, the sample minimum disclosure surprise, the disclosure frequency and the average disclosure surprise.

Proposition 5 *(i) When $c < 0$, the estimator $\hat{\underline{q}}$ is a consistent estimator of q with asymptotic variance given by*

$$\sqrt{N}(\hat{\underline{q}} - \underline{q}) \to_d N(0, \underbrace{\frac{1}{\tau^2}p(v_x + (1-p)(m-\tau)^2)}_{\sigma^2_{\underline{q}}}). \tag{15}$$

(ii) If $c < 0$, the estimator $(1-\hat{p}) - \hat{\underline{q}}$ is a consistent estimator of $1-p-q$ and given by

$$\sqrt{N}(1 - \hat{p} - \hat{\underline{q}} - (1-p) + \underline{q}) \to_d N(0, \frac{1}{\tau^2}p(v_x + (1-p)m^2)). \tag{16}$$

Note that uncertainty about information endowment may only be tested in the case of disclosure benefits, since (otherwise) an interior disclosure threshold would arise even if information endowment were perfect. Specifically, using equation (16), we can reject $\underline{q} = 0$ which provides evidence that information endowment must be incomplete. Equation (17) offers a new test of the theory relying on the set identification of $\hat{q}$. If this set is empty, that is, no information endowment can explain the sample, then the theory can be rejected. We conduct this test by comparing the lower bound $\overline{q}$ to the upper bound on $\hat{q}$. This upper bound follows from the fact that the probability of information endowment cannot be

[13] To be precise, as long as we restrict the analysis to set identifying $c \leq \epsilon$ where $\epsilon > 0$ is a positive number or identifying the true c if $c > \epsilon$. Then, the standard-error of the estimator, as defined by minimum distance between the set where the estimator places c and the true c, will be bounded.

lower than the frequency of disclosure.

Note that the sampling error in the estimation of τ does not affect the asymptotic standard errors. This is because (as also shown formally), the minimum of the sample converges to the disclosure threshold at a rate greater than $\sqrt{N}$. Having noted this, for small samples, the estimator will typically carry positive small sample bias because the sample minimum is always greater than the theoretical minimum. The existence of a bias is not unusual since finite-sample bias exists for most non-linear estimators (such as logit/probit or generalized method of moments) or commonly-used methods such as instrumental variables. Nevertheless, we propose a modified estimator

$$\underline{\hat{c}} \quad = \quad \hat{\tau} - (\hat{\tau_2} - \hat{\tau}) + \frac{\hat{p}}{1-\hat{p}}\hat{m},$$

where $\hat{\tau_2}$ is the second lowest disclosure.[14]

To assess the performance of the estimators, we examine the error and bias of the estimator in finite samples in situations with normally-distributed random variable. Furthermore, we compare the performance of the non-parametric estimators to the parametric estimator of disclosure costs proposed in Bertomeu, Beyer and Taylor (2015a), hereafter BBT. By explicitly solving the threshold equation of Verrecchia (1983) with normal distributions, they derive an estimator of disclosure costs given by

$$\hat{c}_{BBT} = (\Phi^{-1}(1-\hat{p}) + \frac{\phi(\Phi^{-1}(1-\hat{p}))}{1-\hat{p}})\hat{\sigma}_x, \tag{17}$$

where $\hat{\sigma}_x = \sqrt{\frac{\hat{v}_x}{1+\Phi^{-1}(1-\hat{p})\frac{\phi(\Phi^{-1}(1-\hat{p}))}{\hat{p}}-(\frac{\phi(\Phi^{-1}(1-\hat{p}))}{\hat{p}})^2}}$ and $\hat{v}_x$ is the sample variance of the disclosures.[15]

Note that, because $\hat{c}_{BBT}$ uses additional parametric information that neither $\hat{c}$ nor $\underline{\hat{c}}$ use, it may be presumed to be more efficient if private information is normally-distributed and there is no uncertainty about information endowment. Then, by comparing the performance of our estimators to the BBT estimators, we may thus obtain an insight about the potential efficiency loss when not relying on normality or ruling out the theory of Dye (1985).

We simulate three families' data-generating process (DGP) with x drawn from a standard normal - so that costs are measured in standard-errors of x with 40 observations (to model a typical small sample obtained from quarterly time-series data over a 10-year horizon) and structural parameters (q, c) chosen as follows:

[14] Note that $\hat{\tau} - (\hat{\tau_2} - \hat{\tau})$ also converges to the minimum at a rate greater than $\sqrt{N}$ so that the modified estimator satisfies the asymptotic results in Proposition 4.

[15] To save space, we do not reprove their derivation, although it can be found in section 4 of BBT. We use a slightly generalized of the BBT estimator because the expression $\Phi^{-1}(1-\hat{p}) + \frac{\phi(\Phi(1-\hat{p}))}{1-\hat{p}}$ is a cost standardized by the standard-error of x. Proofs for the fact that $\hat{\sigma}_x$ is a consistent estimator of σ_x are available on request and can also be found in the original working paper version of BBT.

1. For $DGP_{c,0}$, we set perfect information endowment and vary the disclosure cost from zero to three. This is the only model for which the BBT estimator is well-specified, so we restrict our comparison with BBT to this case.

2. For $DGP_{0,q}$, we set the disclosure cost to zero and vary the probability of information endowment q from zero to one.

3. For $DGP_{c,.25}$, we vary the disclosure cost from -1.5 to 1.5, assuming that the manager is uninformed with probability .25.

For each case, we draw x from a standard-normal and (for 2. and 3.) the information endowment from a binary distribution with coefficient $1-q$, compute the disclosure threshold from the theoretical equilibrium in equation (12), and generate an observation where the disclosure is set as missing if and only if the information endowment variable was zero or the simulated x was below the threshold. Then, we compute $\hat{c}_i$, $\underline{\hat{c}}_i$ as well as $\hat{c}_{BBT\,i}$ for $DGP_{c,0}$ and derive the bias $Bias_i = (c - \hat{c}_i)$ and the absolute error $Error_i = |c - \hat{c}_i|$. The estimators require at least one non-forecast periods and at least two forecast periods, so all samples for which one of the estimators could not be computed are dropped - this is a source of small sample (selection) bias when applying the estimator to actual data that the simulations can quantify. We repeat this procedure for $i = 1$ to $10,000$, and define $\hat{Bias}$ and $\hat{Error}$ as the average over all $Bias_i$ and $Error_i$, respectively.

In the top panels of figure 1, we plot the performance of the estimators under the pure costly disclosure model $DGP_{c,0}$. We observe that the absolute error and bias can be large for small costs, below .4. Note that this effect is not driven by small-sample deviations from asymptotic theory as, in equation (14), the asymptotic standard-error of the estimator increases when trying to estimate small costs because the estimator may incorrectly find large disclosure benefits. In other words, it is difficult to distinguish the level of benefits for samples close to full disclosure (with no uncertainty about information endowment). The modified estimator corrects most of the bias starting at around $c = .7$. As is usual, the bias-correction introduces additional error - and the two estimators are very similar in terms of absolute error. Nevertheless, errors will tend to average out when examining the average cost or using it in a cross-sectional study, while finite-sample bias will not. Hence, these simulations give the edge to the modified estimator.

The parametric BBT estimator behaves very differently. Contrary to the previous estimators, it performs reasonably well for low disclosure costs below .4. On the other hand, it performs more poorly than our estimators as costs increases - in fact, it tends to have higher absolute error and bias than the modified estimator for costs greater than 1. This loss of efficiency is due to the fact that the BBT

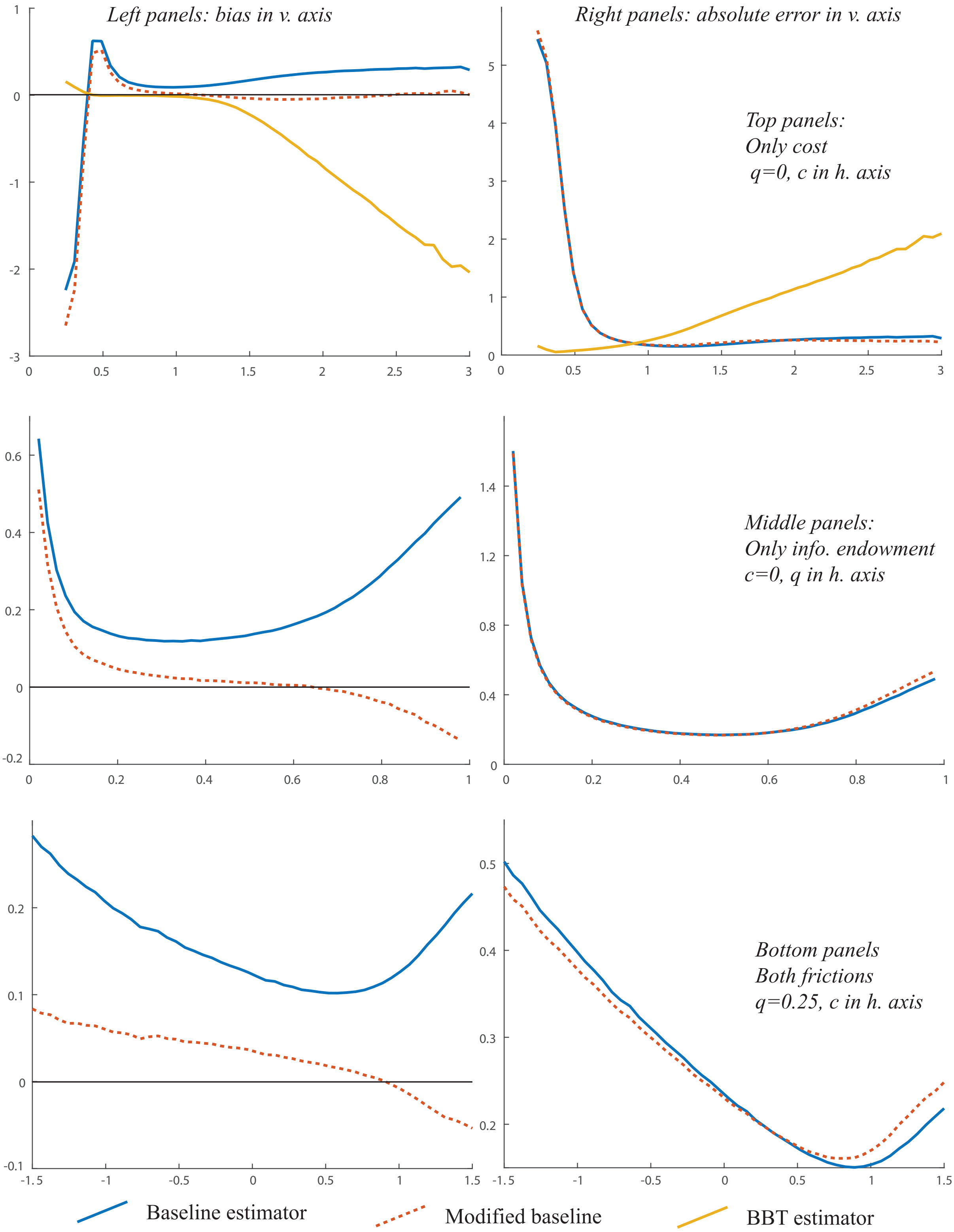

Figure 1. Finite-sample performance

estimator recovers the *actual* disclosure threshold by estimating the variance of forecasts $\hat{\sigma}_x$ which tends to be noisy with few observed forecasts, i.e., when c is large.

In the middle panel, we evaluate whether the estimators correctly predict zero cost in $DGP_{0,q}$ if disclosure is entirely driven by uncertainty about information endowment. The estimators tend to perform more poorly in environments close to full-disclosure or no-disclosure, that is, when the probability of information endowment is lower than .1 or greater than .9. The modified estimator performs very well at identifying zero cost, as the bias and error are below .05 for most information endowment parameters $q \in [.25, .8]$.

In the bottom panel, we conduct simulations of $DGP_{c,.25}$ when the data-generating process features both costs and uncertainty about information endowment. In this setting, we also ask whether the estimators precisely estimate benefits when $c < 0$. Some degree of uncertainty about information endowment increases the performance of the estimator, because it prevents full-disclosure samples in which benefits cannot be distinguished from $c = 0$ — that is, it creates variation in disclosure behavior for low cost. While large benefits or costs remain more difficult to estimate, we find that, in the presence of some uncertainty about information endowment, the two estimators provide plausible estimates when costs are close to zero, in contrast to the top panel. If $c = 0$, the finite-sample bias of the baseline estimator is .12 and .03 for the modified estimator.

In summary, these simulations illustrate a few important properties of the estimators. Using small samples of 40 observations, the estimators perform poorly when there is little cross-sectional variation in disclosure, in particular when both costs and the probability of not receiving information are low. We also illustrated the BBT estimator performs well with low costs, and thus may offer a preferable alternative for samples with few non-disclosure observations.

3 Pooled estimation

3.1 Unobserved heterogeneity

The baseline estimators assume that there are samples with homogenous disclosure costs — in most applications, the natural sample may be repeated observations of a disclosing unit such as a firm or individual. However, certain problems may feature limited-size homogenous samples, and it may be unavoidable to estimate the costs with some degree of heterogeneity using large samples. Naturally, to formally address heterogeneity, we need to make (semi-parametric) assumptions about the nature of the heterogeneity (see, e.g., Arcidiacono and Miller 2011). We shall focus on the following type of heterogeneity which we refer as forced disclosure. Assume that the sample is indexed by $i \in [0, I]$,

$t \in [0, T]$ where one may, for now, interpret i as a homogenous reporting unit (e.g., a firm) and t as a time period. We maintain the assumption that I and T are linear functions of total sample size $N = I\text{x}T$.

Suppose that, with probability $\theta_0 \in (0,1)$, reporting units may play a different game in which they may, when receiving information, disclose their information with probability $\theta_1 \in (0,1)$. Note that the realized θ_1 is never observable, which causes unobserved heterogeneity in the sample. We refer to units that may feature this event as misspecified units. If a unit is misspecified, or if with probability $1-\theta_1$, it does need to report for exogenous reasons, it still faces an homogenous cost c. We make the simplifying assumption that markets are aware of this friction and rationally price the firm but do not observe nor learn about units.[16]

We choose this formulation of mandated disclosure for three primary reasons. First, it can capture reasons orthogonal to the theory why the manager may be required to disclose information, such as a duty to disclose due to particular characteristics of an ongoing event. Second, mandated disclosure implies the sample minimum $\hat{\tau}$ is no longer a consistent estimator of the disclosure threshold since it will converge to the minimum x with probability one. Third, from a modeling perspective, it is symmetric to uncertainty about information endowment, to the extent that while information endowment prevents the manager from disclosing with some probability, mandated disclosure forces disclosure with some probability.

Let us rewrite the disclosure threshold equation (6)

$$\tau - c = \frac{(1-q)(1-\theta)G(\tau)\mathbb{E}(x|x<\tau)}{q+(1-q)(1-\theta)G(\tau)} = \frac{-(1-q)(1-\theta)\int_{\tau}^{+\infty} xg(x)dx}{q+(1-q)(1-\theta)G(\tau)},$$

where $\theta = \theta_0\theta_1$ is the probability of mandated disclosure and reduces the probability of strategic disclosure by $(1-\theta)$ when the manager is informed.

The probability of disclosure is now given by $p = \theta(1-q) + (1-\theta)(1-q)(1-G(\tau))$ and defining the disclosure event by D,

$$\mathbb{E}(x|x \in D) = \frac{(1-q)(1-\theta)\int_{\tau}^{+\infty} xg(x)dx}{p}$$

Thus, we obtain the same equation as in (8) which does not depend on θ, namely,

$$c = \tau + \frac{p}{(1-p)}\mathbb{E}(x|x \in D). \tag{18}$$

[16]Typically, the techniques discussed here are of interest for pooled estimations where time-series (and, thus, potential for learning) are limited. Further, for two of the three estimators proposed here, we assume that the reporting unit is not observable which would make learning difficult, and we attempt to make our treatment of the estimator with observed units consistent with these. Even for the estimator with observed units $\hat{c}_{mm}$, for a sample of units large enough, we show later on that the estimator of the minimum in $\hat{c}_{mm}$ will never be consistent from a misspecified firm. Investor learning is modeled in Bertomeu *et al.* (2015c) and Zhou (2016) and would typically requires more assumptions about distributional forms in order to be able to compute investor expectations.

In particular, equation (19) reveals that the estimator for the disclosure cost may still use the sample frequency $\hat{p}$ and sample forecast surprise $\hat{m}$ — which are consistent estimators of p and $\mathbb{E}(x|x \in D)$ — even in the presence of (this form of) heterogeneity, an implication of the symmetry between information endowment and forced disclosure. Put differently, just as information endowment does not affect the characterization of the cost– nor does forced disclosure. On the other hand, the minimum disclosure in the sample is no longer a consistent estimator of τ, and we develop two robust estimators for the disclosure threshold.

3.2 Median-Minimum estimator

Assume that reporting units i may be observed, although the econometrician does not know whether a unit is misspecified. However, misspecified units will feature a minimum disclosure that converges to the lower bound of x as T, i.e., the number of periods, increases and will thus appear as a group of outliers. Borrowing from the treatment of outliers embedded in quantile estimation, we thus propose to take the median across firms of the minimum within the firm. We refer to this estimator as the median-minimum estimator, and it is defined as follows:

$$\hat{\tau}_{mm} = med(\{\hat{\tau}_i\}), \tag{19}$$

where $\hat{\tau}_i$ is the minimum estimator derived earlier for firm i.

If the fraction of misspecified firms is not too large, that is, $\theta_0 < .5$, the median of the minimum will almost surely be the threshold of a firm that was not misspecified as the sample size increases. Therefore, the median-minimum $\hat{\tau}_{mm}$ is a consistent estimator of τ, and we define $\hat{c}_{mm}$ as the estimator of disclosure costs that uses this estimator for the threshold. Note also that $\hat{\tau}_{mm}$ still converges to τ at a rate greater than $\sqrt{N}$ and, therefore, the asymptotic errors in equation (14) apply to this setting using the total number of observations N. To see this, note that the minimum within firm converges to the true minimum (for non-misspecified firms) at rate strictly greater than $\sqrt{T}$ - as shown in proposition 4 - and the median across firms converges to the true median at rate $\sqrt{I}$, where I corresponds to the number of firms (Koenker 2005). The latter is the source of a finite-sample loss of efficiency relative to the minimum over the entire sample absent heterogeneity, since the convergence rate would have been strictly greater than $\sqrt{I}$ if we had used the minimum across firms. Nevertheless, asymptotically, the rate of convergence of $\hat{\tau}_{mm}$ is strictly greater than $\sqrt{IT}$ while the rate of convergence of $\hat{p}$ and $\hat{m}$ is exactly $\sqrt{IT}$. Thus, the median estimator allows us to improve the quality of the estimation of $\hat{p}$ and $\hat{m}$ using a larger sample, by relying on equation (19).

