How often do managers withhold information?*

Jeremy Bertomeu, Paul Ma, and Iván Marinovic

August 24, 2018

Abstract

We structurally estimate a model of voluntary disclosure with uncertain information endowment (Dye 1985; Jung and Kwon 1988; Einhorn and Ziv 2008). Each period, the manager may have an incentive to strategically withhold information to increase the current market price. We develop a tractable empirical framework that explains serial correlation in disclosure and controls for other observed and unobserved time-varying determinants of disclosure. The model identifies the contribution of short-term price motives to the probability of disclosure. Estimating the model using management forecasts of annual earnings suggest that information is strategically withheld about 10% of the time due to price motives. We show that the model captures a number of empirical facts and quantify the consequences of changes in the reporting environment in various counter-factual analyses.

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

---

*Jeremy Bertomeu is an associate professor at the Rady School of Management, University of California San Diego, 9500 Gilman Dr, La Jolla, CA 92093. Email: mailto:jeremy.bertomeu@gmail.comjbertomeu@ucsd.edu. Paul Ma is an assistant professor at the Carlson School of Management, University of Minnesota, 321 19th Avenue South, Minneapolis, MN 55455. Email: mailto:paulma@umn.edupaulma@umn.edu. Iván Marinovic is an associate professor at Stanford Graduate School of Business, Stanford University, 655 Knight Way, Stanford, CA 94305. Email: mailto:imvial@stanford.eduimvial@stanford.edu. We are grateful for comments on prior drafts from Santiago Bazdresch, Alan Benson, Carlos Corona (Stanford Summer Camp discussant), Joe Gerakos (CFEA discussant), William Greene, Colleen Manchester, Alejandro Molnar, Raj Singh, Aaron Sojourner, John Shoven, Stephen Terry, as well as workshop participants at University of Alberta, University of Chicago, Baruch College, Duke, Penn State, Purdue, University of Minnesota and Stanford.

1 Introduction

Voluntary disclosure theory postulates that managers will reveal some of their private information if they (i) may truthfully disclose information, (ii) choose to do so if it increases the market price, but (iii) are subject to a disclosure friction that may prevent costless communication. To this date, however, the extent to which the theory explains disclosure behavior remains controversial. In this paper, we develop a simple structural framework to test and estimate a family of voluntary disclosure models featuring uncertainty about information endowment (Dye 1985; Jung and Kwon 1988), or DJK, as extended by Einhorn and Ziv (2008) to explain stickiness in disclosure behavior.

In these models, a manager may be subject to a friction which causes her to be unable (or unwilling) to disclose for exogenous reasons. Dye (1985) and Jung and Kwon (1988), hereafter DJK, interpret the friction as uncertainty about information endowment, as an uninformed manager may not credibly convey an absence of information. More generally, in this model, the friction may have alternative interpretations that cause the manager not to disclose for reasons separate from the value of the firm: in certain periods, the manager may have insufficient career or financing motive to care about short-term price, may have received unverifiable information or may not have the time to prepare a credible disclosure. For expositional purposes, we describe the friction in terms of uncertain information endowment but our assumptions apply to other interpretations as well.

In the absence of a disclosure, investors weight the possibility that the manager was uninformed, which leads informed managers to strategically withhold sufficiently unfavorable information. Einhorn and Ziv (2008) extend this model to an information endowment process that is correlated over time with investors dynamically updating their beliefs from observed disclosures and other public information. Specifically, a disclosing firm reveals its information endowment in the current period and, therefore, is perceived as likely to be informed again in the following period, leading investors to increase their assessment of a strategic disclosure and, thus, increasing the price penalty conditional on a non-disclosure. This model provides a tractable foundation for stickiness implied by a rational price response of investors to past disclosures.

We make the theory more amenable to empirical analysis by nesting it within a discrete choice framework (McFadden 1973, 1980). DJK implies a deterministic decision rule whereby the manager discloses if and only if the information is above the theoretical threshold. Unfortunately, this feature is unlikely to fit to actual data that may include confounding reasons for non-disclosure that are not observable to the researcher. To address this, we assume that the manager's preference is a function of both the price reaction to disclosures and preference shocks unobservable to investors. Because managers no longer adopt strict threshold strategies, investors form the non-disclosure price as a function of the distribution of preference shocks. We derive a generalized pricing equation and show that the econometric model has a simple form that can be estimated by maximum likelihood.

Our approach provides several novel benefits. We can evaluate whether managers appear to weight the price consequence of withholding, as quantitatively implied by the equations of the theory. Moreover, the model allows us to recover the probability of information endowment at any period, and estimate the probability that information is concealed for strategic reasons. Therefore, the estimated structural parameters provide detailed knowledge over the assumed economic mechanism. We conduct many additional analyses, showing that the model exhibits properties in line with the sample and demonstrate in counter-factuals how changes to the manager's information endowment process or preference would affect the information conveyed by disclosures.

We evaluate the theory in the context of annual management forecasts and realized earnings from January 1^{st} 2004 to December 31^{st} 2015. In practice, these forecasts are highly publicized, released as part of conference calls recirculated in the press, and command larger price reactions than most other sources of financial information (Beyer, Cohen, Lys, and Walther 2010). Management forecasts are not required by law and, by and large, present the most widespread setting where the theory is tested. For example, Dye (1985) explicitly refers to forecasts as a primary application, noting that "it is commonly believed that managers possess information about the firms they run, such as annual earnings, but not the distribution of their firms' future ' forecasts. The reluctance of managers to disclose such nonproprietary information is the subject of this paper" (p.123 - 124).[1]

Interestingly, the application of the theory to forecasting behavior remains heavily debated, with a large number of studies finding evidence of strategic behaviors consistent with the theory, while other studies documenting other cross-sectional determinants of disclosure unexplained by the theory.[2] A structural model provides one piece of evidence to this debate: using a formal derivation of the empirical model from the theory, we can derive quantitative implications incremental to traditional reduced-form estimation.

For example, in our baseline estimates, we find that managers strategically withhold forecasts about 20% of the time. We quantify economic consequences of this behavior and find that it increases the market's uncertainty about the upcoming earnings by 6% relative to a counterfactual world without strategic concealment. This relatively modest effect is explained by the fact that the act of withholding is per se informative and the manager's private information contains significant noise (forecasts are made six month in advance, on average). We estimate that manager's private information, if any, can explain roughly 40% of the variation in earnings. Finally, by measuring the price given non-disclosure, we find that a naive investor would over-estimate the upcoming earnings by 25% relative to a rational investor.

But, as noted in the existing literature, price motives do not explain all forecasting behavior. Indeed, we estimate that there is a large preference intercept for disclosure that

[1] Many theoretical studies also explicitly use management forecasts as applications. As a representative example, Dye (1985) notes "it is commonly believed that managers possess information about the firms they run, such as annual earnings, but not the distribution of their firms' future ' forecasts. The reluctance of managers to disclose such nonproprietary information is the subject of this paper" (p.123-124).

[2] In particular, King and Wallin (1991) find experimental evidence of the disclosure friction. In the context of management forecasts, Penman (1980) and Lev and Penman (1990) find a positive association between the decision to issue forecast and good news in the cross section. However, perhaps to deter litigation, various studies find evidence that managers pre-empt earning announcement with negative earning warnings (Kasznik and Lev 1995; Skinner 1997; Baginski, Hassell, and Kimbrough 2002).

is not explained by price motives only. In other words, the model suggests that managers often disclose news that would decrease the market price relative to non-disclosure and implies the presence of a latent non-price benefit from disclosure.

We conduct several robustness checks controlling for size, analyst coverage and time trends and in subsamples ranked by size and analyst coverage. The results are consistent with our baseline model, but also reveal additional patterns in subsamples. We also run analyses in a sample of quarterly management forecasts. Contrary to the annual forecast sample, we find price motives to have a muted role in this alternate sample. Presumably, managers care less about accelerating news into a price by a quarter by making a quarter-over-quarter forecast. Indeed, the coefficient on price motives is about times smaller in this sample, and the effect of price motives on the probability of disclosure is about 15% smaller.

Related literature. Several recent studies examine properties of voluntary disclosure within a structural model. Cheynel and Liu-Watts (2015) and Bertomeu, Beyer, and Taylor (2015) estimate the implied disclosure cost in the classic Verrecchia (1983). Their primary focus, however, is not on the DJK model although that, as part of future efforts, more general models may involve estimating and separating both types of friction. Also in the context of costly disclosure theory, the recent study by Zhou (2016) shows that investor learning about fundamentals may create in stickiness in forecasts and develops tools to estimate the model with firm-level heterogeneity using Bayesian statistics. Lastly, Bertomeu, Marinovic, Terry, and Varas (2017) estimate using simulated method of moments a model of reputation-building by a forward-looking manager.