3.3 Non-parametric threshold estimator

There are situations, however, where the previous procedure is not feasible. For example, it may be the case that all firms are potentially misspecified $\theta_0 = 1$, even though $\theta_1 < 1$, or that a reporting unit may not be observable. For example, a firm or an industry may change over time in manners that may change its propensity to make a forced disclosure, causing our notion of reporting unit to be imperfectly mapped to firms. However, even in this case, forced disclosure will imply that the p.d.f. of observed disclosures will be steep (discontinuous with a large enough sample) around the disclosure threshold. Hence, we propose the following non-parametric threshold estimator

$$\hat{\tau}_{np} = argmax_z \hat{f}'(z),$$

where $\hat{f}'$ is a non-parametric estimate of the p.d.f of observed disclosures. We refer to the estimated cost with this procedure as $\hat{c}_{np}$.

Note that $\hat{f}'$ can be computed with any statistical package, by using a non-parametric fit of the density of disclosures and then evaluating its numerical derivative over each sample disclosure observation z, that is, $\hat{f}'(z) \equiv \hat{f}(z'') - \hat{f}(z')/(z'' - z')$ where z'' is the closest disclosure greater than z and z' is the closest disclosure lower than z. The non-parametric fit in STATA, for example, offers a command to evaluate the estimated density over all points of the sample. It is also well-known that the estimation of a discontinuity in the threshold converges at rate greater than $\sqrt{N}$ and thus does not affect the asymptotic properties of the estimator (Chu and Cheng 1996). However, for small finite samples, this procedure may select the inflection point of the distribution of x.

3.4 Dynamics

We expand the baseline model to a multi-period setting and derive the identification of the per-period cost if costs may change over time. This is a special form of heterogeneity in cost caused by dynamics. As an example, a natural presumption might be that the initial disclosure may reveal more proprietary information, and therefore, may be more costly than subsequent disclosures. In principle, any of the estimators may be used as input for the estimation of the dynamics of the model. However, for practical purposes, dynamics imply time-variation in costs so that sample-size considerations will require the use of the pooled estimators introduced earlier in this section.

Time is indexed by $t = 0, \ldots$. At the beginning of period t, investors and managers observe h_t, the public information. Indexing all variables by the history h_t, we assume that the private signal x_t is drawn from a continuous distribution with p.d.f. $g_t(.)$ and c.d.f. $G_t(.)$, and the disclosure cost $c(h_t)$ may be history-dependent. We define the random variable d_t that is equal to 1 if the manager discloses his private signal and 0 otherwise. At the end of each period, the earnings e_t are realized. Investors are risk-neutral and price the expected firm value at $P(d_t, h_t)$. As before, assume that the manager maximizes the current market price netted out of current and future expected disclosure costs discounted at $\beta \in (0, 1)$.[17] We define the random variable d_t that is equal to 1 if the manager discloses his private signal and 0 otherwise, and denote $p(h_t)$ as the equilibrium probability that information is disclosed.[18]

At the beginning of each period t, the manager values the firm as a function of the current expected cash flows at the end of period t and the future expected cash flows as

$$\begin{aligned} V(h_t) &= p(h_t)(\mathbb{E}(x_t|x_t > \tau(h_t)) - c(h_t)) + (1 - p(h_t))\mathbb{E}(x_t|ND) \\ &\quad + \beta\mathbb{E}_t(\Sigma_{i=t+1}^{\infty}\beta^{i-t-1}(p(h_i)(\mathbb{E}(x_i|x_i > \tau(h_i)) - c(h_i)) + (1 - p(h_i))\mathbb{E}(x_i|ND)) \\ &= -p(h_t)c(h_t) - \beta\mathbb{E}_t(\Sigma_{i=t+1}^{\infty}\beta^{i-t-1}p(h_i)c(h_i)). \end{aligned}$$

It follows that the manager discloses if and only if

$$x_t - c(h_t) - \beta\mathbb{E}_t(\Sigma_{i=t+1}^{\infty}\beta^{i-t-1}p(h_i)c(h_i)|d_t = 1) \geq \mathbb{E}(x_t|ND) - \beta\mathbb{E}_t(\Sigma_{i=t+1}^{\infty}\beta^{i-t-1}p(h_i)c(h_i)|d_t = 0).$$

Rearranging the terms, this inequality is identical to the baseline model with a modified cost function:

$$\underbrace{P(\{x_t\})}_{=x_t} - C(h_t) \geq \underbrace{P(ND)}_{=\mathbb{E}(x_t|ND)}$$

The modified cost is given by:

$$C(h_t) = c(h_t) + \beta(\mathbb{E}_t(\Sigma_{i=t+1}^{\infty}\beta^{i-t-1}p(h_i)c(h_i)|d_t = 0) - \mathbb{E}_t(\Sigma_{i=t+1}^{\infty}\beta^{i-t-1}p(h_i)c(h_i)|d_t = 1)).$$

The cost $C(h_t)$ represents the total incremental cost incurred by the manager over current and future periods, that is, considering the effect of a disclosure in the current period on the probability of future disclosures. This term combines the per-period cost $c(h_t)$, which is the economic primitive to be estimated, with the effect on future costs. Therefore, if the baseline model is applied to a dynamic setting where $c(h_t)$ varies with h_t, we may be recovering this total cost $C(h_t)$ instead of the actual cost $c(h_t)$ incurred

[17] Or, equivalently, as in Einhorn and Ziv (2008), future disclosure costs must be substracted from the gross market price.

[18] This empirical model nests the model of einziv08, provided the states are defined to match their model. In their model, firms have uncertainty about information endowment that is correlated over time and disclosure costs. They show that the state variable is the history of past disclosures and earnings, so we would estimate this model by defining states in terms of groups of past histories - where each past history points to a market belief about information endowment. Naturally, their model is continuous and estimation requires discrete groups so that we have enough data points - so one would increase the precision of each group as a function of sample size.

in the period. The intuition for these effects are similar to discrete choices over experience goods, such as in Rust (1987) and the conditional choice probability estimation of Hotz and Miller (1993), in which a decision made in the present may affect utility in future periods.

We borrow from this literature the insight that the value function (which incorporates current and future utility) can be inverted to recover per-period utility, as long as we can estimate the value function for each state where a history h_t may fall. Let $H = \{h_1, ..., h_N\}$ be a complete set of distinct states, where by "state" we mean a subset of possible histories. We require each state to have a probability bounded away from zero conditional on any history. For $h \in H$ and $h' \in H$, $q_{h,h',1}$ and $q_{h,h',0}$ be the transition probabilities defined respectively as the probability to be in the next period in state h' given that the current state is h and the manager discloses and the probability to be in the next period in state h' given that the current state is h and the manager does not disclose. We define $p(h)$, the probability that the manager discloses in state h.

The three variables in the generalized disclosure model $C(h)$, $c(h)$ and $V(h)$ represent respectively the total disclosure cost incurred by the manager in state h, the instantaneous cost in state h if the manager discloses, and the manager's expected firm value in state h and are solutions to

$$\forall h, \quad C(h) = c(h) + \beta\Sigma_{h'\in H}(q_{h,h',0} - q_{h,h',1})V(h'), \tag{20}$$

$$V(h) = -p(h)c(h) + p(h)\beta\Sigma_{h'\in H}q_{h,h',1}V(h') + (1-p(h))\beta\Sigma_{h'\in H}q_{h,h',0}V(h'). \tag{21}$$

Proposition 6 *Let* $\mathbf{c} = (c(h_1), \ldots, c(h_N))'$, $\mathbf{C} = (C(h_1), \ldots, C(h_N))'$, $\mathbf{Q}_k = (q_{h_i,h_j,k})_{i,j}$ *for* $k = 0, 1$. *Then:*

$$\mathbf{c} = (\mathbf{I_N} + \beta(\mathbf{Q}_0 - \mathbf{Q}_1)(I_N - \beta\mathbf{Q}_0)^{-1}\mathbf{P})\mathbf{C}, \tag{22}$$

where $\mathbf{I_N}$ *is the* $N \times N$ *identity matrix and* $\mathbf{P}$ *is a matrix with diagonal term* $(p(h_1), \ldots, p(h_N))$ *and zero off-diagonal.*

Some straightforward observations follow directly from equation (23). To begin with, per-period costs are not identified absent knowledge of the discount factor β.[19] This property is intuitive, as a given level of disclosure would not distinguish between a high per-period cost but high discounting, or a low per-period cost but low discounting. Naturally, a disclosure model is unlikely to be the appropriate model to estimate a discount factor and, in such situations where the cost is incurred by the firm and netted out of the price, it would be preferable to pre-estimate β externally from an observed rate of return

[19]The non-identification of dynamic discrete choice models has been discussed in prior literature, see Rust (1994) and Magnac and Thesmar (2002), and the intuition for the lack of identification of the discount factor is similar to these studies. Note, however, that our model is somewhat different from the discrete choice framework, in that it features a game in which the non-disclosure price is derived from rational expectations — in other words, what the agent is expected to choose will affect his payoff from those choices.

on a firm's securities. Second, if we know β, there exists a simple estimation procedure to consistently estimate per-period costs by using sample estimates of the right-hand side of equation (23). An estimator of $\mathbf{Q}_k$ can be obtained from frequency estimates of the transition probabilities, using the sample ratio of occurrences of state h_j after state h_i to the total occurrences of state h_i. An estimator for $\mathbf{P}$ can be obtained by estimating $\hat{p}(h)$ as the frequency of disclosure conditional on state h since it is a conditional choice probability (Hotz and Miller 1993). Lastly, we know from equation (21) that $C(h)$ can be estimated using one of the previous estimators on the subsample of observations in state h.

4 Empirical application

In this section, we provide an empirical application of the estimation in the context of voluntary management forecasts. We are interested in asking three questions. First, how large are the estimated costs, or benefits, of disclosure? Second, are the estimates related to observable firm characteristics, such as competition, growth or litigation risk? Third, does the theory explain the cross-section of management forecast disclosures?

Management forecast disclosures present a natural setting for our estimation procedure because they are voluntary (i.e., managers are not required by law to issue forecasts) and are arguably issued with incentives related to stock price. It is also an open question as to whether disclosure costs are at the root of disclosure behavior, since forecasts generally present highly aggregated information and might delay competitive actions by at most a year (for annual forecasts). Since our purpose is also to illustrate firm-level versus pooled estimates, we have chosen quarterly forecasts to have, potentially, a useable number of observations per firm. In this context, we attempt to provide evidence as to the existence and magnitude of disclosure costs, if any.

4.1 Management forecast data selection

We begin the sample using quarter period ends from 2004 through 2015 for two reasons. First, the implementation of Regulation Fair Disclosure (Reg. FD) in the U.S. closed private channels of communication to analysts, and thus greatly increased the proportion of forecasting firms for reasons unrelated to costs. Second, in previous years, management forecasts were not systematically collected by the First Call Company Issued Guidance (CIG) database but after 2003, the requirement by Sarbanes-Oxley to record transcripts of conference calls has greatly improved forecast archives.[20] Our data comes from the Compustat and I/B/E/S databases. First, we obtain quarterly earnings announcements of U.S.

[20]Prior to 2003, many forecasts would have been made during unrecorded conference calls, thus leading to systematic omitted data for smaller firms that are less likely to trigger follow-up press releases.

firms from I/B/E/S and match the data to the Compustat fundamentals database. Second, we merge the data by earnings announcement periods to management earnings forecast data, if available, from I/B/E/S guidance (which extends the (formerly) First Call Company Issues Guidelines (CIG) database). As a result, each quarterly earnings announcement may or may not have a corresponding management forecast disclosure, and such variation in forecast disclosure is necessary for our cost estimator.

Table 1 provides the specific details of our sample selection procedure. The sample starts with all quarterly earnings per share (EPS) observations from I/B/E/S with a quarter period ending between the years 2004 and 2015. We rely on the I/B/E/S file, as opposed to Compustat, because I/B/E/S lists the actual date that earnings are announced. We obtain data for 15,967 unique firms with 557,828 firm-quarter observations. We retain observations with non-missing values of quarterly EPS (both adjusted and unadjusted), ibtic, cusip, ticker, pdicity, pends, anndats, lag anndats, and we calculated for each firm its standard deviation of EPS.[21] Because management forecasts of EPS are typically adjusted, we use the ratio of unadjusted to adjusted EPS to convert these forecasts to raw forecasts. Adjusted forecasts are problematic because, for firms that had splits, the variance of adjusted forecasts will be declining over time. In cases of zero adjusted EPS, which can occur because EPS are zero or because of two-digit rounding given very large splits, we use the nearest available adjustment factor, dropping observations that have no adjustment factor. These steps result in 7,960 unique firms with 218,321 firm-quarter observations. We scale EPS by lagged assets per share, which requires asset data from Compustat and the number of shares from CRSP. Our estimation procedure requires the distribution of EPS to be fairly stable across time periods and firms, so we use raw EPS numbers scaled by the firm's standard deviation of EPS and lagged asset per share throughout our analysis.[22] This step reduces our sample to 7,581 firms with 204,535 firm-quarters.

Because our model requires firms that have at least one disclosure and one non-disclosure observation (i.e., period with a management forecast and one period without a management forecast), we further reduce our sample based on whether firms have at least one disclosure and at least one non-disclosure. As shown in table 1 line 4, we eliminate 5,474 out of 7,581 firms (72.2%) that chose to never disclose, which reduces our sample to 2,107 firms with 70,945 firm-quarter observations. We then eliminate 128 out of 7,581 firms (1.7%) that chose to always disclose, which leaves us with 1,979 firms with 68,204 firm-quarter observations. Our final step is to eliminate observations that are clearly incompatible with the theory, removing firms whose expected earnings are significantly negative. As illustrated in step 6 of

[21] We use the firm's standard deviation of EPS in subsequent tests to normalize the variance between firms.

[22] Another option might have been to use EPS adjusted for splits and stock dividends; however, the historical variance of past adjusted EPS after splits have occurred is, by construction, very low. Another popular scaling is lagged price. We have not used it here, however, because this alternative approach carries its own caveats because the daily variance of stock prices may make this alternative measure very noisy.

table 1, performing this test requires each firm to have at least two quarters of management disclosures (to calculate a mean and standard deviation for a t-statistic for expected earnings), which eliminates 399 out of 7,581 firms (5.3%) and reduces our sample to 1,580 firms and 55,641 observations. We then eliminate 297 out of 7,581 firms (3.9%) where expected earnings are significantly negative given disclosure, which results in a sample of 1,283 firms and 44,946 firm-quarter observations spanning 2004 and 2015. Additionally, we perform analysis to determine whether past disclosure costs are likely to influence the firm's propensity to issue subsequent forecasts. To do this, we separate our sample into an estimation period spanning 2004 to 2014 and a holdout period of 2015. We re-estimate the firm's disclosure costs and control variables over the estimation period, then use these values for regressions during our holdout period of 2015. Our holdout sample contains 778 unique firms with 2,998 associated quarters (see line 8 of table 1).

Table 2 provides summary descriptive statistics of our full sample, which is comprised of 44,946 earnings announcements for 1,283 firms from 2004 through 2015. In panel A of table 2, our sample covers a maximum of 49 quarters with an average (median) of 41.18 (48) earnings announcements.[23] The average raw EPS (not adjusted for stock splits) of \$0.37 per share is above the median \$0.27 per share. Conditional on observing a management forecast, the median EPS of \$0.29 is higher than the median EPS conditional on no management forecasts of \$0.26, which is consistent with the strategic selection of good news to forecast.

In panel B of table 2, 17,754 observations out of the 44,946 earnings announcement observations are paired with a management forecast disclosure. Part of the reason for this low frequency of forecasts is due to the fact that our full sample of 44,946 observations is a broad, inclusive sample. The median number of periods with forecast reports is 11, whereas the average number of periods with a forecast is higher at 15.8 or 39.5% of the total periods of earnings announcements. The median forecast of EPS is \$0.25, which is lower than the average forecast \$0.36. Finally, to implement the estimator, we measure management forecast surprise and subtract out the expected component of the forecast. All earnings and forecast variables are scaled by the firm's standard deviation EPS and by lagged assets per share, i.e., lagged at/(shroutm1_EAl/1,000).

We utilize a regression analysis to account for the unexpected component of EPS due to management forecasts. Management forecast surprise is calculate as the difference between the predicted value from

[23]The maximum number of quarters per firm is 49, due to two firms (with tickers CDTC and TTWO) reporting EPS for five quarters in the years 2004 and 2010, respectively

an EPS regression with management forecast:

$$\begin{aligned} \text{EPS}_i &= \beta_0 + \beta_1 \text{EPS\_lag}_i + \beta_2 \text{consensus}_i + \beta_3 \text{ManagementForecast val\_1}_i \\ &\quad + \beta_4 \text{ManagementForecast val\_2}_i + U_i^1 \end{aligned}$$

and the predicted value from an EPS regression without management forecast:

$$\text{EPS}_i = \beta_0 + \beta_1 \text{EPS\_lag}_i + \beta_2 \text{consensus}_i . + U_i^2$$

All variables are scaled by the standard deviation of EPS and then lagged assets. The variable consensus is the market consensus provided by the I/B/E/S Guidance file as the variable "mean_at_date". So ManagementForecast val_1 and val_2 come directly from the IBES CIG Guidance File. ManagementForecast val_1 corresponds to either the single value of the management forecast or the beginning of range of the management forecast. ManagementForecast val_2 is the upper bound for range forecasts. We ran the estimation above separately for range forecasts (such that val_2 is non-missing) and point or open-ended forecasts.

One advantage of using a regression analysis to calculate forecast surprise is that the regression controls for any firm-level forecasting bias (through the intercept), and we can isolate the unexpected component due to forecasting by taking the difference between two fitted values. To facilitate the interpretation of our results when graphing these statistics, we multiply the management forecast surprise by the firm-level standard deviation of EPS to yield a percentage of lagged assets.[24] Accordingly, the mean and median management forecast surprise is 0.41% and 0.22% of lagged assets.

In Panel C of table 2, we report the descriptives for several variables argued to be related to disclosure costs, proprietary costs, general disclosure environment, and litigation risk. Our sample includes firms that are very different in size, growth opportunities, capital intensity and leverage.

We provide more details next on the choice of the explanatory variables. Our first set of variables has been argued to be related to proprietary costs. The variables include the Compustat Herfindahl-Hirsch index (HHI), computed based on revenue for firms in Compustat (public firms) by year and SIC code.

HHI is the average Herfindahl-Hirschman Index over the sample period, calculated for each year and by industry SIC code as the sum of the square of each firm's revenues (revt); this measures the market concentration, i.e., the closer a market is to being a monopoly, the higher the market's concentration (and the lower its competition), the higher this index. Markets become concentrated when some firms

[24]Note that since our EPS variable is earnings per share, then our scaling by lagged assets per share results in the per share term cancelling out to yield earnings per assets; consequently, we interpret our results as a % of lagged assets.

acquire a strategic competitive advantage (Demsetz 1973).

A firm's disclosures are associated with higher proprietary costs in more highly concentrated product-markets because competitors may extract sensitive information or might attract entry of new rivals (Darrough and Stoughton 1990). Alternatively firms' disclosures in highly competitive environments can benefit their competitors (Verrecchia 1990). There is significant cross-sectional variation among characteristics related to competition.

CAPEX_i is defined as the average percentage of capital expenditure divided by total assets for firm i. High capital expenditures have been traditionally used as a measure of barriers to entry (Sutton 1991). Typically firms that are capital intensive have acquired a competitive advantage and therefore a forecasting firm bears less proprietary cost. Alternatively capital intense firms might operate in more sensitive business sectors and any disclosure can be used by their competitors.

TURNOVER_i is defined as the average asset turnover, i.e., the ratio of revenues over total assets for firm i. This variable measures the efficiency of a firm at turning its investments into revenues. If the profits of a firm are essentially driven by volume (over charging high margins), proprietary costs are lower because barriers to entry tend to be stronger.

MTB_i is defined as the average ratio of market capitalization to equity for firm i. Growth opportunities indicate availability of profitable investments such as new products or creation of barriers to entry (Gaver and Gaver 1993). The greater the growth opportunities, the higher the proprietary cost, the more reluctant managers are to reveal information that could dissipate the value of these opportunities.

MKTCAP_i is the market capitalization computed as (previous fiscal year price*shrout)/1,000,000 expressed in billions. We include the log of the firm's market capitalization (in billions) as an additional control for size in our regressions. Larger companies tend to have more captive clients, and competitors benefit less from a forecasting competitor. However, smaller firms might rely more frequently on external capital markets and, to reduce costs (Jensen and Meckling 1976), may be willing to disclose more.

NSEGMENTS_i is the number of business segments per firm, obtained from the busseg variable from Compustat. On one hand, firms with fewer segments have a higher disclosure cost they reveal a finer level of detail than firms with more segments. On the other hand, firms that are more focused have a lower disclosure cost than diversified firms because the processing costs are lower when the information is less aggregated and there is evidence of diversification discount by the markets.

Our second set of variables capture the general disclosure environment and firms' volatility:

LEVERAGE_i is defined as the average ratio of total liabilities over total assets for a firm i. Leveraged firms might choose to disclose to reduce their agency costs (Jensen and Meckling 1976), or the disciplinary effect of debt could reduce disclosure by mitigating the free cash flow problem (Jensen 1986). Leverage is a measure of financial fragility.

INSTITUTIONAL$\%_i$ ownership (the percentage of the company's aggregate common stock held by institutions, obtained from the shares variable from the Thomson Reuters Institutional 13F Holdings database), which controls for a firm's corporate governance. Institutional investors typically ask for more disclosures and might influence the release of management forecasts.

AUDIT$_i$ is an indicator variable equal to one for Big N auditors, based on the Compustat variable au values from one to eight), as Lang and Lundholm (1993) find that firms using BigN auditors tend to have better disclosures.

NUMEST$_i$ is the number of IBES analyst forecasts for a firm, as Lang and Lundholm (1993, 1996) document that firms with a higher number of analysts tend to have higher disclosure quality.

EARNVOL$_i$ is the earnings volatility, i.e., the standard deviation of EPS over the previous 12 quarters, as Waymire (1985) finds an association between higher earnings volatility and frequency of management forecasts.

BETA$_i$ is estimated from daily returns during the quarter from a regression of firm excess returns (returns minus the risk-free rate) on the market return minus the risk-free rate) to proxy for market risk.

Our third set of variables is intended to capture the firm's litigation risk.

LAWSUIT$_i$ is an indicator variable equal to one if the firm had a class action lawsuit filed against it within ninety days of the end of the quarter (data is obtained from the Stanford Securities Litigation Database. Firms to reduce their litigation costs tend to disclose more (Skinner 1994).

LITIGATE$_i$ is an indicator variable equal to 1 if the firm is in the high litigation industries of biotechnology (SIC 2833-2836, 8731-8734), computers (SIC 3570-3577, 7370-7374), electronics (3600-3674), and retail (5200-5961) as defined in Ajinkya *et al.* (2005).

Finally, we consider the stickiness in the management forecasts. NSinceDISC$_i$ is the lagged quarter of the firm's most recent management forecast disclosure; for each firm, we also calculate MAX_NSinceDISC, or the maximum of NSinceDISC. There is extended empirical evidence that firms that did not disclose over a long time horizon are less likely to disclose because they bear higher disclosure costs. Furthemore, to ease the interpretation of our measures, we report LAG ASSETS$_i$ (in millions), which is the lagged value of total assets, scaled by 1,000, or at/1,000.

To run the cross-sectional regressions, we use two measures of quarterly firm-level cost estimates. The quarterly firm-level cost estimates are derived by assuming that a firm incurs the same cost every quarter and to derive the cost, we calculate the estimator of the disclosure threshold, the disclosure frequency and the average of the forecast surprises on all the firm-level observations available on the sample period from 2004 to 2015. The first firm-level measure is COST$_i$, the cost estimates using the minimum of the forecasts to estimate the threshold and the second is COST_mod, which uses instead of the minimum forecast surprise the minimum forecast surprise less the difference between the second lowest forecast surprise minus the minimum. As a result, COST_mod contains an adjustment to mitigate the effect of any potential upward bias.

In table 2 Panels D and E, we report the univariate Spearman correlation matrix of economic characteristics of the firm. Most of the correlations between HHI and the other firm characteristics are highly significant at the 1% level, while the CAPEX is significantly correlated at the 1% level with all the other variables. More concentrated industries have a higher capital intensity, are more efficient at managing their assets to turn them into revenues, face more growth opportunities, and are less risky (with lower leverage). Note that our COST and COST_mod variables exhibit strong correlations at the 1% level for nearly all variables (except for BETA and LAWSUIT).

4.2 Tests

Our first set of tests is a firm level analysis of each firm's disclosure COST. Recall that the inputs to our disclosure cost formula are scaled by the firm-level standard deviation of EPS, then scaled again by the firm's lagged assets per share. We can only estimate a firm-level cost if the firm has at least one period where it issued a management forecast and one period where it did not issue a management forecast (as shown in table 1 lines 4 and 5). Table 3 panel A reports basic descriptives at the firm-level for 1,283 unique companies, while Panels B and C focuses on the 500 firms with positive disclosure COST and the 783 firms with negative COST, respectively. A forecast occurs on average 42.4% of the time. The average (median) forecast surprise is positive, at 0.464% (0.934%) of the standard deviation of EPS and of lagged assets . The average (median) of the minimum forecast surprise is −9.420% (−2.968%) of the standard deviation of EPS and of lagged assets.

For our firms, the average (median) COST estimate is negative, i.e., −8.390% (−1.357%) of the standard deviation of EPS and of lagged assets; therefore, the average (median) firm experiences benefits to disclosure. COST_mod also displays negative costs for the average and median firm slightly more negative, as expected, i.e., −13.542% and −3.203% of the standard deviation of EPS and of lagged assets.

Note that all COST estimates (which uses the minimum forecast surprise), and COST_mod estimates (which uses a modified-minimum forecast surprise) are winsorized at the 1^{st} and 99^{th} percentiles to reduce the effects of outlier firms.

We also split the cost estimates according to whether the costs are positive or negative. In panel B of table 3, firms with positive costs tend to exhibit fewer forecasts than in the full sample, at 39.4%

for the average firm (versus 42.4% for the full sample of firms) and 26.9% for the median firm (versus 35.0% for the population with estimated costs). Firms with positive costs exhibit a higher unconditional management forecast surprise for the average firm of 3.276% of the standard deviation of EPS and of lagged assets (versus 0.464% for the population with estimated costs).

Panel C of table 3 shows the statistics for firms with negative costs. These firms tend to forecast more than other firms (with 13.73 quarters on average with forecasts). The standard deviation of COST in the negative cost sample of 26.06% is larger than the dispersion of 11.371% for the positive costs. Some firms experience large benefits to disclosure, which can be as high as 111.1% and 178.4% of the standard deviation of EPS and of lagged assets for the measures COST and COST_mod respectively.

Note that our COST estimates require the minimum and mean management forecast surprise, as well as the frequency of forecasts. To illustrate these inputs for each of our 1,283 firms, we re-scale these measures (by multiplying by the standard deviation of EPS) so that we can interpret these measures as a percentage of lagged assets. As shown in Charts 1A and 1B, the distributions of the minimum and the average forecast are skewed. Chart 1A shows that around 40% of our firms have a minimum forecast surprise ranging between −1% to 0% of lagged assets. The pattern for the modified minimum forecast surprise is similar to that of the minimum. Chart 1B shows that the management forecast surprise distribution is left-skewed as expected and the average management forecast is positive and ranges from 0% to 1% of lagged assets. The median of the management forecast surprise is at 0.2% of lagged assets. The forecast frequency shown in Chart 1C varies widely across firms and is not smooth. A large portion of our firms disclose less frequently, as around 17% of firms disclose less than 10% of the time, while an equal percentage of firms disclose 10% to 20% of the time. We also re-scale our COST estimates so they can be graphed in Chart 1D as a % of lagged assets. More than 50% of our firms have a small benefit to disclosure in the range of −5% to 0% of lagged assets. However, as shown in Chart 1D, the positive costs are relatively concentrated in the range of 0% to 5% of lagged assets. Since most firms never issue forecasts, we expect most of the sample to have larger costs that we cannot measure directly if costs are very large. Chart 1G shows that the mean costs are more extreme compared to median costs.

Since our estimator makes no distributional assumptions about the forecast surprise, we compare our estimator to a parametric estimator that assumes normality in Bertomeu *et al.* (2015a). Chart 1E compares our estimator versus the BBT estimator. The parametric estimator takes on only positive values. Within the positive cost range, about 37% of our sample has COST values between 0% to 10%, while the 42% of the parametric estimator falls into the same range and in general, the BBT cost estimates returns larger magnitudes. This is intuitive, as the BBT estimator explains all disclosures as a result of disclosure costs.

In Panel D of table 3, we conduct the tests presented in section 2.4 to determine the firms in our sample that are not consistent with our model when we use our more conservative empirical content test based on the bounds of the probability to be uninformed. We classify remaining firms into three categories: firms consistent with a Dye model, hereafter "Dye firms", defined as firms with probability of information endowment significantly below one, firms consistent with a Verrecchia model, hereafter "Verrecchia firms," defined as firms with disclosure costs significantly greater than zero. Based on the COST or the COST_mod we have 110 firms, i.e., 8.6% of the firms in our sample, that are not consistent with the model, a large proportion of Dye firms, roughly 50%, a range of $20\% - 25\%$ of Verrecchia firms and a range of $14\% - 16\%$ of firms that have a disclosure behavior that is not clearly determined.

4.3 Economic determinants of disclosure costs

In our first set of tests, we explore the determinants of a firm's estimated disclosure costs. We run the following OLS regression in table 4:

$$\begin{aligned} \text{COST}_i \quad = \quad & \beta_0 + \beta_1\text{HHI} + \beta_2\text{CAPEX}_i + \beta_3\text{TURNOVER}_i + \beta_4\text{MTB}_i \\ & +\beta_5\text{LOG MKT CAP}_i + \beta_6\text{NSEGMENTS}_i + \beta_7\text{LEVERAGE}_i \\ & +\beta_8\text{INSTITUTIONAL\%}_i + \beta_9\text{AUDIT}_i + \beta_{10}\text{NUMEST}_i + \beta_{11}\text{EARNVOL}_i \\ & +\beta_{12}\text{BETA}_i + \beta_{13}\text{LAWSUIT}_i + \beta_{14}\text{LITIGATE}_i + \beta_{15}\text{MAX\_NSinceDISC}_i + U_i. \end{aligned}$$

We consider different specifications for the variable COST. In Panel A, we examine all values of COST and COST_mod (both positive and negative). In Panel B, we examine only positive COST and COST_mod values in columns (1) to (4). In columns (5) to (8), we run our regressions using negative values of COST and COST_mod.

In Panel A of table 4 with both positive and negative cost estimates, the explanatory power ranges from an adjusted R-square of 15.77% to 18.82%. The regressions with positive COSTS (in Panel B) have a maximum adjusted R-square of 25.68% (in column (3)), whereas the regressions with negative COSTS has a higher maximum adjusted R-square of 29.64% (in column (7)). Therefore, our variables explain more of the variation for firms with a benefit to disclosure compared to those that bear disclosure costs.

In table 4 Panel A, we find strong evidence that more concentrated firms have lower disclosure benefits (coefficient on HHI is positive at the 1% level), possibly because disclosure may attract new rivals (Darrough and Stoughton 1990). Both HHI and CAPEX are positively related to disclosure costs at the 1% level for the overall sample as well as for those firms with negative COSTs. However, both of these coefficients are insignificant for positive COST firms.

Our findings in the overall sample of firms (in Panel A of table 4) tend to be largely attributed to the negative cost sample (in Panel B), and Panel B reveals differences in the positive and negative cost firms. The variables MTB and LOG MKT CAP exhibit different coefficients for the positive versus negative COST firms. For positive cost firms, higher growth opportunities captured by high MTB are associated with large cost of disclosure (with a significantly positive coefficient on MTB ranging from 0.005 to 0.011), which is consistent with proprietary costs; the coefficient on LOG MKT CAP is not statistically significant except in column (1). In contrast, in the negative cost sample, firms with high MTB and low market cap have more negative COST estimates (i.e., more benefits to disclosure).

In the overall sample in Panel A of table 4, the coefficient on LEVERAGE is positive and significant at the 1% level, suggesting that leveraged firms have higher disclosure costs; however, once we examine the positive cost firms alone, we find a negative coefficient on LEVERAGE, which suggests that firms with higher leverage experience less cost to disclosure. In the overall sample, the positive coefficient on MAX_NSinceDISC suggests that firms that have waited a long time since disclosing have higher disclosure costs. This result is driven by our negative cost sample; once we examine our positive cost sample, we find a negative coefficient, which suggests that firms that have recently disclosed bear higher costs, potentially because markets expect a subsequent forecast and firms tend to disclose information that is less beneficial.

4.4 Out-of-sample analysis

In this section, we explore whether disclosure costs can influence a firm's subsequent disclosure behavior. Previous empirical literature conjectures that earlier and more accurate disclosures about competitors' outcomes (such as earnings forecasts) can allow firms to adjust their production schedules more rapidly and improve their own positions at the expense of the forecasting firm. We test whether past estimated costs are predictive of the firms' future disclosure choice.

We divide our sample into an estimation period spanning 2004-2014 and a holdout period of 2015. We estimate firm-level values of the inputs and disclosure COST, as well as firm control variables, as of 2014 to ascertain the effects on the firm's 2015 quarterly disclosures. The purpose of this test is to demonstrate in a holdout sample that the theory helps explain future management forecasts. Our prediction is that firms that bear these costs are less likely to make future disclosures, which means we predict a negative relation between COST and disclosure.

Panel A of table 5 shows the descriptive statistics for our estimates of disclosure COST and COST_mod over the restricted estimation period of 2004-2014. For the estimation, we have 1,274 firms, which is fewer than the 1,283 in our overall sample shown in table 3; additionally, overall COST estimates are

−7.927%, which is less negative than the −8.390% for the full sample. The COST estimates appear similar in both time periods.

We run the following logit regression in Panel B of table 5:

$$\begin{aligned}\text{DISCLOSE}_{2015} \quad = \quad & \beta_0 + \beta_1\text{COST}_{2014} + \beta_2\text{FREQ}_{2014} + \beta_3\text{MIN\_SURPRISE}_{2014} \\ & + \beta_4\text{MEAN\_SURPRISE}_{2014} + U_i.\end{aligned}$$

DISCLOSE is an indicator variable equal to one if the firm issued a management forecast disclosure for the quarter and zero otherwise. We predict the coefficient on COST to be significantly negative because firms that bear high disclosure costs will be less likely to issue a forecast. We incorporate the inputs to our disclosure cost estimate (FREQ, MIN_SURPRISE, and MEAN_SURPRISE) into our regression to test whether COST had any predictive power above and beyond its individual components. The results in Panel B of table 5 confirm the negative relation between DISCLOSE and COST. The coefficient on COST is significantly negative at the 1% level in all columns in Panel B. Of the inputs to disclosure cost, only FREQ is positively related to disclosure, but this finding is not surprising given that disclosure behavior tends to be "sticky" such as firms that disclosed frequently in the past are more likely to continue disclosing in future periods. The findings in this panel support our hypothesis that the past disclosure cost is predictive of future disclosure behavior.

Next, we explore our disclosure regression by incorporating separately three sets of explanatory variables that capture proprietary costs, environmental variables, and litigation costs (all estimated from 2004-2014). We use the following logit model in table 5 in Panel C:

$$\begin{aligned}\text{DISCLOSE}_{2015} \quad = \quad & \beta_0 + \beta_1\text{COST}_{2014} + \beta_2\text{HHI}_{2014} + \beta_3\text{CAPEX}_{2014} + \beta_4\text{TURNOVER}_{2014} \\ & + \beta_5\text{MTB}_{2014} + \beta_6\text{LOG MKT CAP}_{2014} + \beta_7\text{SEGMENTS}_{2014} + \beta_8\text{LEVERAGE}_{2014} \\ & + \beta_9\text{INSTITUTIONAL\%}_{2014} + \beta_{10}\text{AUDIT}_{2014} + \beta_{11}\text{NUMEST}_{2014} \\ & + \beta_{12}\text{EARNVOL}_{2014} + \beta_{13}\text{BETA}_{2014} + \beta_{14}\text{LAWSUIT}_{2014} \\ & + \beta_{15}\text{LITIGATE}_{2014} + \beta_{16}\text{NSinceDISC}_{2014} + U_i.\end{aligned}$$

When we incorporate the disclosure cost estimates from 2004-2014 into our disclosure regression, we find that the significantly negative coefficient on COST remains (see columns (1) through (6) of table 5). Of these variables, MTB exhibits significance with disclosure, and firms with higher MTB (and higher growth options) are more likely to disclose forecasts, and NSEGMENTS is significantly negatively correlated with disclosure, which could be explained by agency costs related to diversified firms.

Firm-specific litigation risk is another plausible determinant of a firm's disclosure behavior. After only controlling for both LAWSUIT and LITIGATE (Columns (3) and (4)), the coefficients on COST and COST_mod in table 5 Panel C remain significantly negative. We find significantly positive coefficients at 1% level on LAWSUIT and LITIGATE. The positive association between LAWSUIT and LITIGATE, and management forecast disclosure does not mean that disclosure triggers more lawsuits. In this context, firms that are exposed to a higher litigation risk will effectively have more lawsuits, and therefore, these firms disclose more to reduce their litigation costs. In fact, we find support for this scenario in our previous tests of the negative association between COST and LITIGATE in table 4 Panel A. We also examine the coefficient on NSinceDISC, which captures the number of periods since the firm last disclosed a management forecast. Since the firm's disclosure tendency tends to be "sticky", then the negative coefficient on NSinceDisc is expected, as the longer the period since disclosure, the less likely the firm will disclose. The pseudo R-square jumps from $3.1\% - 3.3\%$ to $29.7\% - 33.01\%$. Controlling only for the stickiness of past disclosure does not affect the coefficient on COST, which remains significantly negative. Once we combine our litigation variables, various environmental variables and the remaining variables into a single regression, the coefficient on COST is insignificant. Within the environmental variables, firms with higher earnings volatility (EARNVOL) are reluctant to make forecasts. This finding is consistent with the COST results in table 4, which show that firms with higher earning volatility also bear higher disclosure COST. These regressions highlight which variables affect the strength of the COST and disclosure relation; namely, the number of past disclosures is likely to be a dominant determinant of the future disclosure decision.