Our analysis is part of an ongoing effort in accounting research to quantitatively measure the communication by various parties to the financial market, using guidance from existing theoretical models. This effort is not only based on structural estimation but also on learning from a considerable body of existing reduced-form evidence (Beyer, Cohen, Lys, and Walther 2010). Furthermore, other approaches different from structural estimation can bridge the gap between theory and evidence. As examples, the studies by

Chen and Jiang (2006), Gerakos and Kovrijnykh (2013) or Fang, Huang, and Wang (2017) but this is a subset of a broader literature that we cannot review here in its entirety.

Speaking to the growing use of structural estimation as a tool for empirical analysis, there is an emerging literature that uses structural models to understand the release and use of information in accounting numbers, as evidenced by examples of recent working papers spanning a large variety of accounting topics and of which we give some representative examples below. Gayle, Li, and Miller (2015a), Caskey, Gayle, and Li (2017) and Li (2018) estimate models of optimal compensation. Beyer, Guttman, and Marinovic (2014) estimate a dynamic misreporting model within the noisy signaling framework of Fischer and Verrecchia (2000) and Dye and Sridhar (2004). Hemmer and Labro (2015), Liang, Sun, and Tam (2016) and Breuer and Windisch (2018) develop investment models that match empirical properties of earnings.

2 Sample and Motivating Facts

Voluntary disclosure theory is a framework to understand management forecasts. Indeed, some of the core predictions of disclosure models, such as Verrecchia (1983) and Dye (1985), are consistent with several notable empirical facts. Since documenting these facts has been the object of an extensive prior empirical literature (Beyer, Cohen, Lys, and Walther 2010), we describe our sample and then document three motivating facts for our structural model.

We apply the model to annual management forecasts about earnings per share (EPS). As shown in Table 1, the starting sample consists of all US firms present in the IBES annual earnings announcement (EA) database whose fiscal periods end between January 1$^{\text{st}}$ 2004 and December 31$^{\text{st}}$ 2015, with earnings only in dollars.[3] We require each observation to have (pro-forma) EPS and a stock-split adjustment factor.[4] We also require each firm

[3]We set the starting date after Regulation Fair Disclosure, because this regulatory event increased the observed frequency of management forecasts for reasons outside of the model.

[4]We use raw EPS over other measures because, as shown by Cheong and Thomas (2011), raw EPS

to be traded on one of the three major exchanges and remove all fiscal periods that lasted less than 2/3 or more than 4/3 of a year, which shrinks the sample to 49,552 earnings announcements and 7,599 unique firms. Management forecasts (MF) are obtained from the IBES company-issued guidance database. We only retain quantitative guidance, that is, guidance that contains at least one quantitative estimate of EPS and forecasts that are at least six months ahead of the fiscal period end (to select only long-term forecasts) and made earlier than the past earnings announcement. Each guidance comes with a qualitative description (variable *range_desc*) which we classify into three groups: (1) interval ($range\_desc = 1, 6, 8$), (2) point estimate ($range\_desc = 2, 3, 11, 12, 13, 14, 17$), (3) open-ended "greater than" ($range\_desc = 4, 16$). We drop 51 open-ended forecasts "less than" ($range\_desc = 10, 15$) because these (rare) forecasts are not within the scope of the theory. Each category comes with a single quantitative guidance, which we denote MF, except category (1) which includes the upper bound of the interval, which we denote MF_upper.

tend to be relatively invariant with scale and time because, presumably, firms manage the number of shares to maintain comparable changes EPS. By contrast, (split) adjusted EPS tend to become less volatile over past periods for firms with stock splits. Other alternative measures such as earnings scaled by assets or market value are possible, but they are not the variable being forecasted and, relative to the EPS forecast, can feature additional noise due to scaling.

Table 1. Sample Selection

The starting sample of earnings announcements (EA) and management forecasts (MF) is US firms with earnings per share (EPS) solely in USD covered I/B/E/S during fiscal period ends 2004-2015 with non-missing announcement date and EPS. Matching to CRSP and Compustat is sequential, using (i) ibes-permno and gvkey-permno matching tables from WRDS and CRSP, (ii) cusip and (iii) stock ticker. Standard errors of EPS surprises were computed based on a quantile regression of EPS on lag EPS and consensus. sdeps refers to the standard error of EPS and is coded as missing for firms with a single EA.

	Earnings announcements (EA)		Management forecasts (MF)	
	Nb. of EA	Unique firms	Nb. of MF	Unique firms
1) I/B/E/S EA sample 2004-2015	60,710	10,265		
2) Traded on AMEX, NASDAQ or NYSE	52,160	7,888		
3) Fiscal period must be between 2/3 and 4/3 of a year	49,552	7,599		
4) I/B/E/S guidance sample 2004-2015			60,721	3,165
5) Matched to EA			54,412	2,639
6) MF date no more than 30 days before EA and before period end date			48,931	2,540
7) MF at least 6 months ahead			24,660	2,331
8) Drop “lower than” forecasts			24,612	2,313
9 Retain earliest MF only			11,006	2,313
10) Merged I/B/E/S EA and guidance			10,956	2,300
11) Non-missing permno, gvkey, return at EA, lag EPS and total assets	42,471	6,444	10,251	2,155
12) Must have IBES consensus	37,749	5,921	10,129	2,105
13) Firms with 10 or more EA	20,059	1,822	6,815	994
14) Firms with least one period with MF and one period without MF	8,013	725	3,803	725
15) Trim firms 3 highest/lowest MF	8,007	725	3,797	725
16) Remove EA before first MF	6,788	718	3,780	718
Full sample	**6,788**	**718**	**3,780**	**718**

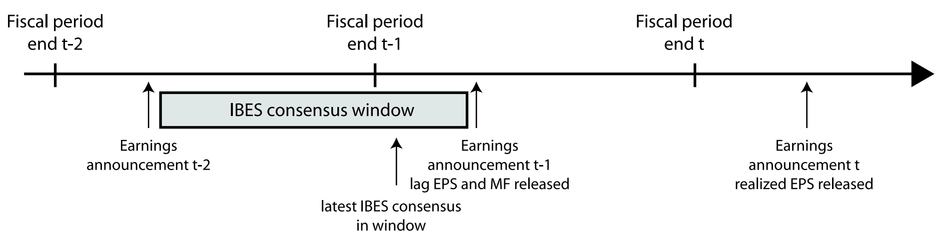
If there is more than one forecast during a period, we only retain the earliest forecast - this is a limitation of our empirical model as it does not explain disclosure dynamics and timing within a forecasting period (Acharya, Demarzo, and Kremer 2011; Marinovic 2013; Guttman, Kremer, and Skrzypacz 2014). We merge the EA sample with the forecast sample, CRSP and Compustat, requiring non-missing EA returns from CRSP, total assets from Compustat and lag EPS, for a total of 10,251 earnings announcements and 2,155 unique firms.

Our model requires the distribution of EPS and forecasts surprises to be stable across time and firms. However, market expectations vary across firms and periods, only because earnings are correlated over time and may respond to aggregate shocks. We rely on prior-year EPS and the IBES analyst consensus to capture market expectations about EPS where, as indicated in Figure 2, the consensus is calculated over the prior fiscal period before a forecast can be made. IBES provides a paired pre-forecast consensus for observations with management forecasts and, otherwise, we use the latest IBES consensus or, if it is missing, we compute an IBES consensus as the median of the latest forecast by analyst. We drop observations for which no consensus could be obtained. In addition, we adjust for differences in EPS levels across firms by scaling management forecasts, EPS, lag EPS and consensus by the standard error of each firm's EPS. Given that our model requires time-series of earnings announcements, we require ten or more earnings announcements. Further, we remove firms that have no time-series variation in forecasting, either because they forecasted in all years or because they never forecasted, and trim the three lowest and highest forecasts in the sample.