4.5 Pooled estimation

We illustrate the use of pooled estimation using the management guidance data, re-estimating disclosure costs within categories in which the observations are assumed to share equal costs and uncertainty about information endowment. We consider three different choice of pools. We pool our estimation of costs by the "stickiness" in the management forecast disclosures, the number of operating segments and firms' exposure to litigation risk. We conduct this pooled analysis by re-estimating the cost using all 44,946 firm-quarter observations within each group level, using the median-minimum estimator and the non-parametric threshold estimation.

For our first grouping, we consider another important characteristic of management forecasts: the number of quarters since the firm's most recent disclosure, also known as disclosure "stickiness" over time. We study whether the stickiness of the most recent disclosure can affect disclosure cost, and we predict that the longer the firm has waited to disclose, the higher the cost of disclosure. To address this issue, we take into account variation in the management forecasts' stickiness, proxied by the number of quarters of non-disclosure. Most firms issue forecasts less than half of the time (as shown in Panel B of table 2), and a significant proportion of occasional forecasters have only a single forecast. Therefore the number of observations per history might vary widely.

We create the variable NSinceDISC as the lagged quarter of the firm's most recent quarterly forecast, or the quarter since last observing DISCLOSE = 1. Because of our methodology, there are several cases in which NSinceDISC will have missing values.[25] We have values of NSinceDISC for 38,145 observations (and missing or 0 values for 6,801 observations), and we group the observations into four categories representing quarters where the most recent disclosure was (a) one quarter ago, (b) two to three quarters ago, (c) four to eight quarters ago, and (d) nine to forty-six quarters ago. We have determined these four groups so that the number of observations per groups is relatively homogenous. The first bucket with the most recent disclosure one quarter ago is by far the bucket with the highest number, i.e., $17,288$ observations but the other groups have a more comparable number of observations, roughly $5,000$ observations for the two next groups and $10,492$ observations for the last group.

Table 6 Panel A shows that cost estimates are negative for firms that disclose one quarter ago and generally become positive (indicating disclosure costs) for those with a long time since disclosure with both cost estimators for the pooled estimations. For the remaining groups (with the last disclosure more than a quarter ago) with positive costs, the stickiness in management forecasts does not seem to change the magnitude of the costs. This table suggests that if disclosure benefits exist, they may be present for firms that disclose frequently. The group with the most recent disclosures, i.e., a quarter ago, is consistent with the Dye model and we can set identify a confidence interval for q within the range $3.5\% - 19.0\%$ or $11\% - 19.0\%$ based on the median-minimum and non-parametric estimations respectively.

We conduct the pooled estimation on stickiness using the dynamic model in Panel C of table 6. The stickiness model strongly suggests to use the dynamic model since, by construction, a time-series will feature transition across states that depend on disclosure behavior - unlike segments or litigation risk which may stay similar during the period. We define each group (or state) as the number of periods since the last disclosure, up to group 12 which includes all firms that have not disclosed for 11 or more periods. We set the discount factor at $\beta = 0.9$. Standard-errors and confidence intervals were computed by bootstrap, by resampling a data set composed of the same number of firms 1,000 times and estimating costs for each sample. Based on the median-minimum cost estimators, the first group with last disclosures

[25]For example, if the firm did not issue a disclosure in quarter 1 and issued a disclosure in quarters 2 and 3, then NSinceDISC will have values for quarters 1, 2, and 3 of missing, 0, and 1. If the firm did not issue a disclosure in quarters 1 and 2, then issued a disclosure in quarters 3, then did not issue a disclosure in quarter 4, then NSinceDISC will have values of for quarters 1, 2, 3, and 4 of missing, 0, 0, then 1. Note that we have 6,801 observations that have either missing or 0 values, which are considered non-disclosures and excluded from the pooled analysis.

made a quarter ago displays a disclosure benefit but not significantly different from zero and the non-parametric threshold estimator returns a statistically disclosure benefit of 5.1% of the standard deviation of EPS and of lagged assets in the same group. The groups with the last disclosures occurring between two quarters and five quarters have significant positive disclosure costs with the median-minimum cost estimator and not statistiscally significant with the non-parametric estimator. Frequent forecasters seem to experience disclosure benefits and occasional forecasters tend to have disclosure costs.

For our second grouping, we investigate whether disclosure costs differ based on the number of business segments. We obtain from Compustat the number of business segments, and we conjecture that firms with fewer segments might bear more disclosure costs (because their disclosure is less aggregated than for firms with more segments). We form four groups based on the number of segments. In panel A of table 7, firms with less than three segments have negative disclosure costs (i.e., a disclosure benefit). Firms with one segment have disclosure benefits of roughly 5.6% (3.3%) of the standard deviation of EPS and of lagged assets with the median-minimum (non parametric) cost estimator. These benefits are slightly lower if firms have between two or three segments of roughly 1.5% (0.7%) of the standard deviation of EPS and of lagged assets with the median-minimum (non parametric) cost estimator. Those disclosure benefits are unexpected from the proprietary costs hypothesis but in line with potential agency costs in opaque diversified conglomerates. Firms with more than three segments have disclosure costs or benefits, depending on the cost estimators used. There is weak evidence that firms with more than four segments have disclosure costs. Across both estimations, firms with less than three segments behave like Dye firms. We can set identify a confidence interval for q within the range $53.6\% - 55.1\%$ or $53.1\% - 55.1\%$ based on the median-minimum estimation and non-parametric estimation respectively for the firms with one segment whereas the confidence interval for q is larger for the firms between two or three segments, i.e., $39.9\% - 63.2\%$ or $27.3\% - 63.2\%$ based on the median-minimum estimation and non-parametric estimation respectively.

For our last grouping, we separate the observations depending on litigation risk. We consider two variables to conduct the pooled estimations based on LAWSUIT and then based on LITIGATION. For LAWSUIT in Panel A of table 8, we have three groups based on LAWSUIT: the first group without any lawsuits, the second with one lawsuit and the last one with at least two lawsuits, and two groups based on LITIGATION in Panel C of table 8, the industries with a higher litigation risk and the remaining ones. Both the median-minimum and the non-parametric cost estimates exhibit higher disclosure benefits for industries with higher litigation risk 6.4% and 4.56% of the standard deviation of EPS and of lagged assets with the median-minimum cost estimator and the non parametric cost estimator. A robust finding across both estimations for LITIGATION is that the group exposed to high litigation risk is consistent with a Dye model and the confidence interval for the probability to be uninformed q is within the range $49.3\%-53.2\%$ or $48.1\%-53.2\%$ based on the median-minimum estimation and non-parametric estimation respectively. We find mixed evidence based on LAWSUIT that more litigation risk encourages firms to disclose. The cost estimates in the non-parametric threshold estimator are statistically negative costs, i.e., disclosure benefits, for firms with more than one lawsuit compared to the firms without lawsuits. The median-minimum estimator does not exhibit a clear distinctive pattern between firms with or without lawsuit.

5 Conclusion

In this study, we offer an empirical framework to estimate and tests several classic models of strategic voluntary disclosure. A key insight from the estimation is that we can test a joint model of disclosure costs (or benefits) and information endowment, and we can distinguish disclosure costs from uncertainty about information endowment as long as there is variation in observed disclosure. We also prove that frequency, average disclosure and minimum disclosure are three important pieces of evidence that must be combined and are not necessarily proxies for the friction when taken in isolation. This latter insight is crucially different from an existing literature that takes frequency of disclosure as a measure of frictions and suggests an entirely new approach to explaining disclosure data.

Using the framework, an econometrician may potentially answer new questions, such as

1. Can we reject disclosure theory and, if not, can we test theories of costly disclosure against uncertainty about information endowment?
2. How large are disclosure cost or benefits, if any? What are these costs associated with?
3. How do costs or benefits change over time, or in response to certain events?

Most of these questions can be answered with simple methods that can obtain the costs using standard statistical packages and without solving the disclosure game. In this respect, we have attempted (as much possible) to derive estimators that can be computed as easily as other measures used in empirical research, but anchoring the derivation of that measure as a consistent estimator of a theoretical parameter. We hope that this can make the parameter easier to interpret and help clarify the information about the frictions contained in disclosure data.

As an illustration, we used the estimators in the context of the (commonly-used) management guidance data. This is, however, only one of possible samples where it may be plausible to expect voluntary disclosure to apply. In fact, to the extent that the estimator may capture benefits, we should be able to apply the estimation more generally to any setting in which the manager has some choice to withhold, even if withholding may cause some future losses. Thus, the framework may apply more generally to voluntary press releases, non-GAAP numbers, disclosures of medical trials or other relevant information in 8K filings, or the decision to delay certain items in financial statements.

Appendix

Proof of Proposition 1: Let us define $q, q' \in (0,1)$ and $q \neq q'$ such that for $\xi = (q, h(x,e), c), \xi' = (q', h'(x,e), c') \in S$, $\Psi(\xi) = q$ and $\Psi(\xi') = q'$, where $\Psi : S \rightarrow (0,1)$ is a function. q and q' are observationally equivalent, i.e., $F_{Y,e}(.;\xi) = F_{Y,e}(.;\xi')$. For $x \geq \tau$and $\forall e$, $h(x,e) = \frac{f(e)f(x,e)}{1-q}$. From equations (10) and (10), we can obtain the following expressions:

$$\forall x < \tau, \int eh(x,e)de = x \int h(x,e)de$$
$$\forall e, \int_{-\infty}^{\tau} h(x,e)dx = f(e)(1 - \frac{p(e)}{1-q}).$$

We will now build a function $h'(x,e)$ that satisfies the above two equations for $x < \tau$ and has the properties of a probability density function. Let us define

$$\begin{aligned} h'(x,e) &= \alpha h(x,e) \text{ for } x < \tau, \text{ where } \alpha \in (0,1) \\ h'(x,e) &= \beta h(x,e) \text{ for } x > \tau \text{ for } \beta > 0. \end{aligned}$$

Substituting this function into equation (24),

$$\begin{aligned} \int_{-\infty}^{\tau} \alpha h(x,e)dx &= f(e)(1 - \frac{p(e)}{1-q'}) \\ \text{Rearranging, } \frac{1}{\alpha}(1 - \frac{p(e)}{1-q'}) &= 1 - \frac{p(e)}{1-q} \\ q' &= 1 - \frac{p(e)}{1 - \alpha(1 - \frac{p(e)}{1-q})}. \end{aligned}$$

We also need to choose β to verify that $h'(x,e)$ is a pdf:

$$\begin{aligned}
1 &= \int\int h'(x,e)dxde \\
1 &= \int\int_{-\infty}^{\tau} h'(x,e)dxde + \int\int_{\tau}^{+\infty} h'(x,e)dxde \\
1 &= \alpha(1 - \int\int_{\tau}^{+\infty} h(x,e)dxde) + \int\int_{\tau}^{+\infty} h'(x,e)dxde \\
\int\int_{\tau}^{+\infty} h'(x,e)dxde &= 1 - \alpha + \alpha\int\int_{\tau}^{+\infty} h(x,e)dxde \\
\int\int_{\tau}^{+\infty} h(x,e)dxde &= \frac{1}{\beta}(1 - \alpha + \alpha\int\int_{\tau}^{+\infty} h(x,e)dxde) \\
\beta &= \frac{1 - \alpha + \alpha\frac{\int p(e)f(e)de}{1-q}}{\frac{\int p(e)f(e)de}{1-q}} > 0
\end{aligned}$$

Given $h'(x,e)$ and q', from equation (12), $c = c'$.

We now determine the bounds on q. By construction, $q < 1 - \int p(e)f(e)de$. Given that $\mathbb{E}(x|x \leq \tau) < \tau$, and $c < 0$, and using equation (6), we have the following inequality:

$$\begin{aligned}
(1-q)G(\tau)\tau &> (\tau - c)(1-p) \\
(1-q)(1 - \frac{p}{1-q})\tau &> (\tau - c)(1-p) \\
q &> (1-p)(1 - \frac{\tau - c}{\tau}).
\end{aligned}$$

Rewriting the above inequality as a function of the observed distributions yields:

$$q > (1 - \int p(e)f(e)de)(1 - \frac{\tau - c}{\tau}).$$

If $c > 0$ and $q \in [0,1)$ or $q \in (0,1)$, let us take two values c and c' such that $c \neq c'$, then the two vectors ξ and $\xi' \in S$ returning c and c' in equation (12), are such that $F_{Y,e}(.;\xi) \neq F_{Y,e}(.;\xi')$. □

Proof of Proposition 2: Let us denote $ND =$ "$\tilde{d} = 0$" as a non-disclosure event. We demonstrate the empirical content of the joint theory by showing that the data generating process must satisfy $\mathbb{E}(e|ND) \geq 0$, so that any $F_{Y,e}$ with $\mathbb{E}(e|ND) < 0$ is inconsistent with any $\xi \in S$.

Note that $\mathbb{E}(e|ND) = \mathbb{E}(\mathbb{E}(e|x)|ND) = \mathbb{E}(x|ND)$.

$$\begin{aligned}
\mathbb{E}(x) = 0 &= \int_{-\infty}^{\tau} xg(x)dx + \int_{\tau}^{+\infty} xg(x)dx \\
0 &= (1-q)\int_{-\infty}^{\tau} xg(x)dx + (1-q)\int_{\tau}^{+\infty} xg(x)dx \\
0 &= (1-p)\mathbb{E}(x|ND) + p\mathbb{E}(x|x > \tau)
\end{aligned}$$

If $\mathbb{E}(x|x>\tau)>\mathbb{E}(x|ND)$ then from the above equation, $\mathbb{E}(x|ND)<0$.. Conversely, if $\mathbb{E}(x|ND)<0$, from the above equation $\mathbb{E}(x|x>\tau)>0$ and thus $\mathbb{E}(x|x>\tau)>\mathbb{E}(x|ND)$. By construction, for any threshold equilibrium, $\mathbb{E}(x|x>\tau)>\mathbb{E}(x|ND)$ and thus if this condition is violated, no ξ exists that rationalizes this property. □

Proof of Proposition 4: The proof of consistency is immediate. By continuity, the estimator $\hat{c}$ is consistent, satisfying

$$plim\ \hat{c} = plim\ \hat{\tau} + \frac{plim\ \hat{p}}{1-plim\ \hat{p}} plim\ \hat{m} = \tau + \frac{p}{1-p}\mathbb{E}(x_t|x_t \geq \tau) = c.$$

Let us derive the asymptotic variance next. First we show that $\hat{\tau}$ is a super-efficient estimator of τ and plays no role in $\sqrt{N}$ expansion. Let us define $J(.)$, the distribution of $x|x>\tau$. Let $\tilde{m}$ be a random variable that follows a binomial distribution $B(n,p)$.

$$\begin{aligned} \forall t>0,\ Prob(\sqrt{n}(\hat{\tau}-\tau)\leq t) &= \Sigma_{m=0}^{n} C_n^m p^m (1-p)^{n-m} Prob(\hat{\tau} \leq \frac{t}{\sqrt{n}}+\tau|m) \\ &= \Sigma_{m=0}^{n} C_n^m p^m (1-p)^{n-m} (1-(1-J(\frac{t}{\sqrt{n}}+\tau))^m) \end{aligned}$$

$Prob(\hat{\tau} \leq \frac{t}{\sqrt{n}}+\tau|m) = 1-(1-J(\frac{t}{\sqrt{n}}+\tau))^m = 1-exp(mlog(\frac{t}{\sqrt{n}}+\tau)) = \varphi(m)$. Thus $\varphi(m)$ is concave. Applying the Jensen inequality,

$$\forall t>0,\ Prob(\sqrt{n}(\hat{\tau}-\tau)\leq t) \leq 1-(1-J(\frac{t}{\sqrt{n}}+\tau))^{np}$$

$$\begin{aligned} 1-(1-J(\frac{t}{\sqrt{n}}+\tau))^{np} &= 1-\exp(nplog(1-J(\frac{t}{\sqrt{n}}+\tau)) \sim 1-exp(-npJ(\tau)-\frac{npt}{\sqrt{n}}j(\tau)) \\ &\sim 1-exp(-\frac{npt}{\sqrt{n}}j(\tau)). \end{aligned}$$

Taking the limit when $n\to+\infty$ yields $lim_{n\to+\infty} 1-exp(-\frac{npt}{\sqrt{n}}j(\tau)) = 0$. Thus $lim_{n\to+\infty} Prob(\sqrt{n}(\hat{\tau}-\tau)\leq t) = 0$.