In the multi-period model, the probability that the manager is informed will be a time-varying latent variable. Therefore, with a finite time-series of earnings announcements, we do not observe the probability that the manager is informed for the first observations in the sample and, unfortunately, estimating it is computationally difficult because it would involve simulating many disclosure paths to estimate the likelihood of each p_t before the first disclosure (hence, the likelihood would no longer be in closed-form). Therefore, we drop observations before the first forecast, starting the sample at a point where the researcher knows the market's starting belief about the manager's information endowment. For observations after the first forecast, the model implies that $p_1 = k_1$ since, in this case, the manager is known to be informed in the prior period.[5]

The final sample has a total of 6,788 earnings announcements, of which 3,780 were followed by a forecast, among 718 unique firms. It is broadly descriptive of firms traded on major exchanges after they made their first forecast and provided they did not follow a perfectly predictable forecasting policy.

Figure 1. Measurement window



In order to measure the forecast surprise we need to estimate market expectations. Ideally, we would observe the consensus before and after a forecast, holding other sources of information fixed. Unfortunately, this is infeasible because four out of five forecasts are bundled with or within five days of the earnings announcement. Hence, we impute market expectations by regressing realized earnings the analyst consensus with and without the management forecast. The details of this procedure are described in Appendix A-3.

As shown in Table 2, firms forecast on average once every two years. The median firm in our sample has 10 years of data. The average number of consecutive disclosures is 2.05. On average, firms switch from disclosure to non-disclosure (or the other way round) 27% of the time. Though disclosure policies are clearly persistent, firms vary their disclosure choices significantly over time.

[5]In robustness analyses, we re-estimate the model without dropping these observations, setting the prior for the first observation at the steady-state expected probability of information endowment (this is only an approximation as, to be exact, the likelihood should be calculated by integrating over the entire stready-state distribution of priors).

EPS surprises are, on average, zero as expected from the regressions in Table 2. Average EPS revisions following a forecast are positive, equal to 21% of a standard-deviation of the EPS surprise and are consistent with DJK. Firms in the sample tend to be large, with more than 95% with assets above 100 million dollars and median assets of 2 billion dollars. The distribution of market capitalization is similar, with a median market capitalization of 1.7 billion dollars. The median book-to-market and debt-to-asset ratios are about .5, and the median firm spent 2% in $R\&D$ and 3% in capital expenditures.

Table 2. Descriptive Statistics

The sample includes all US firms covered by IBES for fiscal periods 2004-2015, with non-missing IBES or Compustat EPS and for fiscal periods that last 365 days, plus or minus one-third, and which can be merged to Compustat and CRSP using the WRDS link tables, cusip or ticker and satisfying the criteria in table 1. Management forecast (MF) frequency is the number of periods with at least one MF to the total number of periods by firm. EPS and MF are the raw USD earnings per share and raw imputed management forecast from models (1)-(3) in Table 2, and MF error is the difference between EPS and MF for periods with an MF. EPS surprise and MF surprise are the EPS and MF minus the imputed raw market expectation from models (1)-(3) in table 2 in raw USD (multiplying the predicted variable in these models by the standard error of EPS by firm). Standardized EPS and MF surprise are the standardized variables used in the structural model.

	Mean	SD	Min	5%	25%	Median	75%	95%	Max
Forecast Frequency	0.54	0.31	0.08	0.09	0.27	0.5	0.83	1	1
Number of years	9.94	1.77	1	7	9	10	11	12	12
EPS Surprise	0.00	1	-4.04	-2.00	-0.41	0.16	0.56	1.38	2.77
MF Surprise	0.21	0.37	-2.84	-0.38	0.1	0.25	0.37	0.67	2.77
Total Assets USD	19,564	114,182	11.23	137.44	622.66	2,066	7,056	56,846	2,261,264
Market Cap USD	7,755	24,551	7.73	145.75	584.25	1,679	5,224	28,938	391,708
Book to Market	0.59	0.42	-0.26	0.12	0.30	0.49	0.76	1.34	3.16
Debt/TA	0.55	0.21	0.10	0.20	0.40	0.55	0.70	0.91	1.21
R&D/TA	0.05	0.07	0	0	0.00	0.02	0.07	0.16	0.94
CAPEX/TA	0.05	0.04	-0.03	0.01	0.02	0.03	0.06	0.13	0.41

As a preliminary reduced-form analysis to capture facts about disclosure behavior, we run in Table 3 the following logistic regression

$$d_t = \mathbf{a}'X_t + b_1 t + b_2 EPS_t + b_3 nb_{nd_t} + b_4 d_{t-1} + v_t, \tag{1}$$

where d_t is an indicator variable equal to one when at least one forecast about EPS_t is observed, X_t is a set of characteristics including number of analysts, log total assets, log market capitalization, book-to-market, debt-to-assets, R&D to assets and capital expenditures to assets, t is a time-trend defined as calendar year minus 2004, EPS_t is the realized forecasted EPS, nb_{nd} is the number of consecutive periods without forecast starting in the prior period.

Consistent with DJK, the realized EPS is positively associated with the existence of a forecast, which indicates that managers select positive information. The variables nb_{nd} and d_{t-1} provide further evidence on the mechanism in Einhorn and Ziv (2008), since their model predicts that the number of consecutive periods of non-disclosure starting in the last period reduces the likelihood of disclosure in the current period. Indeed, the coefficient on these two variables is positive. Their model also implies that the effect becomes weaker over time, which is seen from the greater coefficient on d_{t-1}.

We also observe that firm characteristics have a significant, but small, incremental association to disclosure beyond the variables that are explained by the structural model. The adjusted R^2 changes from 32% to 35% when including other characteristics so that stickiness and selection explain most of the cross-sectional variation in disclosure. The pseudo R^2 does not change much when excluding other characteristics.[6] If characteristics are to be considered, the results suggest to include the number of analysts and size (market capitalization), as well as a time-trend to capture an increase in forecasting propensity over time. Hence, in some specifications, we will control for size and number of analysts as determinants of information endowment. These determinants may also affect other structural parameters or have a non-linear effect. To (partly) address this, we shall also estimate the model separately as in Li (2013) and Gayle, Li, and Miller (2015b) dividing the sample into groups of firms, ranked as above-median (H) to below-median (L) when they enter the sample on book-to-market and size. These analyses are given in additional robustness checks in Section 6.

[6]The slight increase is primarily caused by the larger sample when excluding characteristics for which data can be missing - especially $R\&D$. If we compare identical samples, the pseudo R^2 is nearly unchanged.

Table 3. Logit Analysis

This table reports the results of a logistic regression where the dependent variable is an indicator equal to one in the presence of a management forecast (zero otherwise) during a quarter-firm observation. Number of analysts is the number of unique analysts with updated EPS forecasts in the consensus window. All variables are measured at the quarter prior to the quarter to be forecasted, except for Market capitalization which is calculated with the closing price and the number of shares one day after the earnings announcement preceding the forecasted quarter. The variables R&D, Capital expenditures and financing CF are divided by total assets. ND periods is a variable indicating the number of periods since the last disclosure (first observation is set to zero). Standard errors clustered by firm are in parenthesis with significance at the 1%, 5% and 10% level indicated as *, ** and ***, respectively.

	Management Forecast		
	(1)	(2)	(3)
N Analysts	-0.03**	-0.04***	
	(0.01)	(0.01)	
Log TA	-0.35*	-0.13	
	(0.21)	(0.17)	
Log MCAP	0.73***	0.44**	
	(0.21)	(0.17)	
Book-To-Market	0.23	-0.25	
	(0.35)	(0.29)	
Debt/TA	0.18	0.056	
	(0.49)	(0.39)	
R&D/TA	-3.33***	-2.47***	
	(1.22)	(0.89)	
Capital Exp./TA	-0.93	-0.72	
	(1.68)	(1.37)	
Time Trend	-0.14***	0.07***	0.08***
	(0.02)	(0.02)	(0.01)
EPS		0.20***	0.32***
		(0.05)	(0.04)
NB_ND		-0.48***	-0.63***
		(0.07)	(0.05)
Disclosure Lag		1.57***	1.48***
		(0.17)	(0.13)
Pseudo R2	0.078	0.32	0.35
Obs.	3,250	2,866	5,947

3 Static Model

In this section, we present a stand-alone implementation of the static disclosure theory developed by Dye (1985) and Jung and Kwon (1988). This implementation is intended both as a first step to building a richer structural model (which we develop in more detail in the next section) but also, importantly, as an easily implementable tool that captures most of the first-order aspects of disclosure behavior. To begin with, we summarize three stylized facts, described earlier, that this model aims to explain.