Next, we denote x_i as the manager's information and d_i as the disclosure for each observation i, where $d_i=1$ if a forecast is issued or $d_i=0$ otherwise.

$$\hat{c} = \tau + \frac{\hat{p}}{1-\hat{p}}\hat{m}^s = \tau + \frac{\sum d_i/N}{1-\hat{p}}\frac{\sum d_i x_i}{\sum d_i} = \tau + \frac{1}{1-\hat{p}}\underbrace{\frac{\sum d_i x_i}{N}}_{\hat{w}}.$$

In what follows, let us denote $\tilde{x}$ as the random variable corresponding to the manager's private information and $\tilde{d} = 1$ if a forecast is issued $d = 0$ otherwise. The associated moments to these random variables are denoted $m = \mathbb{E}(\tilde{x}|\tilde{x} \geq \tau)$, $v_x = Var(\tilde{x}|\tilde{x} \geq \tau)$ and $\mathbb{E}(\tilde{d}\tilde{x}) = p\mathbb{E}(\tilde{x}) = pm$. Note that $\hat{p}$ and $\hat{w}$ are sample means, therefore, by the central limit theorem,

$$\sqrt{N}(\begin{pmatrix} \hat{p} \\ \hat{w} \end{pmatrix} - \begin{pmatrix} p \\ pm \end{pmatrix}) \to_d N(\mathbf{0_2}, \underbrace{\begin{pmatrix} Var(\tilde{d}) & cov(\tilde{d}, \tilde{d}\tilde{x}) \\ cov(\tilde{d}, \tilde{d}\tilde{x}) & Var(\tilde{d}\tilde{x}) \end{pmatrix}}_{\mathbf{V_0}}).$$

Simplifying this variance-covariance matrix and denoting $m = \mathbb{E}(\tilde{x}|\tilde{x} \geq \tau)$ and $v_x = Var(\tilde{x}|\tilde{x} \geq \tau)$,

$$\mathbf{V_0} = \begin{pmatrix} p(1-p) & (1-p)pm \\ (1-p)pm & p(v_{x+}(1-p)m^2) \end{pmatrix} \tag{23}$$

$$\begin{aligned} \text{because} \quad Var(\tilde{d}) &= p(1-p); \\ cov(\tilde{d}, \tilde{d}\tilde{x}^s) &= \mathbb{E}(\tilde{d}^2\tilde{x}) - \mathbb{E}(\tilde{d})\mathbb{E}(\tilde{d}\tilde{x}) \\ &= \mathbb{E}(\tilde{d}\tilde{x}) - \mathbb{E}(\tilde{d})\mathbb{E}(\tilde{d}\tilde{x}) = (1-p)\mathbb{E}(\tilde{d}\tilde{x}) \\ &= (1-p)p\mathbb{E}(\tilde{x}|\tilde{x} \geq \tau) = (1-p)pm; \end{aligned}$$

$$\begin{aligned} \text{and} \quad Var(\tilde{d}\tilde{x}) &= \mathbb{E}(\tilde{d}^2(\tilde{x})^2) - \mathbb{E}(\tilde{d}\tilde{x})^2 \\ &= \mathbb{E}(\tilde{d}(\tilde{x})^2) - (p\mathbb{E}(\tilde{x}|\tilde{x} \geq \tau))^2 \\ &= p\mathbb{E}((\tilde{x})^2|\tilde{x} \geq \tau) - (p\mathbb{E}(\tilde{x}|\tilde{x} \geq \tau))^2 \\ &= p(Var(\tilde{x}|\tilde{x} \geq \tau) + \mathbb{E}(\tilde{x}|\tilde{x} \geq \tau)^2) - (p\mathbb{E}(\tilde{x}|\tilde{x} \geq \tau))^2 \\ &= p(Var(\tilde{x}|\tilde{x} \geq \tau) + (\mathbb{E}(\tilde{x}|\tilde{x} \geq \tau))^2(1-p)) = p(v_{x+}(1-p)m^2). \end{aligned}$$

Next, note that $\hat{c} = G(\hat{p}, \hat{w})$ where $G(X) = \tau + \frac{z}{1-y}$ and $X = (y, z)$. Hence, applying the delta method,

$$\sqrt{N}(\hat{c} - c) \to_d N(0, \underbrace{AV_0A}_{\sigma_c^2})$$

such that $A = \frac{\partial G}{\partial X'}|_{X=(p,pm)} = (\frac{pm}{(1-p)^2}, \frac{1}{1-p})$. Therefore,

$$\begin{aligned}\sigma_c^2 &= (\frac{pm}{(1-p)^2}, \frac{1}{1-p}) \begin{pmatrix} p(1-p) & (1-p)pm \\ (1-p)pm & p(v_{x+}(1-p)m^2) \end{pmatrix} \begin{pmatrix} \frac{pm}{(1-p)^2} \\ \frac{1}{1-p} \end{pmatrix} \\ &= \frac{p(pm+m(1-p))^2 + (1-p)pv_x}{(1-p)^3}.\square\end{aligned}$$

Proof of Proposition 6: Combining equations (21) and (22), $\forall h \in H, V(h) = -p(h)C(h) + \beta\Sigma_{h' \in H} q_{h,h',0} V(h')$. This equation can be written in matrix form as $\mathbf{AV} = -\mathbf{C}$, where $\mathbf{V} = (V(h_1), \ldots, V(h_N))'$, $\mathbf{C} = (C(h_1), \ldots, C(h_N))'$ and $\mathbf{A}$ is given by:

$$\mathbf{A} = \begin{pmatrix} 1-\beta q_{h_1,h_1,0} & -\beta q_{h_1,h_2,0} & \ldots & -\beta q_{h_1,h_N,0} \\ -\beta q_{h_2,h_1,0} & 1-\beta q_{h_2,h_2,0} & \ldots & -\beta q_{h_2,h_N,0} \\ \vdots & \vdots & \ddots & \vdots \\ -\beta q_{h_N,h_1,0} & -\beta q_{h_N,h_2,0} & \ldots & 1-\beta q_{h_N,h_N,0} \end{pmatrix}. \quad (24)$$

The matrix $\mathbf{A}$ is strictly diagonally dominant because $\forall i, 1-\beta q_{h_i,h_i,0} > \Sigma_{j\neq i} q_{h_i,h_j,0} > \beta\Sigma_{j\neq i}|-q_{h_i,h_j,0}|$. Hence, it is invertible and implies that $\mathbf{V} = -\mathbf{A}^{-1}\mathbf{C}$. Note, then, that equation (21) can be written as $\mathbf{C} = \mathbf{c} + \mathbf{\Delta V}$, where $\mathbf{c} = (c(h_1), \ldots, c(h_N))'$ is the per-period cost vector and $\mathbf{\Delta}$ is a square matrix with term $\mathbf{\Delta}_{i,j} = \beta(q_{h_i,h_j,0} - q_{h_i,h_j,1})$. Combining these equations:

$$\mathbf{c} = \mathbf{C} - \mathbf{\Delta V} = (\mathbf{I_N} + \mathbf{\Delta A}^{-1}\mathbf{P})\mathbf{C}, \quad (25)$$

where $\mathbf{I_N}$ is the $N \times N$ identity matrix and $\mathbf{P}$ is the diagonal matrix with diagonal terms $\mathbf{P_{ii}} = p(h_i)$.$\square$

References

Ajinkya, B., Bhojraj, S. and Sengupta, P. (2005). The association between outside directors, institutional investors and the properties of management earnings forecasts. *Journal of accounting research*, **43** (3), 343–376.

Arcidiacono, P. and Miller, R. A. (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica*, **79** (6), 1823–1867.

Armstrong, C. S., Core, J. E., Taylor, D. J. and Verrecchia, R. E. (2011). When does information asymmetry affect the cost of capital? *Journal of Accounting Research*, **49** (1), 1–40.

Bertomeu, J., Beyer, A. and Taylor, D. J. (2015a). From casual to causal inference in accounting research: The need for theoretical foundations. *Available at SSRN 2694105*.

—, Evans III, J. H., Feng, M. and Tseng, A. (2015b). *Tacit collusion and voluntary disclosure: Theory and evidence from the US automotive industry*. Tech. rep., Technical report, University of Pittsburgh, Working Paper.

—, Ma, P. and Marinovic, I. (2015c). How often do managers withhold information? *Stanford GSB Working Paper*.

Beyer, A., Cohen, D. A., Lys, T. Z. and Walther, B. R. (2010). The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics*, **50** (2-3), 296–343.

—, Guttman, I. and Marinovic, I. (2014). Earnings management and earnings quality: Theory and evidence. *Available at SSRN 2516538*.

Botosan, C. A. (1997). Disclosure level and the cost of equity capital. *The Accounting Review*, **72** (3), 323–349.

Caskey, J. (2013). The pricing effects of securities class action lawsuits and litigation insurance. *Journal of Law, Economics, and Organization*, p. ews048.

—, Hughes, J. and Liu, J. (2012). Strategic informed trades, diversification, and cost of capital.

Chen, Q. and Jiang, W. (2006). Analysts' weighting of private and public information. *Review of Financial Studies*, **19** (1), 319–355.

Cheong, F. S. and Thomas, J. (2011). Why do EPS forecast error and dispersion not vary with scale? implications for analyst and managerial behavior. *Journal of Accounting Research*, **49** (2), 359–401.

Cheynel, E. (2013). A theory of voluntary disclosure and cost of capital. *Review of Accounting Studies*, **18** (4), 987–1020.

Chu, C. and Cheng, P. (1996). Estimation of jump points and jump values of a density function. *Statistica Sinica*, pp. 79–95.

Darrough, M. (1993). Disclosure policy and competition: Cournot vs. Bertrand. *The Accounting Review*.

— and Stoughton, N. M. (1990). Financial disclosure policy in an entry game. *Journal of Accounting and Economics*, **12**, 219–243.

Demsetz, H. (1973). Industry structure, market rivalry, and public policy. *Journal of Law and Economics*, pp. 1–9.

Dye, R. (2013). *Voluntary disclosure and the duty to disclose*. Tech. rep., Working Paper, Northwestern University.

Dye, R. A. (1985). Disclosure of nonproprietary information. *Journal of Accounting Research*, **23** (1), 123–145.

— (1986). Proprietary and nonproprietary disclosures. *Journal of Business*, **59** (2), 331–366.

Einhorn, E. and Ziv, A. (2008). Intertemporal dynamics of corporate voluntary disclosures. *Journal of Accounting Research*, **46** (3), 567–589.

Gaver, J. J. and Gaver, K. M. (1993). Additional evidence on the association between the investment opportunity set and corporate financing, dividend, and compensation policies. *Journal of Accounting and Economics*, **16** (1), 125–160.

Gayle, G.-L., Li, C. and Miller, R. A. (2015). Was sarbanes-oxley costly? evidence from optimal contracting on ceo compensation. *Evidence from Optimal Contracting on CEO Compensation (2015-08-20). FRB St Louis Paper No. FEDLWP2015-017.*

Gerakos, J. and Kovrijnykh, A. (2013). Performance shocks and misreporting. *Journal of Accounting and Economics*, **56** (1), 57–72.

— and Syverson, C. (2015). Competition in the audit market: Policy implications. *Journal of Accounting Research.*

Gow, I., Larcker, D. and Reiss, P. (2015). Causal inference in accounting research. *Unpublished Working Paper, Stanford University.*

Hotz, V. J. and Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, **60** (3), 497–529.

Hughes, J., Liu, J. and Liu, J. (2009). On the relation between expected returns and implied cost of capital. *Review of Accounting Studies*, **14** (2-3), 246–259.

Indjejikian, R. and Matějka, M. (2009). CFO fiduciary responsibilities and annual bonus incentives. *Journal of Accounting Research*, **47** (4), 1061–1093.

Jensen, M. C. (1986). Agency cost of free cash flow, corporate finance, and takeovers. *Corporate Finance, and Takeovers. American Economic Review*, **76** (2).

— and Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of financial economics*, **3** (4), 305–360.

Jung, W.-O. and Kwon, Y. K. (1988). Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting Research*, **26** (1), 146–153.

Kasznik, R. and Lev, B. (1995). To warn or not to warn: Management disclosures in the face of an earnings surprise. *The Accounting Review*, pp. 113–134.

Koenker, R. (2005). *Quantile regression*. 38, Cambridge university press.

Lang, M. and Lundholm, R. (1993). Cross-sectional determinants of analyst ratings of corporate disclosures. *Journal of Accounting Research*, **31** (2), 246–271.

Lev, B. and Penman, S. H. (1990). Voluntary forecast disclosure, nondisclosure, and stock prices. *Journal of Accounting Research*, pp. 49–76.

Liang, P. J. and Wen, X. (2007). Accounting measurement basis, market mispricing, and firm investment efficiency. *Journal of Accounting Research*, **45** (1), 155–197.

Magnac, T. and Thesmar, D. (2002). Identifying dynamic discrete decision processes. *Econometrica*, **70** (2), 801–816.

Marinovic, I. (2013). Internal control system, earnings quality, and the dynamics of financial reporting. *The RAND Journal of Economics*, **44** (1), 145–167.

Matějka, M. and Ray, K. (2014). Balancing incentive weights and difficulty of performance targets: Theory and evidence. *Available at SSRN 2480185*.

Matzkin, R. L. (2007). Nonparametric identification. *Handbook of Econometrics*, **6**, 5307–5368.

Penman, S. H. (1980). An empirical investigation of the voluntary disclosure of corporate earnings forecasts. *Journal of Accounting Research*, pp. 132–160.

Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society*, pp. 999–1033.

— (1994). Structural estimation of markov decision processes. *Handbook of econometrics*, **4**, 3081–3143.

Skinner, D. J. (1994). Why firms voluntarily disclose bad news. *Journal of Accounting Research*, **32** (1), pp. 38–60.

— (1997). Earnings disclosures and stockholder lawsuits. *Journal of Accounting and Economics*, **23** (3), 249–282.

Stocken, P. C. and Verrecchia, R. E. (2004). Financial reporting system choice and disclosure management. *The Accounting Review*, **79** (4), 1181–1203.

Sutton, J. (1991). *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*. MIT Press.

Verrecchia, R. E. (1983). Discretionary disclosure. *Journal of Accounting and Economics*, **5**, 179–194.

— (1990). Information quality and discretionary disclosure. *Journal of Accounting and Economics*, **12** (4), 365 – 380.

Wagenhofer, A. (1990). Voluntary disclosure with a strategic opponent. *Journal of Accounting and Economics*, **12**, 341–363.

Waymire, G. (1985). Earnings volatility and voluntary management forecast disclosure. *Journal of Accounting Research*, pp. 268–295.

Zakolyukina, A. A. (2012). Measure of intentional manipulation: A structural approach. *Unpublished Working Paper, Stanford University*.

Zhou, F. (2016). Disclosure dynamics and investor learning.

Table 1 – Sample Selection Procedures

Sample Selection Procedures

		N Firms	N Firm-Quarters
1.	Earnings announcements (EPS) and earnings announcement dates (anndats) are obtained from the I/B/E/S Earnings Announcement file for 2004-2015 for firms with US Dollars as the currency. The I/B/E/S file is used (as opposed to Compustat) because it lists the date earnings were announced (anndats). The maximum number of observations per firm is 49 (due to two firms with tickers CDTC and TTWO reporting EPS for five quarters in the years 2004 and 2010, respectively).	15,967	557,828
2.	Observations with non-missing values of quarterly EPS adjusted, quarterly EPS unadjusted (used to compute an adjustment factor), ibtic, cusip, ticker, pdicity, pends, anndats, lag anndats, lag EPS, and standard deviation of EPS.	7,960	218,321
3.	Observations with non-missing values of gvkey, total assets (at), and shares outstanding (shrout).	7,581	204,535
4.	Observations are matched to management forecasts (from the I/B/E/S guidance file) of quarterly EPS, when available. Remaining firms must have at least one quarter with a management forecast disclosure. This eliminates 5,474 firms with no management forecast disclosures.	2,107	70,945
5.	Remaining firms must have at least one quarter with no management forecast disclosure. This eliminates 128 firms with 100% management forecast disclosures.	1,979	68,204
6.	Observations in which there are at least two quarters with management forecast disclosures (to calculate whether expected earnings are consistent with the model). This eliminates 399 firms with 1 management disclosure only.	1,580	55,641
7.	Observations in which expected earnings given disclosure is consistent with the model. We exclude 297 firms in which expected earnings given disclosure are significantly negative at the 5% level, as this is not consistent with the model.	1,283	44,946
	Full Sample	**1,283**	**44,946**
8.	**Holdout Sample: Observations for quarters for 2015 only (using estimates of disclosure costs as of 2014).**	**778**	**2,998**

Table 2 – Descriptive Statistics for 44,946 Firm-Quarter Observations

Panel A. EPS Characteristics

	N	Mean	Median	St.Dev.	Min	Max
Number of Earnings Announcements per Firm	44,946	41.18	48.00	10.90	3.00	49.00
EPS (raw)	44,946	$ 0.37	$ 0.27	$ 0.91	-$ 46.80	$ 31.92
EPS (raw) When There is a Management Forecast Disclosure	17,754	$ 0.41	$ 0.29	$ 0.61	-$ 13.87	$ 14.38
EPS (raw) When There is No Management Forecast Disclosure	27,192	$ 0.35	$ 0.26	$ 1.06	-$ 46.80	$ 31.92

Panel B. Management Forecast Characteristics

	N	Mean	Median	St.Dev.	Min	Max
N Periods With Forecasts	44,946	15.80	11.00	13.66	2.00	47.00
N Periods For Each Firm	44,946	41.18	48.00	10.90	3.00	49.00
% of Periods with Management Forecasts = (N Forecasts/N Periods)	44,946	39.5%	29.3%	31.0%	4.2%	97.9%
Management Forecast of EPS (raw)	17,754	$ 0.36	$ 0.25	$ 0.54	-$ 13.00	$ 13.57
Forecast Surprise (scaled by standard deviation of EPS and lagged assets)	17,751	0.0038	0.0097	0.1239	-6.0673	3.1489
Firm-Level St.Dev. of EPS	44,946	38.8%	24.7%	56.0%	0.8%	1066.4%
Forecast Surprise (scaled by lagged assets)	17,751	0.0041	0.0022	0.0243	-1.1643	0.4960

Panel C. Firm Characteristics

	N	Mean	Median	St.Dev.	Min	Max
HHI	44,946	0.07	0.05	0.06	0.00	0.27
CAPEX	44,946	0.02	0.01	0.03	0.00	0.16
TURNOVER	44,946	0.27	0.23	0.19	0.00	0.99
MTB	44,946	2.75	2.01	2.92	-3.39	19.55
MKT CAP (in billions)	44,946	4.28	1.06	8.96	0.02	56.63
NSEGMENTS	44,946	3.83	3.00	3.17	1.00	10.00
LEVERAGE	44,946	0.52	0.52	0.23	0.08	1.16
INSTITUTIONAL%	44,946	0.58	0.71	0.36	0.00	1.00
AUDIT	44,946	0.71	1.00	0.45	0.00	1.00
NUMEST	44,946	11.03	9.00	8.34	1.00	80.00
EARNVOL	44,946	0.25	0.14	0.44	0.00	14.61
BETA	44,946	1.17	1.06	0.59	-5.45	10.68
LAWSUIT	44,946	0.01	0.00	0.09	0.00	1.00
LITIGATE	44,946	0.37	0.00	0.48	0.00	1.00
MAX_NSinceDISC	44,946	16.26	13.00	13.15	0.00	46.00
LAG ASSETS (in millions)	44,924	7.75	1.17	37.81	0.00	919.79

Table 2 – continued

Panel D. Spearman Correlations for 44,946 Firm-Quarter Observations

	COST	HHI	CAPEX	TURNOVER	MTB	MKT CAP	NSEGMENTS	LEVERAGE
HHI	0.0530***							
CAPEX	-0.0532***	0.0826***						
TURNOVER	-0.0707***	0.2949***	0.3263***					
MTB	-0.2002***	0.0283***	0.1769***	0.1340***				
MKT CAP	0.2045***	0.0342***	0.0768***	-0.1055***	0.3579***			
NSEGMENTS	0.246***	0.2697***	0.0729***	-0.0604***	-0.1079***	-0.0754***	0.2399***	
LEVERAGE	0.3244***	0.0410***	-0.0613***	-0.0063	0.0327***	0.2599***	0.2824***	
INSTITUTIONAL%	0.0978***	-0.0919***	-0.0583***	-0.1186***	0.0540***	0.2124***	0.0679***	0.0448***
AUDIT	-0.0148***	0.0130***	0.0891***	0.0741***	0.0605***	0.2593***	-0.0806***	0.0472***
NUMEST	0.1002***	-0.0077	0.2080***	-0.0480***	0.2343***	0.7153***	0.0751***	0.1290***
EARNVOL	0.2911***	0.0614***	0.0727***	0.1170***	-0.0378***	0.2642***	0.1344***	0.2300***
BETA	0.0009	0.0158***	0.0212***	-0.0013	-0.0063	-0.0097**	-0.0104**	-0.0070
LAWSUIT	-0.0013	-0.0019	0.0115**	-0.0119**	0.0123***	0.0235***	0.000	0.0011
LITIGATE	-0.2575***	-0.1046***	0.0809***	0.0951***	0.0936***	-0.1715***	-0.2628***	-0.2615***
MAX_NSinceDISC	0.1042***	0.0362***	-0.1045***	-0.0380**	-0.1075***	0.0574***	0.1512***	0.2010***
LAG ASSETS	0.3874***	0.0336***	-0.0318***	-0.1887***	-0.0036	0.8641***	0.3587***	0.4931***

*, **, *** indicates correlation is significant at the 10%, 5%, 1% level, respectively.