First, firms do not consistently issue forecasts over time, implying that there are some underlying frictions that stop unravelling to full disclosure (Viscusi 1978; Grossman 1981). Second, comparing EPS surprises for periods with versus periods without management forecasts, disclosure periods are associated with positive earnings surprises while non-disclosure periods are associated with negative earnings surprises. This suggests that disclosures do not come entirely at random and unfavorable information is selectively withheld. Third, the dispersion of management forecasts is significantly lower than the dispersion of earnings, with a standard-deviation in forecast surprises at about a third of the standard deviation of earnings surprises. Hence, managers are imperfectly informed about earnings when forecasting over horizons of up to a year. Together, these stylized facts suggest that a classic disclosure model á la Dye (1985), in which managers selectively disclose but do not always have information, is a reasonable starting point to examine management forecast data.

A brief description of the model is in order and, for expositional purposes, we state the model in terms of a single firm observed over time $t = 1, \ldots, T$. In each period, the manager may receive information about future earnings, which we model as an expectation $x_t \sim N(0, \sigma_x^2)$ about earnings surprises $e_t \sim N(0, \sigma_e^2)$.[7] With probability p, the manager bears a friction and does not observe x_t or, equivalently, has soft information

[7]We specify the manager's information directly in terms of a posterior expectation about earnings to save space. Formally, assume that the manager receives some signal s_t about e_t, and define $x_t \equiv \mathbb{E}(e_t|s_t)$. It then follows from standard results about posteriors that $cov(e_t, x_t) = Var(x_t) = \sigma_x^2$. The variance-covariance matrix of the vector $(x_t, e_t)'$ follows readily.

that cannot be disclosed or does not care about current price. With probability $1-p$, the manager is not subject to the friction and may either truthfully disclose ($d_t = 1$) or withhold ($d_t = 0$). The disclosure choice maximizes the market price, which we assume to be linear in the post-disclosure market expectation $\mathbb{E}(x_t|d_t, d_t x_t)$. The market expectation conditional on non-disclosure is denoted $P^{nd} \equiv \mathbb{E}(x_t|d_t = 0)$.

In the unique equilibrium of this game, there exists a threshold y such that an informed manager discloses if and only if $x_t \geq y$. The threshold y is determined by the indifference condition

$$P^{nd} = E(x_t|d_t = 0) = y. \tag{2}$$

From equation (7) in Jung and Kwon (1988), specialized to normal distributions, the threshold satisfies

$$-\frac{p}{1-p}z = \int_{-\infty}^{z} \Phi(x)dx, \tag{3}$$

where $\Phi(.)$, with derivative $\phi(.)$, indicates the distribution of the standard normal and $z \equiv \frac{y}{\sigma_x}$ is the disclosure threshold standardized by the standard deviation of x.

Equation (3) defines an implicit relation $z = Z(p)$ that maps the probability of any friction to a standardized disclosure threshold z. The function $Z(.)$ is increasing, because a higher probability of the friction p makes it more likely that the firm was subject to the friction and did not strategically withhold. This increases the non-disclosure price causing more firms to withhold information.

Equation (3) is not yet in a form suitable for empirical analysis, because we observe a sample of forecast and realizations $(d_t, d_t x_t, e_t)$ but do not directly observe the true parameters (p, σ_x, σ_e). However, noting that the model implies that $d_t = 1_{x_t \geq y}$, we can derive several moment restrictions.

The expected probability of a disclosure predicted by the model is $\mathbb{E}(d_t)$ which, since $d_t = 1$ if and only if (a) the manager is not subject to the friction, which has probability

p, and (b) x_t is greater than y, which has probability $\Phi(-Z(p))$, implies

$$\mathbb{E}(d_t) = (1-p)\Phi(-Z(p)). \tag{4}$$

The left-hand of this moment can be estimated using the sample frequency of a forecast $\hat{d}$ and yields an implied estimated value for p. Of course, the greater the observed forecasting frequency, the lower the probability of the friction.

Next, we examine the prediction of the model for $\mathbb{E}(d_t x_t)$ which captures information from both the forecast frequency and the expected forecast.[8] Note that $\mathbb{E}(d_t x_t) = \mathbb{E}(d_t)\mathbb{E}(x_t | \frac{x_t}{\sigma_x} \geq Z(p_0))$ is the conditional expectation of a truncated normal, implying

$$\mathbb{E}(d_t x_t) = (1-p)\phi(Z(p))\sigma_x. \tag{5}$$

The left-hand side of this equation can be estimated by the sample average of $\hat{dx}$ over all observed forecasts and zeros for periods without a forecast. If, say, we observe a large average sample forecast, the moment condition requires that either (a) the probability of the friction is high, so that only managers with very good news would optimally disclose, or (b) the volatility of the private information is large, which reduces the level of the expectation relative to the ex-ante variance.

After substituting the sample moments, equations (4)-(5) form a system of two equations in two unknowns $(\hat{p}, \hat{\sigma}_x)$, to which we can supplement the volatility of realized earnings surprises to obtain an estimate the standard-error of realized earnings $\hat{\sigma}_e$. In summary, the method of moments estimator $(\hat{p}, \hat{\sigma}_x, \hat{\sigma}_e)$ of the parameters (p, σ_x, σ_e) is

[8] For the method of moments application developed here, the objective function would be the same if we had used the conditional moment $\mathbb{E}(x_t | d_t = 1)$, since, once we match the frequency moment $\mathbb{E}(d_t)$, $\mathbb{E}(x_t | d_t = 1) = \mathbb{E}(d_t x_t)/\mathbb{E}(d_t)$ (implying that matching frequency and conditional moment yields the same predictions as matching frequency and $\mathbb{E}(d_t x_t)$). Of course, because the objective functions under both set of moments have the same optimum by construction, it is also the case that the standard-errors of both methods are the same and the choice between conditional and unconditional moment is only a matter of exposition. The presentation in terms of $\mathbb{E}(d_t x_t)$ is slightly less cumbersome because it allows for a quicker derivation of the standard-errors without having to apply the delta method on the sample estimates of $(\mathbb{E}(d_t), \mathbb{E}(d_t x_t))$.

defined as the solution to the system of equations

$$\begin{cases} \hat{d} = (1-\hat{p})\Phi(-Z(\hat{p})) \\ \hat{dx} = (1-\hat{p})\phi(Z(\hat{p}))\hat{\sigma_x} \\ \hat{e} = \hat{\sigma}_e \end{cases} . \tag{6}$$

A valid solution (with positive $\hat{\sigma}_x$) requires the sample average forecasts to be positive, or $\hat{dx} \geq 0$. This equation has, at most, a unique solution, implying that the parameters of interest are identified.[9]

Proposition 1 (Identification) *For any $\hat{d} \in (0,1)$ and $\hat{dx}, \hat{e} > 0$ there is a unique $(\hat{p}, \hat{\sigma})$ solving equation* (6)*.*

The existence of a monotone mapping between the probability of disclosure $(1-p)\Phi(-Z(p))$ and p guarantees the identification of the parameters. Put differently, there is a single p that is consistent with a given frequency of disclosure. Implementing an estimation procedure with method of moments is then straightforward. Given $\hat{d}$, we can solve the first equation of (6) to obtain $\hat{p}$. Then, we can plug $\hat{dx}$ and $\hat{p}$ in the second equation of (6) to obtain $\hat{\sigma}_x = \frac{(1-\hat{p})\phi(Z(\hat{p}))}{\hat{dx}}$ and $\hat{\sigma}_x$ is simply the sample standard error of realized earning surprises.

In the sample, $\hat{d} = .5537$ and $\hat{dx} = .1136$ and $\hat{e} = 0.7324$. Plugging these empirical moments in (6) yields the estimates as presented in Table 4.

The interpretation of these coefficients is straightforward. The manager faces a disclosure friction roughly one of out four reporting periods. In these cases, the manager does

[9]One complication in (6) is that finding a solution requires solving a non-linear equation. However, the solution can be very closely approximated as a closed-form expression of the sample moments that can be plugged in any statistical package. Setting

$$\hat{p}_a = 0.9\hat{d}^2 - 1.9\hat{d} + 1 \quad \text{and} \tag{7}$$

$$\hat{\sigma}_a = \frac{\hat{dx}}{\max(0, -0.35\hat{d}^2 + 0.77\hat{d} - 0.02)}, \tag{8}$$

the maximum approximation error $|\hat{p}_a - \hat{p}|$ and $|\hat{dx}/\hat{\sigma}_a - \hat{dx}/\hat{\sigma}|$ remains below .01 for all $\hat{p} \in [0,1]$, which implies that we can use these approximations as very good substitutes for the correct estimator $(\hat{p}, \hat{\sigma})$.