Panel E. Spearman Correlations for 44,946 Firm-Quarter Observations

	INSTITUTIONAL%	AUDIT	NUMEST	EARNVOL	BETA	LAWSUIT	LITIGATE	MAX_NSinceDISC
AUDIT	0.0920***							
NUMEST	0.1780***	0.2250***						
EARNVOL	0.0604***	0.1353***	0.1804***					
BETA	0.1027***	0.1196***	0.0344***	0.0774***				
LAWSUIT	-0.0001	-0.0138***	0.0419***	0.0103**	0.0059			
LITIGATE	-0.1857***	-0.0647***	0.0217***	-0.1939***	-0.0151***	0.0291***		
MAX_NSinceDISC	0.0245*	-0.0155***	-0.0800***	0.1222***	-0.0302***	-0.0112**	-0.1993***	
LAG ASSETS	0.1889***	0.2226***	0.6202***	0.3388***	0.0045	0.0263***	-0.2892***	0.1479***

*, **, *** indicates correlation is significant at the 10%, 5%, 1% level, respectively.

Table 2 – continued

Variable Definitions: EPS (raw) refers to a company's reported earnings per share EPS (unadjusted), as obtained from the I/B/E/S Earnings Announcement Database. Management forecasts of EPS are taken from the I/B/E/S CIG variable value_1, then converted to unadjusted values by multiplying by the ratio (EPS unadjusted/EPS adjusted). Management forecast surprise is calculate as the difference between the predicted value from an EPS regression with management forecast ($EPS_{scaled} = a + b_1 EPS\_lag_{scaled} + b_2 consensus_{scaled} + e$) and the predicted value from an EPS regression without management forecast ($EPS_{scaled} = a + b_1 EPS\_lag_{scaled} + b_2 consensus_{scaled} + b_3 Management\ Forecast\ val\_1_{scaled} + b_4 Management\ Forecast\ val\_2_{scaled} + e$) ; for the EPS regressions, all variables are scaled by the firm's standard deviation EPS and by lagged assets per share, i.e., lagged at/(shroutm1_EAl/1,000). CONSENSUS forecast is the mean forecast obtained from the variable mean_at_date from the I/B/E/S CIG Guidance file. DISCLOSE is an indicator variable equal to 1 if there is a Management Forecast issued and 0 otherwise. The following data items are obtained from Compustat unless otherwise noted and are calculated for each firm-quarter and each firm: HHI is the Herfindahl-Hirschman Index, calculated for each year and industry SIC code as the summation of the square of each firm's revenues (revt). CAPEX is capital expenditures (capex) divided by total assets (at). TURNOVER is revenues (revt) divided by total assets (at). MTB is the market value of equity divided by the book value of equity, or [(prcc_f*shrout)/1,000]/(at-lt); missing values of MTB are set to one. MKT CAP (in billions) is (previous fiscal year price *shrout)/1,000,000. NSEGMENTS is the number of business segments per firm, obtained from the busseg variable from Compustat. LEVERAGE is total liabilities (lt) divided by total assets (at); missing values of LEVERAGE are set to zero. INSTITUTIONAL% is calculated as the ratio for each firm and quarter of institutional shares / (shrout*1,000), where institutional shares are obtained from the Thomson Reuters Institutional 13F Holdings and shrout is the lagged shares outstanding from CRSP. AUDIT is an indicator variable equal to 1 for Big N auditors (Compustat variable au values from 1 to 8), and 0 otherwise. NUMEST is the number of analyst forecasts for a firm, obtained from IBES. EARNVOL is the standard deviation of EPS over the previous 12 quarters. BETA is estimated from daily returns during the quarter from a regression of firm excess returns (returns minus the risk-free rate) on the market return minus the risk-free rate. LAWSUIT is an indicator variable equal to 1 if the firm had a class action lawsuit filed against it within ninety days of the end of the quarter, and 0 otherwise; data is obtained from the Stanford Securities Litigation Database. LITIGATE is an indicator variable equal to 1 if the firm is in the biotechnology (SIC 2833-2836, 8731-8734), computers (SIC 3570-3577, 7370-7374), electronics (3600-3674), and retail industries (5200-5961). NSinceDISC is the lagged quarter of the firm's last management forecast disclosure. MAX_NSinceDISC is the maximum quarters of the firm's last management forecast disclosure. LAG ASSETS is the lagged value of total assets /1,000, or at/1,000. The variables HHI, CAPEX, TURNOVER, MTB, MKT CAP, and LEVERAGE are winsorized at the 1^{st} and 99^{th} percentiles.

Table 3 – Descriptive Statistics for 1,283 Firm-Level Observations

Panel A. Firms with Positive and Negative Cost Estimates

	N	Mean	Median	St.Dev.	Min	Max
St.Dev.(EPS)	1,283	0.36	0.22	0.59	0.01	10.66
Minimum Forecast Surprise	1,283	-9.420%	-2.968%	28.729%	-606.734%	38.617%
Modified Minimum Forecast Surprise	1,283	-15.123%	-4.724%	52.218%	-1207.059%	22.020%
Average Forecast Surprise	1,283	0.464%	0.934%	7.318%	-123.126%	79.174%
N Periods with Forecasts	1,283	13.84	9.00	12.65	2.00	47.00
N Periods for Each Firm	1,283	35.03	43.00	14.69	3.00	49.00
% of Periods with Forecasts	1,283	42.4%	35.0%	30.6%	4.2%	97.9%
COST ($\hat{c}$)	**1,283**	**-8.390%**	**-1.357%**	**24.463%**	**-111.1%**	**53.7%**
COST_mod ($\underline{\hat{c}}$)	**1,283**	**-13.542%**	**-3.203%**	**33.277%**	**-178.4%**	**50.4%**

Panel B. Firms with Positive Cost Estimates

	N	Mean	Median	St.Dev.	Min	Max
St.Dev.(EPS)	500	0.59	0.36	0.87	0.01	10.66
Minimum Forecast Surprise	500	-0.011%	0.628%	5.864%	-56.354%	38.617%
Modified Minimum Forecast Surprise	500	-1.606%	0.009%	9.616%	-160.209%	22.020%
Average Forecast Surprise	500	3.276%	1.848%	5.073%	0.302%	50.201%
N Periods with Forecasts	500	14.01	8.00	14.08	2.00	47.00
N Periods for Each Firm	500	37.03	46.50	14.16	4.00	49.00
% of Periods with Forecasts	500	39.4%	26.9%	32.1%	4.2%	97.9%
COST ($\hat{c}$)	**500**	**6.083%**	**1.975%**	**11.371%**	**0.0%**	**53.7%**
COST_mod ($\underline{\hat{c}}$)	**500**	**4.488%**	**1.519%**	**11.938%**	**-103.8%**	**50.4%**

Table 3 – continued

Panel C. Firms with Negative Cost Estimates

	N	Mean	Median	St.Dev.	Min	Max
St.Dev.(EPS)	783	0.21	0.17	0.18	0.01	1.48
Minimum Forecast Surprise	783	-15.429%	-6.609%	35.190%	-606.734%	-0.105%
Modified Minimum Forecast Surprise	783	-23.755%	-9.738%	64.960%	-1207.059%	-0.324%
Average Forecast Surprise	783	-1.331%	-0.066%	7.943%	-123.126%	79.174%
N Periods with Forecasts	783	13.73	10.00	11.65	2.00	47.00
N Periods for Each Firm	783	33.76	39.00	14.88	3.00	49.00
% of Periods with Forecasts	783	44.2%	38.5%	29.5%	4.2%	97.9%
COST (ĉ)	**783**	**-17.632%**	**-6.596%**	**26.060%**	**-111.1%**	**0.0%**
COST_mod (ĉ)	**783**	**-25.056%**	**-9.929%**	**37.201%**	**-178.4%**	**-0.2%**

Panel D. Sample Composition by Firm

	COST (ĉ)		COST_mod (ĉ)	
	N Firms	% Firms	N Firms	% Firms
Full Sample Composition:	1,283	100.0%	1,283	100.0%
A. Firms Not Consistent with Model	110	8.6%	110	8.6%
B. Dye Firms	516	40.2%	614	47.9%
C. Verrecchia Firms	344	26.8%	270	21.0%
D. “Ambiguous” Firms (not in A, B, or C)	313	24.4%	289	22.5%

Variable Definitions: St.Dev.(EPS) is each firm’s standard deviation of EPS. COST (ĉ) is calculated as the minimum forecast surprise + [frequency of management forecasts /(1-frequency of management forecasts)]*average forecast surprise. Modified Cost, or COST_mod (ĉ) is calculated using a modified minimum with the formula [minimum forecast surprise – (2nd lowest forecast surprise-minimum forecast surprise)] + [frequency of management forecasts /(1-frequency of management forecasts)]*average forecast surprise. Management forecast surprise is calculate as the difference between the predicted value from an EPS regression with management forecast and the predicted value from an EPS regression without management forecast; for the EPS regressions, all variables are scaled by the firm’s standard deviation EPS and by lagged assets per share or lagged at/(shroutm1_EAl/1,000), and the dependent variable is EPS, while the independent variables are EPS_lag, the IBES consensus, and the management forecast variables val_1 and val_2. The modified minimum forecast surprise is the [minimum forecast surprise – (2nd lowest forecast surprise-minimum forecast surprise)]. Firms are classified into one of four categories (A. Firms Not Consistent with the Model, B. Dye Firms, C. Verrecchia Firms, and D. “Ambiguous Firms” (not in A, B, or C)) based on tests described in sections 2.3 and 2.4.

Chart 1A. Minimum and Modified Minimum Management Forecast Surprise (as % of Lagged Assets) for 1,283 Firms with Positive and Negative Costs

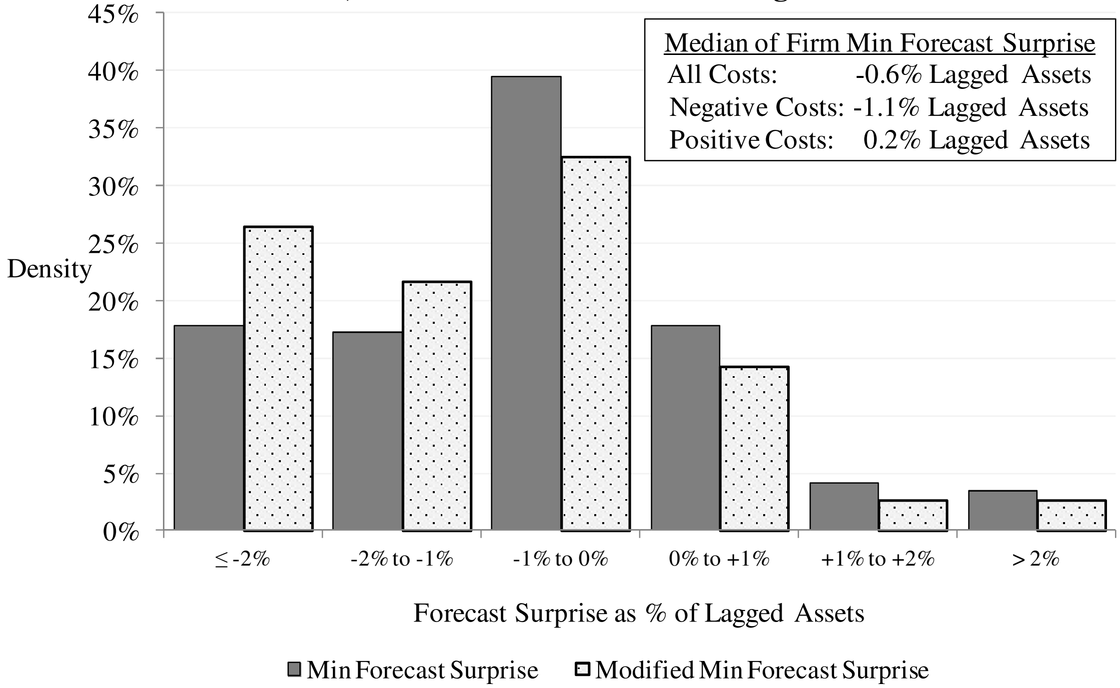

Chart 1B. Mean Management Forecast Surprise (as % of Lagged Assets) for 1,283 Firms with Positive and Negative Disclosure Costs

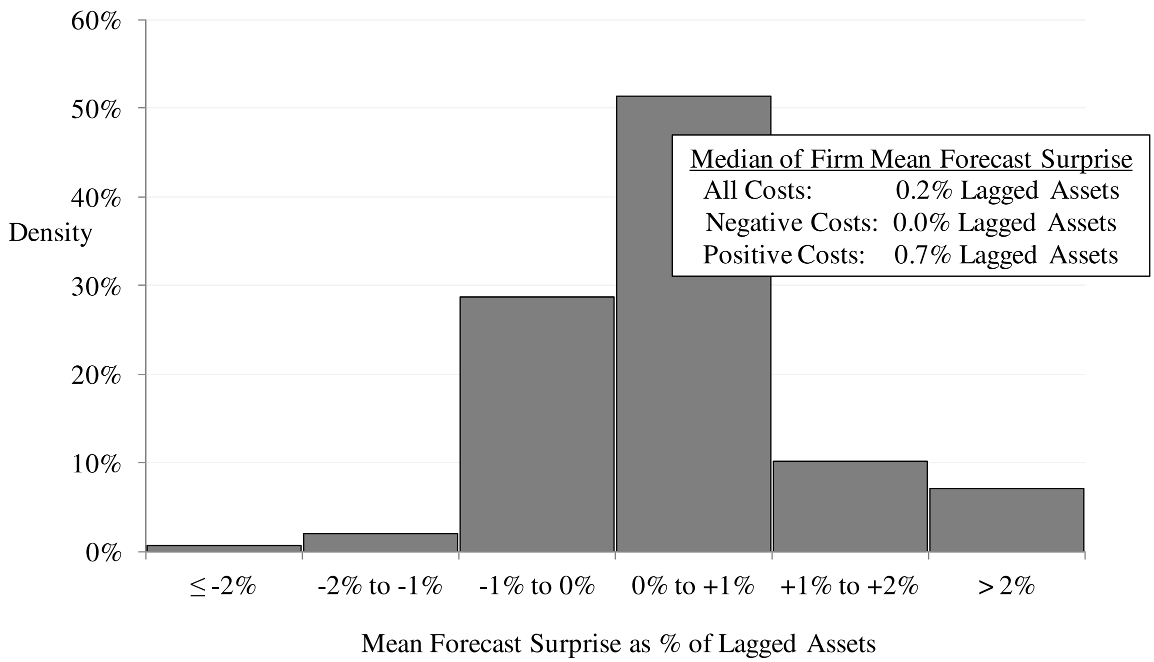

Chart 1C. Frequency of Management Forecasts for 1,283 Firms with Positive and Negative Disclosure Costs

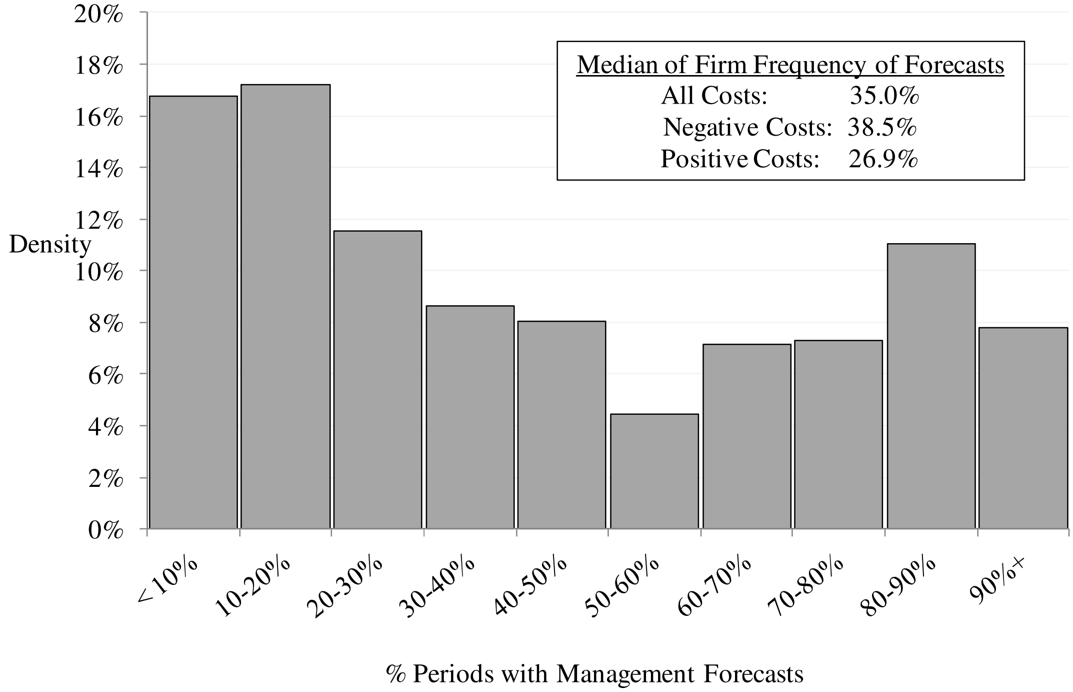

Chart 1D. Distribution Costs (as % of Lagged Assets) for 1,283 Firms with Positive and Negative Disclosure Costs