Table 4. Method of Moment Estimates

Standard errors are reported in parentheses.

p	σ_x	σ_e
0.2293	0.3198	0.7324
(.0053)	(.0095)	(.0092)

not know or cannot disclose information. The estimate $\frac{\hat{\sigma}_x^2}{\hat{\sigma}_e^2} = 44\%$ measures the quality of the manager's information: the manager's private information, at the forecast date, is able to explain less than half of the variation in earnings surprises. Qualitatively, this result suggests that managers have significant uncertainty about the upcoming earnings when they disclose their forecasts.

A benefit of the structural approach is that the assumptions allow us to quantify the economic consequences of strategic behavior and their sensitivity to primitives of the model. This would very difficult (if not impossible) using a reduced form approach. In particular, we can use our estimates to measure the probability that the manager withholds a forecast strategically and calculate how changes in the quality of the manager's information would affect the disclosure strategy. We can also measure the amount of information loss due to the manager's strategic behavior. Lastly, we can estimate the over-valuation incurred by a naive investor who ignores that managers strategically conceal information (Kartik, Ottaviani, and Squintani 2007).

Consider first the probability that the manager conceals his information strategically, hereafter, the probability of strategic withholding. From equation (6), we can estimate the probability of strategic withholding ξ_1 from

$$\hat{\xi_1} = (1 - \hat{p})\Phi(Z(\hat{p}). \tag{9}$$

In equation (9), if the friction is very small, unraveling to full disclosure predicts that nearly all information is reported; vice-versa, if the friction is very large, no manager will be reporting strategically. The maximum level of strategic reporting is attained for intermediate values of the friction.[10] This observation has an important consequence: because strategic withholding is not a monotonic function of the disclosure frequency, disclosure frequency is not a proxy for strategic disclosure.

Second, to measure the information loss caused by strategic withholding, we can estimate the average residual variance of earnings, conditional on all public information. Using results in Dye and Hughes (2017), the average residual variance of earnings boils down to

$$EVar(\frac{e}{\sigma_e}|d) = 1 - (1-p)\Phi(-Z(p))\frac{\sigma_x^2}{\sigma_e^2}. \tag{10}$$

The residual variance will be zero if the manager is perfectly informed ($\frac{\sigma_x}{\sigma_e} = 1$) and non-strategic. By contrast, the residual variance will approach one if the quality of the manager's information is poor ($\sigma_x \to 0$) or the disclosure friction is very likely ($p \to 1$). If the manager is non-strategic, removing the term $\Phi(-Z(p))$, the residual variance will be

$$E_{ns}Var(\frac{e}{\sigma_e}|d) \equiv 1 - (1-p)\frac{\sigma_x^2}{\sigma_e^2}.$$

To measure the overvaluation of earnings incurred by a naive investor given no-disclosure, we can use

$$\frac{E(e) - E(e|d=0)}{\sigma_e} = -E(\frac{e}{\sigma_e}|d=0). \tag{11}$$

Our estimates of these three measures are presented in Table 5.

Table 5. Prevalence and Impact of Strategic Withholding.
(Standard errors in parentheses, computed based on Delta Method)

ξ_1	$E[Var(\frac{e}{\sigma_e}|d)]$	$E_{ns}[Var(\frac{e}{\sigma_e}|d)]$	$E(\frac{e}{\sigma_e}|ND)$
0.2162	0.8944	0.8531	-0.2525
(0.0026)	(0.0073)	(0.0102)	(0.0085)

[10] It is uncertainty about the friction, not its level, that determines strategic withholding. This feature of strategic models is relatively transparent in the original DJK, yet lost to empirical studies that use frequency of disclosure as a proxy for discretionary disclosure. Note that the same aspect is also true in models of noise trading, where the maximum level of information asymmetry is attained when there is maximal uncertainty about the noise trade (Bertomeu, Beyer, and Dye 2011).

The manager withholds information strategically 22% of the time. On average, the residual variance, at the forecast date, is 89% of the unconditional variance of earnings, e. If the manager were non-strategic, the residual variance would fall to 85% . This effect is modest relative to the estimated probability of strategic withholding and reveals a key quantitative property of voluntary disclosure theory. Even when the manager withholds, the market partially reflect the information in price by considering the negative information conveyed by a non-disclosure. Ignoring strategic withholding would have severe consequence. On average, a naive investor would over-value earnings given non-disclosure by roughly 25%.

To conclude this section, we perform several counterfactuals, i.e., analyses of what would occur if the characteristics of the friction changed. Specifically, we study the consequences of improving the manager's information (increasing σ_x), or reducing the incidence of disclosure frictions (i.e, reducing p). The result of this exercise is presented in Figure 2 and are in line with comparative statics derived in the existing theoretical literature (Dye 1985; Jung and Kwon 1988; Dye and Hughes 2017).

The qualitative lessons emerging from this analysis are:

i. the probability of withholding is relatively insensitive to p, when evaluated around $\hat{p}$ and, in fact, the estimated $\hat{p}$ achieves close to the maximal probability of strategic withholding;

ii. the residual variance of earnings does not change much as one varies p, because the information of the manager is noisy and prices rationally anticipate the probability that the manager is strategic;

iii. by contrast, the residual variance is highly sensitive to σ_x, consistent with the notion that σ_x represents the quality of the manager's information;

iv. the price of non-disclosure displays a high sensitivity to both p and σ_x: the less likely the friction, the stronger the negative selection effect associated with non-disclosure; the better the manager's information, the stronger the price reaction to non-disclosure.

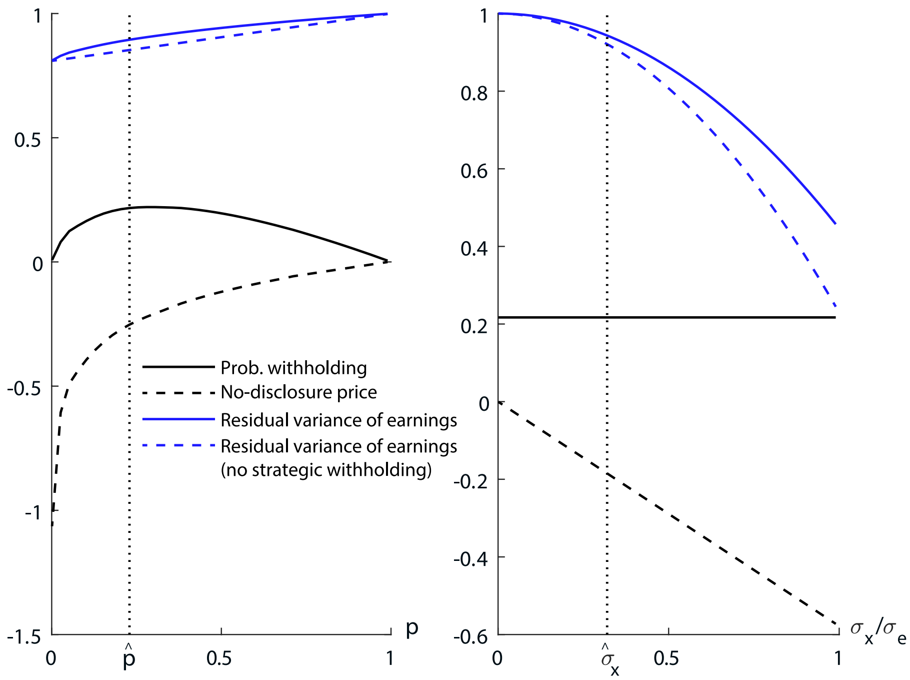


Figure 2. Comparative Statics.

4 Multi-period model

We expand next on the static model to explain two additional patterns in observed disclosure behavior. In doing so, our primary objective is to show how the Dye-Jung-Kwon (DJK) model can be used as a flexible framework to incorporate features of interest to the researchers and, in particular, offer one of many possible empirical applications of this model. Specifically, we focus on two additional empirical facts that are not the focus of the static model.

First, disclosures are persistent, namely, the more a firm has disclosed in the past, the more likely it is to continue disclosing in the future. To illustrate this, figure 3 documents the conditional probability of disclosure conditional on a disclosure at date t, revealing an exponential decay in disclosure behavior over time. This suggests that frictions may

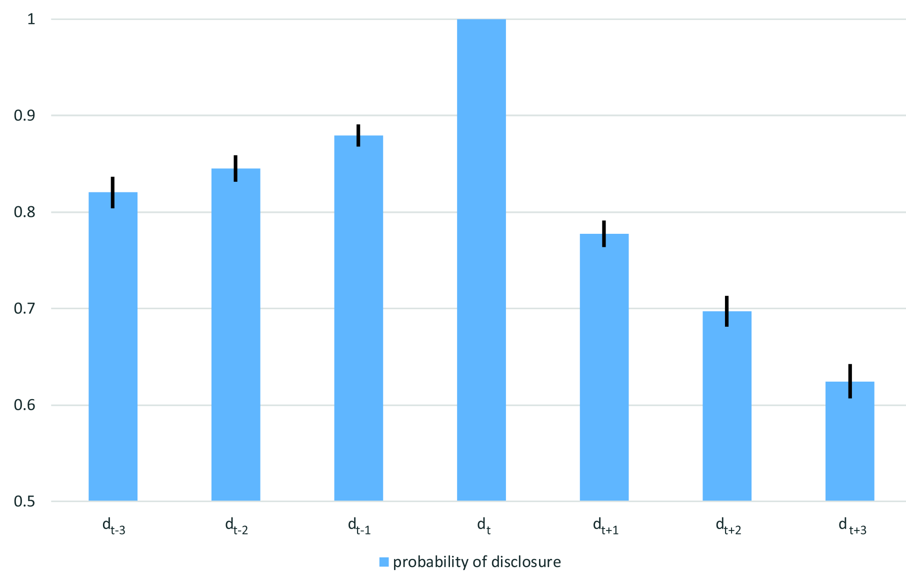


Figure 3. Disclosure frequency conditional on $d_t = 1$ (95% confidence interval in solid line).

exhibit serial correlation and, if the friction is not directly observable but must be inferred from observed disclosures, this correlation may reflect as time-varying strategic incentives to withhold.

Second, the DJK model is a stylized model that predicts a truncation of the distribution of forecasts, with no forecast being made below the disclosure threshold. This is, again, not what we observe empirically (see Figure 4) and suggests there are unobserved factors affecting individual choices to withhold. As a result, the non-disclosure price, and, hence, the equilibrium disclosure strategy, must be modelled to reflect a rational anticipation of these factors.

To build an empirical model of noisy persistent forecasting, we follow Einhorn and Ziv (2008) to allow for serially correlated information endowment. We use the notation of the static model but explicitly denote $\mathbb{E}_t(.)$ the expectation conditional on public information prior to period t. Each period, the manager may be uninformed, which we represent by a serially-correlated binary variable $\rho_t \in \{0, 1\}$. Let $p_t \equiv \mathbb{E}_t(\rho_t) \in (0, 1)$ denote the probability the manager is uninformed (as perceived by investors); note that the static model is nested as a special case when assuming that ρ_t is i.i.d. and $E_t(\rho_t) = p$ is not a function of t.

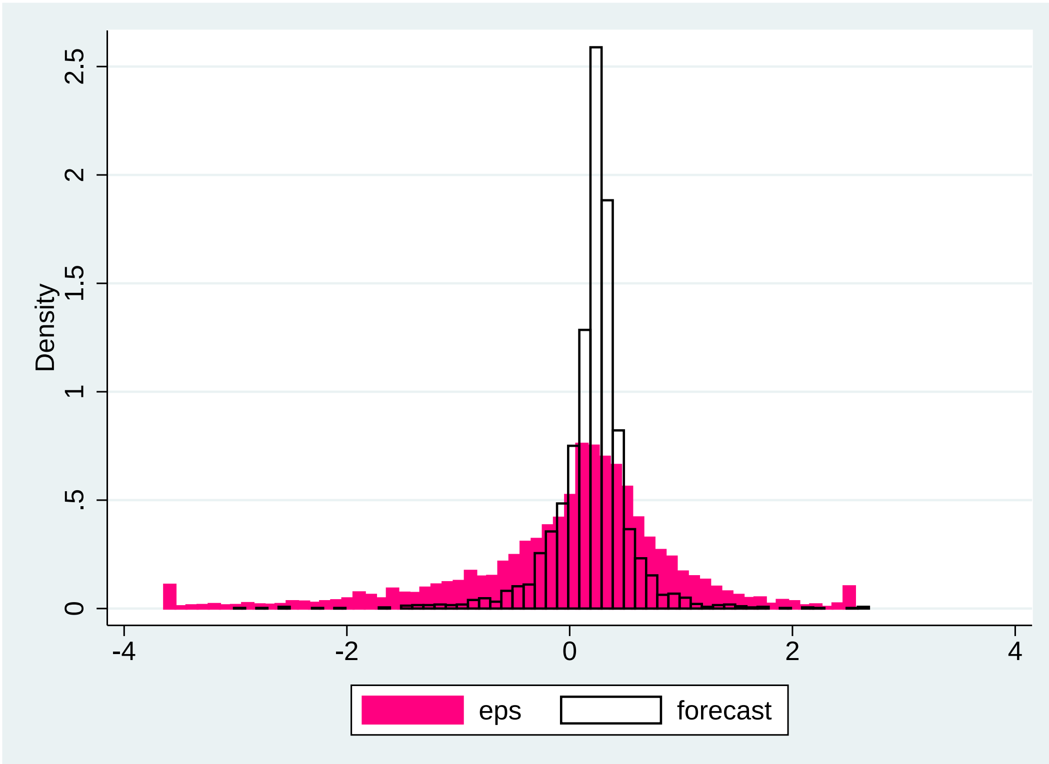


Figure 4. Histogram of EPS and MF

As before, the manager does not observe x_t when $\rho_t = 1$, in which case no disclosure can be made ($d_t = 0$); otherwise, when $\rho_t = 0$, the manager chooses whether to disclose, choosing $d_t \in \{0, 1\}$ to maximize the current market expectation $\mathbb{E}(x_t|d_t, d_t x_t)$ where $P_t(x_t) = \mathbb{E}(x_t|d_t = 1, x_t)$ is the price conditional on disclosure and $P_t^{nd} = \mathbb{E}_t(x_t|d_t = 0)$ is the non-disclosure price conditional on all prior public information.

As noted earlier, the standard DJK model predicts that no disclosure below a certain threshold would ever be observed. By contrast, we assume that there is some unobserved noise in the manager's action that may affect disclosure incentives. We model the manager's preference as in discrete choice theory (McFadden 1973, 1980) modified to incorporate price motives. When informed, the manager chooses whether to disclose to maximize his current utility

$$u_t(d_t|x_t) = \alpha(d_t P_t(x_t) + (1 - d_t)P_t^{nd}) + d_t\beta + d_t\epsilon_t, \tag{12}$$

where β is a fixed disclosure benefit, $\alpha P_t(x_t)$ captures a short-term price motive and ϵ_t is a standard normal white noise that is observed by the manager and captures other factors affecting the disclosure decision.[11] β could be interpreted, for example, as the expected litigation costs induced by non-disclosure.

Investors update their expectation about the manager's information endowment based on the multi-period model of Einhorn and Ziv (2008). The information endowment $\rho_t \in \{0,1\}$ follows a Markov chain with transitions given by $\Pr(\rho_{t+1} = 1|\rho_t = 1) = \Phi(k_{t,0})$ and $\Pr(\rho_{t+1} = 1|\rho_t = 0) = \Phi(k_{t,1})$, where $\Phi(.)$ is the c.d.f. of the standard normal. In the baseline model, we set $k_{t,j} = k_j$ for $j = 0, 1$ with $k_0 \geq k_1$, as assumed by Einhorn and Ziv. In additional analyses, we also consider time-series variation in the probability of information endowment, so the transitions obey $k_{t,j} = k_j + \gamma' X_t$ where X_t can be observable covariates.

At the end of each period, a normally-distributed public signal $e_t = x_t + v_t$, representing the firm's earnings, realizes where v_t captures uncertainty realized after the management forecast is made and is orthogonal to x_t. The process for forecast and earnings surprises is then given by

$$\begin{pmatrix} x_t \\ e_t \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & \sigma_x^2 \\ \sigma_x^2 & \sigma_e^2 \end{bmatrix} \right), \tag{13}$$

where $cov(x_t, e_t) = cov(x_t, x_t + v_t) = Var(x_t) = \sigma_x^2$ captures the amount of private information known in advance by the manager.

Given P_t^{nd}, an informed manager discloses his signal if and only if $u_t(d_t = 1|x_t) \geq u_t(d_t = 0|x_t)$. The non-disclosure price P_t^{nd} is consistent with the manager's policy if the following condition is satisfied.