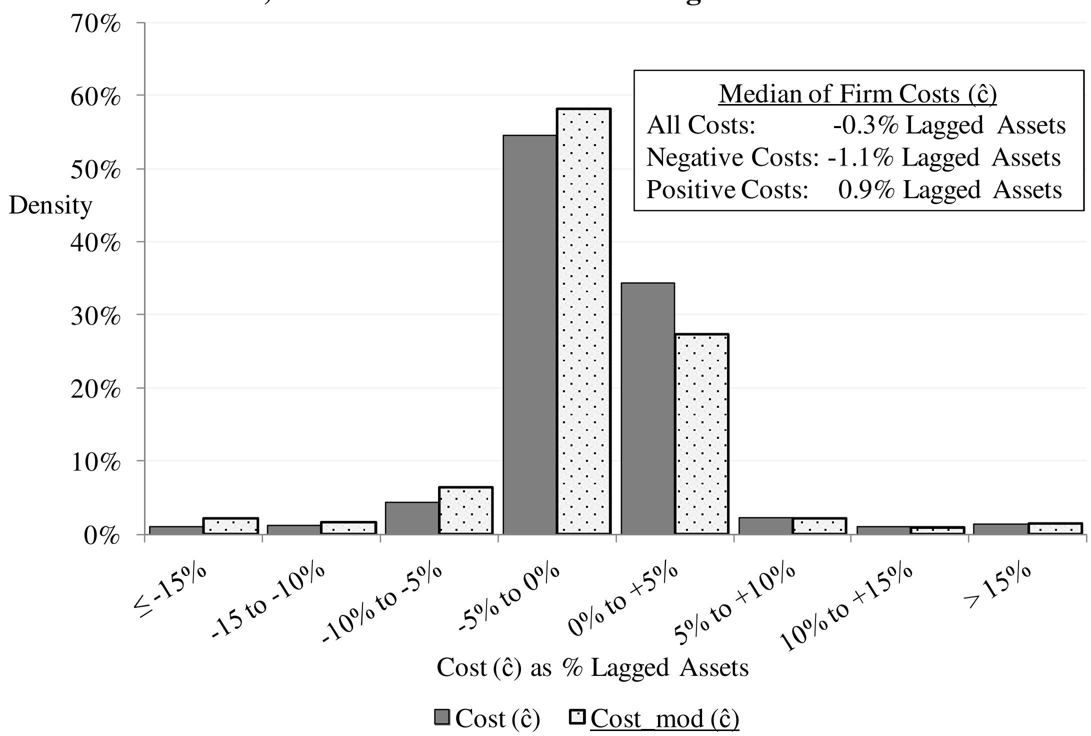

Chart 1E. Distribution of Cost (ĉ) vs. $Cost_{BBT}$ Estimator (as % of Lagged Assets) for 1,283 Firms with Positive and Negative Disclosure Costs

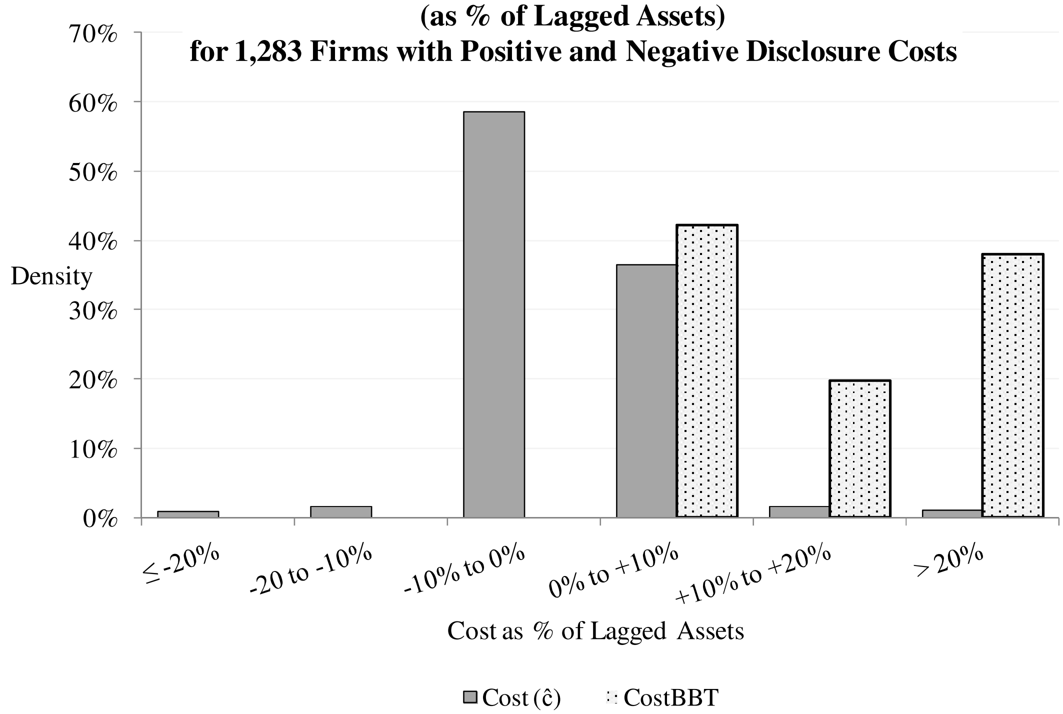

Chart 1F. Distribution of 1,283 Firms Based on Cost (ĉ)

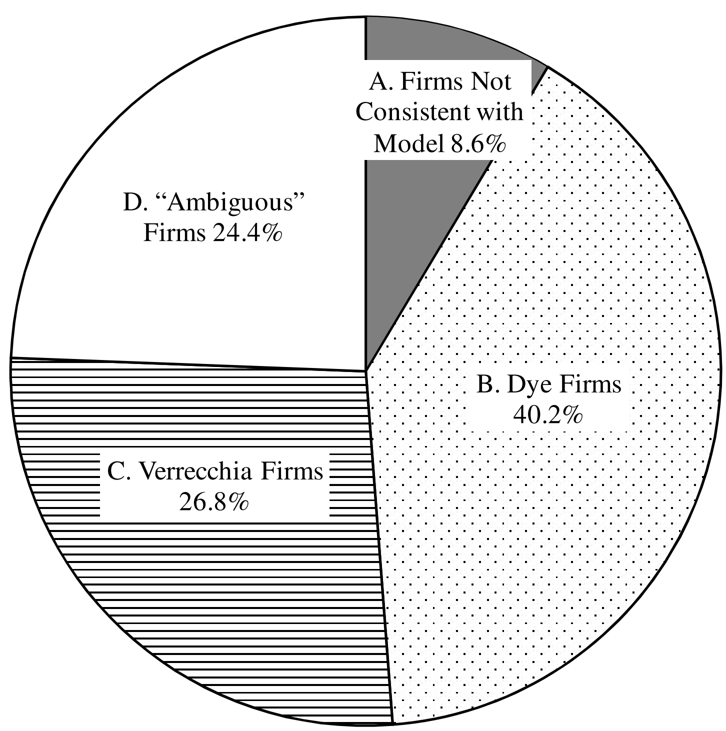

Chart 1G. Cost (as % of Lagged Total Assets)
Firm Mean and Median Values for 1,283 Firms

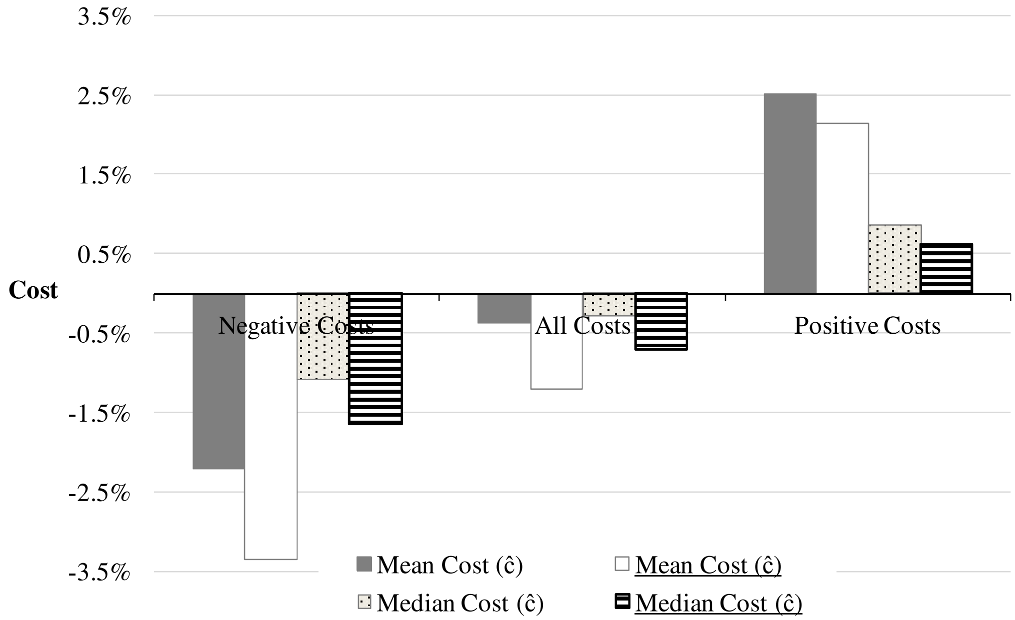

Table 4 – Firm-Specific Cost Determinants, Full Sample 2004-2015

Panel A. Firm-Level Regressions for 1,283 Firms with Positive and Negative Cost Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS Model Where the Dependent Variable is the Cost							
	COST (ĉ)	COST_mod (ĉ)	COST (ĉ)	COST_mod (ĉ)	COST (ĉ)	COST_mod (ĉ)	COST (ĉ)	COST_mod (ĉ)
Intercept	-0.145***	-0.219***	-0.147***	-0.222***	-0.124***	-0.189***	-0.147***	-0.223***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.002]	[0.001]	[0.000]	[0.000]
HHI	0.400***	0.421***	0.393***	0.412***	0.353***	0.356**	0.330***	0.323**
	[0.000]	[0.004]	[0.000]	[0.004]	[0.001]	[0.014]	[0.001]	[0.023]
CAPEX	0.714***	1.046***	0.712***	1.043***	0.692***	1.015***	0.714***	1.047***
	[0.003]	[0.001]	[0.003]	[0.001]	[0.004]	[0.002]	[0.003]	[0.001]
TURNOVER	0.037	0.079**	0.038	0.081**	0.043	0.088**	0.045	0.091**
	[0.196]	[0.034]	[0.181]	[0.031]	[0.130]	[0.019]	[0.119]	[0.017]
MTB	-0.031***	-0.043***	-0.031***	-0.043***	-0.030***	-0.041***	-0.029***	-0.040***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
LOG MKT CAP	0.021***	0.039***	0.022***	0.040***	0.018***	0.034***	0.016**	0.031***
(in billions)	[0.001]	[0.000]	[0.001]	[0.000]	[0.006]	[0.000]	[0.014]	[0.001]
NSEGMENTS	0.004	0.005	0.004*	0.005	0.003	0.004	0.003	0.004
	[0.103]	[0.138]	[0.091]	[0.125]	[0.189]	[0.245]	[0.188]	[0.244]
LEVERAGE	0.172***	0.221***	0.172***	0.221***	0.156***	0.199***	0.143***	0.180***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]
INSTITUTIONAL%	0.033	0.047*	0.034*	0.049*	0.023	0.033	0.023	0.033
	[0.104]	[0.075]	[0.091]	[0.066]	[0.249]	[0.207]	[0.248]	[0.205]
AUDIT	0.005	0.017	0.006	0.018	0.003	0.014	0.003	0.013
	[0.793]	[0.528]	[0.761]	[0.503]	[0.864]	[0.594]	[0.883]	[0.610]
NUMEST	-0.001	-0.001	-0.001	-0.002	0.000	0.000	0.000	0.000
	[0.531]	[0.432]	[0.516]	[0.420]	[0.980]	[0.926]	[0.733]	[0.802]
EARNVOL	0.075***	0.091***	0.075***	0.091***	0.068***	0.081***	0.061***	0.072**
	[0.000]	[0.002]	[0.000]	[0.002]	[0.001]	[0.004]	[0.003]	[0.011]
BETA	-0.028	-0.032	-0.028	-0.032	-0.026	-0.028	-0.024	-0.025
	[0.218]	[0.316]	[0.214]	[0.312]	[0.257]	[0.368]	[0.292]	[0.414]
LAWSUIT			0.063	0.080	0.062	0.079	0.069	0.089
			[0.222]	[0.222]	[0.224]	[0.223]	[0.179]	[0.174]
LITIGATE					-0.042***	-0.059***	-0.037**	-0.052**
					[0.007]	[0.007]	[0.015]	[0.015]
MAX_NSinceDISC							0.002***	0.002***
							[0.000]	[0.000]
N	1,283	1,283	1,283	1,283	1,283	1,283	1,283	1,283
Adj. R-Square	15.77%	16.86%	15.88%	16.96%	16.42%	17.54%	17.01%	18.20%

*, **, ***, significant at 10%, 5%, and 1%, respectively, with robust p-values in brackets; standard errors are clustered by firm IBTIC.

Note: Firm-average values are calculated for all independent variables except for NSEGMENTS, LAWSUIT, LITIGATE, and MAX_NSinceDISC, where the most recent value of each firm-quarter is used. See Tables 2 and 3 for variable definitions.

Table 4 – continued

Panel B. Firm-Level Regressions for Firms Splits by Positive and Negative Cost Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS Model Where the Dependent Variable is the Cost							
	500 Firms with Positive Costs				783 Firms with Negative Costs			
	COST (ĉ)	COST_mod (ĉ)	COST (ĉ)	COST_mod (ĉ)	COST (ĉ)	COST_mod (ĉ)	COST (ĉ)	COST_mod (ĉ)
Intercept	0.065**	0.032	0.095***	0.058*	-0.177***	-0.249***	-0.229***	-0.305***
	[0.028]	[0.257]	[0.001]	[0.051]	[0.001]	[0.002]	[0.000]	[0.000]
HHI	-0.083	-0.12	-0.075	-0.112	0.461***	0.496**	0.264**	0.243
	[0.435]	[0.307]	[0.458]	[0.311]	[0.001]	[0.018]	[0.042]	[0.228]
CAPEX	-0.228	-0.175	-0.336*	-0.270	0.993**	1.482***	0.936**	1.414**
	[0.214]	[0.330]	[0.059]	[0.118]	[0.013]	[0.009]	[0.017]	[0.012]
TURNOVER	0.015	0.017	0.000	0.003	0.107***	0.199***	0.100***	0.190***
	[0.608]	[0.581]	[0.992]	[0.928]	[0.006]	[0.000]	[0.010]	[0.001]
MTB	0.011***	0.009***	0.009**	0.007**	-0.037***	-0.051***	-0.034***	-0.047***
	[0.003]	[0.008]	[0.016]	[0.034]	[0.000]	[0.000]	[0.000]	[0.000]
LOG MKT CAP	-0.009**	-0.003	-0.007	0.000	0.048***	0.077***	0.033***	0.058***
(in billions)	[0.036]	[0.639]	[0.106]	[0.980]	[0.000]	[0.000]	[0.000]	[0.000]
NSEGMENTS	0.000	0.001	0.000	0.001	0.003	0.003	0.000	0.000
	[0.770]	[0.429]	[0.918]	[0.590]	[0.476]	[0.606]	[0.949]	[0.968]
LEVERAGE	-0.098***	-0.068**	-0.063***	-0.036	0.148**	0.154	0.111*	0.107
	[0.000]	[0.023]	[0.005]	[0.225]	[0.027]	[0.113]	[0.085]	[0.261]
INSTITUTIONAL%	-0.026	-0.013	-0.017	-0.004	0.054**	0.068*	0.045*	0.054
	[0.174]	[0.491]	[0.360]	[0.856]	[0.032]	[0.053]	[0.060]	[0.109]
AUDIT	0.012	0.008	0.015	0.010	0.023	0.051	0.019	0.046
	[0.401]	[0.682]	[0.285]	[0.574]	[0.383]	[0.180]	[0.449]	[0.224]
NUMEST	0.001	0.001	0.000	0.000	-0.004**	-0.006*	-0.001	-0.001
	[0.324]	[0.467]	[0.974]	[0.845]	[0.046]	[0.053]	[0.658]	[0.624]
EARNVOL	-0.021***	-0.013*	-0.008	-0.001	0.175**	0.181	0.142*	0.136
	[0.006]	[0.062]	[0.297]	[0.908]	[0.036]	[0.115]	[0.057]	[0.191]
BETA	0.035**	0.031**	0.032**	0.028**	-0.051*	-0.056	-0.040	-0.042
	[0.011]	[0.025]	[0.014]	[0.037]	[0.086]	[0.197]	[0.164]	[0.320]
LAWSUIT			-0.002	0.008			0.095**	0.124*
			[0.958]	[0.868]			[0.047]	[0.084]
LITIGATE			0.009	0.012			-0.035*	-0.054**
			[0.448]	[0.445]			[0.055]	[0.049]
MAX_NSinceDISC			-0.003***	-0.002***			0.005***	0.006***
			[0.000]	[0.000]			[0.000]	[0.000]
N	500	500	500	500	783	783	783	783
Adj. R-Square	10.55%	4.51%	19.32%	11.21%	21.19%	20.00%	26.58%	24.01%

*, **, ***, significant at 10%, 5%, and 1%, respectively, with robust p-values in brackets; standard errors are clustered by firm IBTIC. See Tables 2 and 3 for variable definitions.

Table 5 – Logit Regression of the Probability of Disclosure for 2015, Using Estimates of Cost from 2004-2014

Panel A. Firms with Positive and Negative Cost Estimates (for Estimation Period 2004-2014)

	N	Mean	Median	St.Dev.	Min	Max
Minimum Forecast Surprise	1,274	-9.169%	-2.806%	28.601%	-606.734%	38.617%
Modified Minimum Forecast Surprise	1,262	-14.995%	-4.630%	52.384%	-1207.059%	22.020%
Average Forecast Surprise	1,274	0.396%	0.918%	7.500%	-123.126%	79.174%
N Periods with Forecasts	1,274	12.92	8.00	11.79	0.00	44.00
N Periods for Each Firm	1,274	32.65	41.00	13.40	2.00	45.00
% of Periods with Forecasts	1,274	42.3%	34.8%	30.6%	0.0%	100.0%
COST $(\hat{c})_{2014}$	**1,268**	**-7.927%**	**-1.225%**	**23.454%**	**-110.6%**	**48.3%**
COST_mod $(\underline{\hat{c}})_{2014}$	**1,256**	**-13.173%**	**-3.016%**	**32.498%**	**-178.8%**	**42.9%**

Table 5 – continued

Panel B. Regressions with Positive and Negative Cost Estimates from 2004-2014

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Logit Model Where the Dependent Variable is the Probability of Making Management Forecast in 2015							
Intercept	-0.567***	-0.588***	-2.671***	-2.803***	-2.671***	-2.804***	-2.664***	-2.791***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
COST ($\hat{c}$)$_{2014}$	-1.346***		-1.437***		-1.479**		-1.448**	
	[0.003]		[0.003]		[0.012]		[0.018]	
COST_mod ($\underline{\hat{c}}$)$_{2014}$		-0.991***		-1.105***		-1.215**		-1.190**
		[0.003]		[0.002]		[0.011]		[0.014]
FREQ$_{2014}$			4.809***	4.966***	4.812***	4.968***	4.816***	4.970***
			[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
MIN_Surprise$_{2014}$					0.039		0.111	
					[0.882]		[0.793]	
MIN_mod_Surprise$_{2014}$						0.066		0.127
						[0.657]		[0.595]
MEAN_Surprise$_{2014}$							-0.49	-0.786
							[0.809]	[0.692]
N	2,998	2,950	2,998	2,950	2,998	2,950	2,998	2,950
Pseudo R-Square	1.11%	1.09%	30.23%	31.54%	30.23%	31.55%	30.24%	31.56%

*, **, ***, significant at 10%, 5%, and 1%, respectively, with robust p-values in brackets; standard errors are clustered by firm IBTIC.