---

[11] It is well-known in discrete choice theory that the noise term in the utility function cannot be non-parametrically identified, see McFadden (1973) and McFadden (1980). Further, parameters can only be identified per unit of standard deviation of the noise term. Some of the research in this area uses an extreme value distributions for ϵ_t for practicality. However, because the distribution of ϵ_t also affects the pricing equation in the structural model, extreme value distributions do not yield closed-form solutions in our setting.

Lemma 2 *Let $p_t = p$, then the non-disclosure price is the unique negative solution to the fixed point $\Gamma(P_t^{nd}) = P_t^{nd}$ where*

$$\Gamma(y) = \frac{1-p}{p} \int x\Phi(-\alpha x - \beta) \frac{1}{\sigma_x} \phi(\frac{x+y}{\sigma_x}) dx, \tag{14}$$

where $\Phi(.)$ and $\phi(\cdot)$ are the c.d.f. and p.d.f. of the standard normal, respectively.

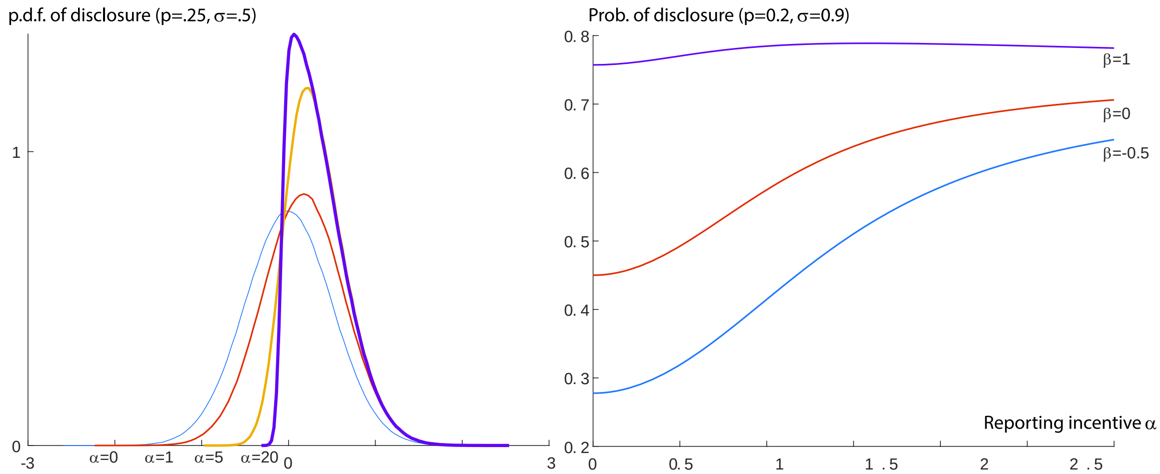
The literature considers several alternative frictions to explain the failure of unraveling. Unlike Einhorn and Ziv (2008), in our setting, the firm is not directly bearing disclosure costs. By contrast, we focus our estimation on uncertainty about the manager's information endowment which can, on its own, explain serial correlation in disclosure behavior. This restriction is also imposed to focus on a single type of friction and distinguish from other studies focusing on estimating costly voluntary disclosure models anchored on Verrecchia (1983).[12]

Our empirical model includes a stochastic disturbance ϵ_t in the preference so that, from the perspective of investors, the manager does not follow a strict threshold equilibrium - unlike standard disclosure models. Equation (14) is a fixed point that determines the price conditional on non-disclosure as a function of the structural parameters (α, β, σ), as well as the time-varying market belief about information endowment p_t.

In the left-hand side of Figure 5, we illustrate the implication of the model on the distribution of disclosures by plotting the density of disclosure for various price incentives. At one extreme, at $\alpha = 0$, the manager ignores the effect of disclosure on price, so that the manager does not select disclosures that increase price. As α increases, the theoretical equilibrium threshold of Jung and Kwon (1988) without noise pins down disclosure behavior. The model predicts some selection over favorable disclosures coexisting with disclosures driven by non-price incentives, $\beta + \epsilon_t$. Price motives cause observed disclosures

[12] As noted by Einhorn and Ziv, incorporating firm-level disclosure costs is non-trivial because prices will be forward-looking and incorporate future disclosure costs; computationally, solving this generalized model would then require deriving prices by value-function iteration. See also Bertomeu, Beyer, and Taylor (2015), Zhou (2016) and Cheynel and Liu-Watts (2016) for recent estimations of disclosure costs.

Figure 5. Properties of the voluntary disclosure equilibrium



to become more skewed and, at $\alpha = 20$, disclosures are a truncated normal distribution similar to the theoretical threshold equilibrium in standard disclosure models.

To further illustrate the effect of price incentives, we plot the probability of disclosure as a function of α in the right-hand side of Figure 5. For the most part, the probability of disclosure is increasing in α because, holding the probability of non-disclosure fixed, the direct effect of increasing α is to increase the importance of $P_t(x_t) - P_t^{nd}$, which is positive on average and thus makes the manager more willing to disclose.

Strictly speaking, the probability of disclosure is not always monotonic in the price incentive α. An increase in α may cause more adverse selection and withholding because managers are more exposed to their private information and are less willing to reveal it when it's relatively adverse. In turn, this can cause a decrease in the non-disclosure price and reduce the probability of disclosure (top curve of Figure 5). In the limit, the probability of disclosure is $(1 - p_t)\Phi(\beta)$ when α is small, and, conditional on being informed, depends only on the noise in the preference and the parameter β. When α is large, the model converges to DJK with benefits β, so that the probability of disclosure is $(1 - p_t)\Pr(x_t \geq \tau_\beta)$ where τ_β is the disclosure threshold within the DJK model extended to incorporate the intercept β.[13] When β is large, the probability of disclosure is greatest absent price incentives, while, when β is low, the opposite happens.

The model is estimated by maximum likelihood. Below, we write the likelihood of an observation $(d_t x_t, e_t)$ in period t given the two possible disclosure outcomes:

(i) if $d_t = 1$, so that both the forecast x_t and realized earnings e_t are observed, the conditional distribution of the forecast is given by $x_t | e_t \sim N(\frac{\sigma_x^2}{\sigma_e^2} e_t, \frac{\sigma_x^2 \sigma_v^2}{\sigma_e^2})$ which yields a likelihood

$$L_t(x_t, e_t) = \frac{(1-p_t)K_t(x_t)}{\frac{\sigma_x \sigma_v}{\sigma_e}} \phi\left(\frac{x_t - \frac{\sigma_x^2}{\sigma_e^2} e_t}{\frac{\sigma_x \sigma_v}{\sigma_e}}\right), \tag{15}$$

where $K_t(x) = \Pr(\alpha x_t + \epsilon_t > \alpha P_t^{nd} - \beta | e_t, x_t = x)$ is the probability of withholding conditional on observing x. Given disclosure in period t, in period $t+1$, the probability of the friction is updated to $p_{t+1} = \Pr(\rho_{t+1} = 1 | \rho_t = 0) = \Phi(k_1 + \gamma' X_{t+1})$;

(ii) if $d_t = 0$, the likelihood can be written as

$$L_t(0, e_t) = p_t + (1-p_t) \Pr(\alpha x_t + \epsilon_t < \alpha P_t^{nd} - \beta | e_t). \tag{16}$$

In the absence of disclosure, and after observing the earnings realization, investors update the probability the manager will be uninformed in the next period, given that the probability the manager was uninformed in period t is $p_t / L_t(0, e_t)$ and the probability the manager was informed in period t with probability $1 - p_t / L_t(0, e_t)$. This leads to

$$p_{t+1} = \frac{p_t}{L_t(0, e_t)} \Phi(k_0 + \gamma' X_{t+1}) + \left(1 - \frac{p_t}{L_t(0, e_t)}\right) \Phi(k_1 + \gamma' X_{t+1}). \tag{17}$$

[13] Specifically, this can be easily derived by modifying the threshold equation in DJK as

$$p_t(\tau_\beta + \beta) = \int_{-\infty}^{\tau_\beta} F(x) dx.$$

.

We can now write the (log) likelihood of a time-series of disclosures as given by

$$\mathcal{L}((d_t x_t, e_t)_{t=1}^T) = \sum_{t=1}^{T} \ln L_t(d_t x_t, e_t),$$

where p_t can be updated recursively using equation (17).