Note: The dependent variable is DISCLOSE$_{2015}$, which is an indicator variable equal to 1 for management forecasts for each quarter of 2015, and 0 otherwise. The variables for disclosure COST ($\hat{c}$), frequency of management forecasts (FREQ), minimum forecast surprise (MIN_Surprise), and average forecast surprise (MEAN_Surprise) are estimated for quarterly forecasts from the estimation period 2004-2014; the holdout period is 2015. See Tables 2 and 3 for variable definitions.

Table 5 – continued

Panel C. Regressions with Positive and Negative Cost Estimates from 2004-2014

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Logit Model Where the Dependent Variable is the Probability of Making Management Forecast in 2015							
Intercept	-0.365*	-0.355*	-0.964***	-0.996***	1.391***	1.465***	1.136***	1.205***
	[0.083]	[0.100]	[0.000]	[0.000]	[0.000]	[0.000]	[0.002]	[0.002]
COST (ĉ)$_{2014}$	-0.821*		-0.944**		-0.829**		-0.344	
	[0.070]		[0.034]		[0.024]		[0.379]	
COST_mod (ĉ)$_{2014}$		-0.620*		-0.693**		-0.755**		-0.367
		[0.064]		[0.034]		[0.016]		[0.253]
HHI$_{2014}$	-0.562	-0.764						
	[0.666]	[0.562]						
CAPEX$_{2014}$	-1.938	-1.965						
	[0.548]	[0.546]						
TURNOVER$_{2014}$	-0.102	-0.124						
	[0.806]	[0.769]						
MTB$_{2014}$	0.092***	0.091***						
	[0.006]	[0.007]						
LOG MKT CAP$_{2014}$	-0.003	0.005						
(in billions)	[0.945]	[0.923]						
NSEGMENTS	-0.101***	-0.102***						
	[0.000]	[0.001]						
LAWSUIT			1.005***	1.010***			0.436	0.400
			[0.008]	[0.008]			[0.501]	[0.547]
LITIGATE			0.688***	0.680***			0.335**	0.305**
			[0.000]	[0.000]			[0.027]	[0.050]
NUMEST$_{2014}$			0.013	0.014			0.001	0.001
			[0.174]	[0.137]			[0.900]	[0.933]
NSinceDISC					-0.404***	-0.442***	-0.392***	-0.431***
					[0.000]	[0.000]	[0.000]	[0.000]
EARNVOL$_{2014}$							-1.298***	-1.236***
							[0.001]	[0.001]
BETA$_{2014}$							0.309	0.322
							[0.176]	[0.168]
N	2,998	2,950	2,998	2,950	2,998	2,950	2,998	2,950
Pseudo R-Square	3.14%	3.13%	3.36%	3.35%	46.29%	48.23%	47.26%	49.06%

*, **, ***, significant at 10%, 5%, and 1%, respectively, with robust p-values in brackets; standard errors are clustered by firm IBTIC.

Note: The dependent variable is DISCLOSE$_{2015}$, which is an indicator variable equal to 1 for management forecasts for each quarter of 2015, and 0 otherwise. The independent variables HHI, CAPEX, TURNOVER, MTB, LOG MKT CAP, NUMEST, EARNVOL, and BETA are estimated for quarterly forecasts from the estimation period 2004-2014; the holdout period is 2015. The independent variables NSEGMENTS, LAWSUIT, LITIGATE, and NSinceDISC are calculated for each quarter of 2015. See Tables 2 and 3 for variable definitions.

Table 6 – Estimates of Costs Within Stickiness Categories of Most Recent Forecast

Panel A. Cost Estimated Within Categories of Most Recent Forecast

Cost_mm ($\hat{c}_{mm}$)	N	Cost ($\hat{c}$)	Med-Minimum Forecast Surprise	Mean Forecast Surprise	Frequency of Forecasts	St.Dev.$\hat{c}$
1 Quarter Ago	17,288	-1.176%	-2.352%	0.267%	81.490%	0.440%
2-3 Quarters Ago	5,018	0.993%	0.391%	1.416%	29.813%	0.141%
4-8 Quarters Ago	5,347	1.298%	1.166%	1.002%	11.595%	0.115%
9-46 Quarters Ago	10,492	0.951%	0.927%	0.927%	2.526%	0.015%
Total	38,145					
Cost_np ($\hat{c}_{np}$)						
1 Quarter Ago	17,288	-3.583%	-4.759%	0.267%	81.490%	0.440%
2-3 Quarters Ago	5,018	1.638%	1.036%	1.416%	29.813%	0.141%
4-8 Quarters Ago	5,347	4.390%	4.258%	1.002%	11.595%	0.115%
9-46 Quarters Ago	10,492	1.676%	1.652%	0.927%	2.526%	0.015%
Total	38,145					

Panel B. Information Endowment for Cost Estimated Within Categories of Most Recent Forecast

				St.Dev		Confidence Interval for q				Model Not
Cost_mm ($\hat{c}_{mm}$)	N	q_low	q_high	q_low	q_high	Min	Max	Dye Test	Verr Test	Consistent?
1 Quarter Ago	17,288	9.25%	18.51%	3.47%	0.30%	3.54%	19.00%	yes	no	consistent
2-3 Quarters Ago	5,018	0.00%	70.19%	0.00%	0.65%	0.00%	71.25%	no	yes	consistent
4-8 Quarters Ago	5,347	0.00%	88.40%	0.00%	0.44%	0.00%	89.12%	no	yes	consistent
9-46 Quarters Ago	10,492	0.00%	97.47%	0.00%	0.15%	0.00%	97.73%	no	yes	consistent
Total	38,145									
Cost_np ($\hat{c}_{np}$)										
1 Quarter Ago	17,288	13.94%	18.51%	1.74%	0.30%	11.08%	19.00%	yes	no	consistent
2-3 Quarters Ago	5,018	0.00%	70.19%	0.00%	0.65%	0.00%	71.25%	no	yes	consistent
4-8 Quarters Ago	5,347	0.00%	88.40%	0.00%	0.44%	0.00%	89.12%	no	yes	consistent
9-46 Quarters Ago	10,492	0.00%	97.47%	0.00%	0.15%	0.00%	97.73%	no	yes	consistent
Total	38,145									

Note: Most recent forecast information is obtained from the variable Disclosure based on management forecasts. Of our sample of 44,946 observations, there were 6,801 observations that did not have a history of Disclosure and were excluded from this analysis. Cost_mm ($\hat{c}_{mm}$) is estimated within each group and calculated as the median of (firm minimum forecast surprise) + [frequency of management forecasts /(1-frequency of management forecasts)]*average forecast surprise. $Cost_{np}$ ($\hat{c}_{np}$) is calculated as the np threshold estimator + [frequency of management forecasts /(1-frequency of management forecasts)]*average forecast surprise. See Tables 2 and 3 for variable definitions.

Table 6 – continued

Panel C. Dynamic Cost Estimated Within Categories of Most Recent Forecast (Until 12 Quarters or 3 Years Ago)

			Confidence Interval for $\hat{c}_{dyn}$				Confidence Interval for $\hat{c}_{dyn}$	
Most Recent Forecast:	Cost ($\hat{c}_{dyn}$)	St.Dev.$\hat{c}_{dyn}$	Min	Max	Cost ($\hat{c}_{np\_dyn}$)	St.Dev.$\hat{c}_{np\_dyn}$	Min	Max
1 Quarter Ago	-0.006%	0.422%	-0.96%	0.74%	-5.135%	2.196%	-7.71%	-0.82%
2 Quarters Ago	2.431%	0.425%	1.56%	3.20%	3.090%	3.148%	-3.88%	7.19%
3 Quarters Ago	2.442%	0.511%	1.33%	3.30%	2.604%	4.538%	-6.86%	9.32%
4 Quarters Ago	2.464%	0.708%	0.84%	3.58%	0.220%	6.433%	-11.40%	10.11%
5 Quarters Ago	2.161%	0.794%	0.32%	3.33%	5.047%	8.356%	-13.46%	15.80%
6 Quarters Ago	2.136%	0.919%	-0.01%	3.47%	4.737%	10.296%	-17.09%	20.43%
7 Quarters Ago	1.501%	1.125%	-1.03%	3.23%	4.040%	12.254%	-18.19%	24.71%
8 Quarters Ago	1.881%	1.269%	-1.24%	3.36%	4.871%	14.811%	-21.22%	29.65%
9 Quarters Ago	1.161%	1.476%	-2.27%	3.11%	3.789%	17.315%	-24.01%	34.10%
10 Quarters Ago	1.865%	1.678%	-2.30%	3.88%	4.194%	20.886%	-27.22%	39.91%
11 Quarters Ago	2.059%	1.945%	-2.99%	4.11%	2.516%	44.054%	-32.78%	105.95%
12-46 Quarters Ago	1.648%	2.442%	-4.70%	3.86%	3.978%	19.539%	-44.25%	15.06%

Note: Most recent forecast information is obtained from the variable Disclosure based on management forecasts. Dynamic Cost ($\hat{c}_{dyn}$) is estimated within each group. See Tables 2 and 3 for variable definitions.

Table 7 – Estimates of Costs Within Number of Business Segments

Panel A. Cost Estimated Within Number of Business Segments

Cost_mm ($\hat{c}_{mm}$)	N	Cost ($\hat{c}$)	Med-Minimum Forecast Surprise	Mean Forecast Surprise	Frequency of Forecasts	St.Dev.$\hat{c}$
1 Segment	17,457	-5.626%	-5.482%	-0.172%	45.540%	0.134%
2 to 3 Segments	8,518	-1.552%	-2.090%	0.889%	37.685%	0.128%
4 to 6 Segments	9,804	-0.586%	-0.973%	0.727%	34.710%	0.044%
7 or More Segments	9,167	0.026%	-0.442%	0.877%	34.777%	0.126%
Total	44,946					
Cost_np ($\hat{c}_{np}$)						
1 Segment	17,457	-3.354%	-3.210%	-0.172%	45.540%	0.134%
2 to 3 Segments	8,518	-0.799%	-1.336%	0.889%	37.685%	0.128%
4 to 6 Segments	9,804	1.753%	1.366%	0.727%	34.710%	0.044%
7 or More Segments	9,167	-5.813%	-6.280%	0.877%	34.777%	0.126%
Total	44,946					

Panel B. Information Endowment for Cost Estimated Within Number of Business Segments

				St.Dev		Confidence Interval for q				
Cost_mm ($\hat{c}_{mm}$)	N	q_low	q_high	q_low	q_high	Min	Max	Dye Test	Verr Test	Model Not Consistent?
1 Segment	17,457	55.89%	54.46%	1.38%	0.38%	53.62%	55.08%	yes	no	consistent
2 to 3 Segments	8,518	46.29%	62.32%	3.87%	0.53%	39.92%	63.18%	yes	no	consistent
4 to 6 Segments	9,804	39.34%	65.29%	3.01%	0.48%	34.38%	66.08%	yes	no	consistent
7 or More Segments	9,167	0.00%	65.22%	0.00%	0.50%	0.00%	66.04%	no	no	consistent
Total	44,946									
Cost_np ($\hat{c}_{np}$)										
1 Segment	17,457	56.91%	54.46%	2.30%	0.38%	53.13%	55.08%	yes	no	consistent
2 to 3 Segments	8,518	37.26%	62.32%	6.00%	0.53%	27.38%	63.18%	yes	no	consistent
4 to 6 Segments	9,804	0.00%	65.29%	0.00%	0.48%	0.00%	66.08%	no	yes	consistent
7 or More Segments	9,167	60.37%	65.22%	1.43%	0.50%	58.02%	66.04%	yes	no	consistent
Total	44,946									

Note: Number of business segments is obtained from the Compustat variable busseg. Cost_mm ($\hat{c}_{mm}$) is estimated within each group and calculated as the median of (firm minimum forecast surprise) + [frequency of management forecasts /(1-frequency of management forecasts)]*average forecast surprise. Cost_np ($\hat{c}_{np}$) is calculated as the np threshold estimator + [frequency of management forecasts /(1-frequency of management forecasts)]*average forecast surprise. q_low = max(0,1-frequency + frequency/ minimum*average). q_high = 1 - frequency. Min Confidence Interval for q = max(0 , q_low - 1.64485* [sqrt((1/minimum^2)*frequency*(variance of forecast surprise + (1-frequency)*(average-minimum)^2))*(1/sqrt(number of observations))]). Max Confidence Interval for q = min(1, (1-frequency) +1.64485* [sqrt(frequency*(1-frequency))/sqrt(number of observations)]). See Tables 2 and 3 for variable definitions.

Table 8 – Estimates of Costs Within Number of Lawsuits per Firm and High and Low Litigation Industries

Panel A. Cost Estimated Within Number of Lawsuits per Firm

Cost_mm ($\hat{c}_{mm}$)	N	Cost ($\hat{c}$)	Med-Minimum Forecast Surprise	Mean Forecast Surprise	Frequency of Forecasts	St.Dev.$\hat{c}$
0 Lawsuits	32,776	-2.791%	-3.008%	0.340%	38.873%	0.062%
1 Lawsuit	9,994	-2.336%	-2.824%	0.653%	42.776%	0.180%
2+ Lawsuits	2,176	-2.754%	-2.494%	-0.510%	33.778%	0.238%
Total	44,946					
Cost_np ($\hat{c}_{np}$)						
0 Lawsuits	32,776	2.457%	2.240%	0.340%	38.873%	0.062%
1 Lawsuit	9,994	-5.729%	-6.217%	0.653%	42.776%	0.180%
2+ Lawsuits	2,176	-1.541%	-1.280%	-0.510%	33.778%	0.238%
Total	44,946					

Panel B. Information Endowment for Cost Estimated Within Number of Lawsuits per Firm

				St.Dev		Confidence Interval for q				
Cost_mm ($\hat{c}_{mm}$)	N	q_low	q_high	q_low	q_high	Min	Max	Dye Test	Verr Test	Model Not Consistent?
0 Lawsuits	32,776	56.73%	61.13%	1.30%	0.27%	54.59%	61.57%	yes	no	consistent
1 Lawsuit	9,994	47.33%	57.22%	3.69%	0.49%	41.26%	58.04%	yes	no	consistent
2+ Lawsuits	2,176	73.13%	66.22%	6.37%	1.01%	62.65%	67.89%	yes	no	consistent
Total	44,946									
Cost_np ($\hat{c}_{np}$)										
0 Lawsuits	32,776	0.00%	61.13%	0.00%	0.27%	0.00%	61.57%	no	yes	consistent
1 Lawsuit	9,994	52.73%	57.22%	1.74%	0.49%	49.87%	58.04%	yes	no	consistent
2+ Lawsuits	2,176	79.67%	66.22%	12.32%	1.01%	59.40%	67.89%	yes	no	consistent
Total	44,946									

Note: Lawsuit dates are obtained from the Stanford Securities Litigation Database. Cost_mm ($\hat{c}_{mm}$) is estimated within each group and calculated as the median of (firm minimum forecast surprise) + [frequency of management forecasts /(1-frequency of management forecasts)]*average forecast surprise. Cost_np ($\hat{c}_{np}$) is calculated as the np threshold estimator + [frequency of management forecasts /(1-frequency of management forecasts)]*average forecast surprise. q_low = max(0,1-frequency + frequency/ minimum*average). q_high = 1 - frequency. Min Confidence Interval for q = max(0 , q_low - 1.64485* [sqrt((1/minimum^2)*frequency*(variance of forecast surprise + (1-frequency)*(average-minimum)^2))*(1/sqrt(number of observations))]). Max Confidence Interval for q = min(1, (1-frequency) +1.64485* [sqrt(frequency*(1-frequency))/sqrt(number of observations)]). See Tables 2 and 3 for variable definitions.

Table 8 – continued

Panel C. Cost Estimated Within High and Low Litigation Industries

Cost_mm ($\hat{c}_{mm}$)	N	Cost ($\hat{c}$)	Med-Minimum Forecast Surprise	Mean Forecast Surprise	Frequency of Forecasts	St.Dev.$\hat{c}$
Low Litigation Industry	28,138	-0.909%	-1.212%	0.569%	34.754%	0.047%
High Litigation Industry	16,808	-6.410%	-6.544%	0.149%	47.430%	0.159%
Total	44,946					
Cost_np ($\hat{c}_{np}$)						
Low Litigation Industry	28,138	0.357%	0.054%	0.569%	34.754%	0.047%
High Litigation Industry	16,808	-4.560%	-4.694%	0.149%	47.430%	0.159%
Total	44,946					

Panel D. Information Endowment for Cost Estimated Within High and Low Litigation Industries

				St.Dev		Confidence Interval for q		Dye	Verr	Model Not
Cost_mm ($\hat{c}_{mm}$)	N	q_low	q_high	q_low	q_high	Min	Max	Test	Test	Consistent?
Low Litigation Industry	28,138	48.93%	65.25%	2.56%	0.28%	44.71%	65.71%	yes	no	consistent
High Litigation Industry	16,808	51.49%	52.57%	1.34%	0.39%	49.29%	53.20%	yes	no	consistent
Total	44,946									
Cost_np ($\hat{c}_{np}$)										
Low Litigation Industry	28,138	0.00%	65.25%	0.00%	0.28%	0.00%	65.71%	no	yes	consistent
High Litigation Industry	16,808	51.07%	52.57%	1.83%	0.39%	48.06%	53.20%	yes	no	consistent
Total	44,946									

Note: We classify high litigation industries as those with the following SIC codes: biotechnology (2833-2836, 8731-8734), computers (3570-3577, 7370-7374), electronics (3600-3674), and retail (5200-5961), based on Ajinkya et al (2005). Cost_mm ($\hat{c}_{mm}$) is estimated within each group and calculated as the median of (firm minimum forecast surprise) + [frequency of management forecasts /(1-frequency of management forecasts)]*average forecast surprise. Cost_np ($\hat{c}_{np}$) is calculated as the np threshold estimator + [frequency of management forecasts /(1-frequency of management forecasts)]*average forecast surprise. q_low = max(0,1-frequency + frequency/ minimum*average). q_high = 1 - frequency. Min Confidence Interval for q = max(0 , q_low - 1.64485* [sqrt((1/minimum^2)*frequency*(variance of forecast surprise + (1-frequency)*(average-minimum)^2))*(1/sqrt(number of observations))]). Max Confidence Interval for q = min(1, (1-frequency) +1.64485* [sqrt(frequency*(1-frequency))/sqrt(number of observations)]). See Tables 2 and 3 for variable definitions.