Table 9 reports baseline results. We analyze two specifications: column (1) estimates the baseline model and column (2) estimates a generalized version of the model where the probability of information endowment depends on size, analyst coverage and a trend. Later on, in robustness checks, we estimate the model by subgroups of size and analyst coverage to allow all coefficients to change as a function of these characteristics.

In both specifications, the probability of information endowment is very persistent. Specifically, we can measure persistence as $\Delta = \Phi(k_0 + \gamma'\overline{X}) - \Phi(k_1 + \gamma'\overline{X})$, where $\overline{X}$ is a vector of zeros for column (1) and the average analyst coverage, size and trend variable in column (2).

Next we assess the amount of private information observed by the manager. Note first that the standard deviation of forecast surprise is .37 (Table 2), accounting for a little over a third of the information in earnings. The observed sample of forecasts is a truncation of the true information observed by the manager (given that unfavorable information tends to be withheld) so the sample standard deviation of forecasts is not an unbiased estimate of the standard deviation of the manager's information. Specifically, truncations of the normal exhibit lower variance and, thus, .37 is a lower bound on the standard deviation of the manager's private information. The model yields an estimate of σ_x that accounts for this truncation. With point estimates of around .44, it remains slightly above but close to the standard deviation of observed forecasts.

These numbers are quite remarkable because they suggest that managers are able to reveal almost half of the residual uncertainty in earnings over six months prior to the end of a year.

The small difference between .44 and .37 suggests that price incentives only play a

Table 6. Baseline estimates

This table reports results from the baseline model in section 2. Column (1) is estimated for all firms in the sample. Column (2) is estimated for all firms assuming that the probability of information endowment is linear in Time trend, Nb. of analysts and log MCAP. Time trend is measured in the number of years since 2004 divided by 10. Nb. of analysts is measured as the number of analysts revising a forecast during the consensus window, divided by 10. k_0 (k_1) is the probability of not being informed in the next period conditional on being uninformed (informed). σ is the precision of the manager's private information. α is the manager's weight on current price. and β is the firm's exogenous preference for disclosure. Standard errors are in parenthesis are calculated using the information matrix. The Log-likelihood is reported scaled by the number of firms.

	(1)	(2)
k_0	1.823	3.432
	(0.075)	(0.181)
k_1	-1.256	0.156
	(0.032)	(0.142)
$\frac{\sigma_x}{\sigma_e}$	0.444	0.444
	(0.007)	(0.007)
α	2.667	2.680
	(0.114)	(0.112)
β	0.613	0.595
	(0.051)	(0.049)
σ_e	0.732	0.732
	(0.006)	(0.006)
Time trend		-0.678
		(0.100)
Nb. of analysts		0.208
		(0.039)
log MCAP		-0.182
		(0.022)
Log Lik	-14.92	-14.83
Obs.	6,857	6,857

moderate role, as more truncations would create a greater wedge between the standard deviation of forecasts and the standard deviation of information. Nevertheless, price incentives α are non-zero, are around 2.67. To interpret these numbers, note that $\alpha\sigma_x$ is the total price incentive effect of a one standard deviation increase in the manager's private information. By and large, this term is close to one, so that price incentives have an impact similar to one standard deviation change in the random shock ϵ_t. Hence, price incentives are as important to explain disclosure patterns as the noise component of the utility function.

Price incentives may also be compared to the point estimate for β, which varies between .61 to .59. The effect of this constant is sizeable, being about two-thirds of price incentives for one standard deviation in the private information for small firms, and a little over a third of a standard deviation for large firms.

In model (2), we examine the effect of three covariates, a time trend, analyst coverage and size, assuming that these covariates enter the model by affecting the information endowment process. The estimates suggest that managers have become more informed over the sample period. This result may be explained by the gradual phasing in of better internal controls, as required by the Sarbanes-Oxley Act, as well as, possibly, a general trend (that may have started prior to 2004) to improve information technology and forecasting systems.

Also consistent with this conjecture, larger firms tend to more informed, since larger firms receive greater scrutiny and may benefit from greater returns to scale on their forecasting efforts. Managers with a greater number of analysts are more likely to be uninformed. We cannot easily know the reason for this result, but two explanations (which we cannot distinguish with our data) may be considered, (i) selection in analyst coverage, in that analysts are compensated based on their informational advantage and thus may prefer to avoid firms in which their information output is superseded by voluntary forecasts, or (ii) the fact that being informed is relative to the current information environment so a better information environment mechanically translates into less informed managers. These two explanations may be particularly relevant in our sample period, since it is after Regulation Fair Disclosure and thus all management forecasts must be communicated to the public at the same time as they are released to analysts.[14]

5 Additional Analyses

Table 7 contrasts key moments in the data versus moments in the baseline model in column (1), as an indirect means of assessing the fit of the model. Our findings are as follows. The model slightly over-predicts the disclosure probability relative to the data. Based on average earnings, we also find evidence suggestive of strategic withholding: earnings are higher during disclosure periods than during non-disclosure periods, and the difference is about 13% of the standard deviation of earnings.

The model under-predicts the variation of forecasts relative to the data ($Var(d_t x_t)$). Also, the dispersion of forecasts suggests that managers observe relevant private information about earnings but their information contains significant noise at the disclosure date. On average, the residual variance conditional on d_t is roughly 93% of the variation in earnings, both in the data and in the model. This relatively modest decrease in the uncertainty attained by management forecasts has a two-fold explanation: a significant likelihood of non-disclosure and noise in management forecasts.

The persistence of disclosure, which was unmodelled in our static model, emerges clearly in the data, with the probability that a firm discloses in the next period being nearly 78% when there was a disclosure in the current period. The model closely matches this probability.

We examine next several counterfactuals, by, first, changing strategic behavior and, second, altering the information endowment process. We aim to predict how changes in some of the model parameters would affect some outcome variables.

[14]There is evidence on this question from exogenous shocks to the number of analysts which, therefore, are not confounded by selection and which suggest that selection is not the only explanation why analyst coverage is associated to price (Balakrishnan, Billings, Kelly, and Ljungqvist 2014). Nevertheless, in our observational data set, selection may also contribute to the estimated parameters.

Table 7. Moments: Data vs Model

	$\Pr(d=1)$	$\frac{E(e\|d=1)-E(e\|d=0)}{\sigma_e}$	$\sqrt{\frac{Var(x\|d=1)}{\sigma_e^2}}$	$\Pr(d_t=1\|d_{t-1}=1)$	$\sqrt{\frac{EVar(e\|d)}{\sigma_e^2}}$
Data	0.553	0.133	0.358	0.777	0.928
	(0.012)	(0.015)	(0.012)	(0.007)	(0.010)
Model	0.572	0.149	0.321	0.761	0.923
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)

We first develop two measures of strategic non-disclosure, where we use the term of "strategic" for non-disclosures that were chosen by the manager rather than being due to the fact the manager was uninformed. First, consistent with the static model, we measure the unconditional probability that the manager is informed and withholds information, which we define as

$$\xi_1 = \mathbb{E}((1-\rho_t)\Pr(\alpha x_t + \epsilon_t < \alpha P_t(0) - \beta)).$$

To estimate ξ_1, note that the above equation can be written as

$$\xi_1 = \mathbb{E}(1 - d_t - p_t), \tag{18}$$

so that an estimator $\hat{\xi}_1$ can be computed as the difference between the sample frequency of non-disclosure $1-\hat{d}$ and the average probability of being uninformed $\hat{p}$ (which we compute each period when estimating the model with MLE and simply need to average over the sample). Intuitively, $\hat{\xi}_1$ is the estimated probability of withholding relative to a counter-factual in which the manager always discloses when informed.

The previous measure captures the probability of withholding. But there are two factors driving that probability: price incentives, measured by α, and the random component, $Var(\epsilon_t)$. To isolate the effect of price incentives, we construct a measure which examines the probability of withholding relative to a counter-factual in which the manager has no price motives, i.e., with $\alpha = 0$. Specifically, we define ξ_2 as

$$\xi_2 = \xi_1 - \mathbb{E}(1 - \rho_t) \Pr(\epsilon_t < -\beta).$$

Note that $p_t = \mathbb{E}_t(\rho_t)$ so that, from iterated expectations, $\mathbb{E}(p_t) = \mathbb{E}(\rho_t)$; hence, simplifying terms,

$$\xi_2 \;\; = \;\; \mathbb{E}(1 - d_t + (1 - p_t)\Phi(\beta)) - 1,$$

which we can recover from the estimated preference parameter $\hat{\beta}$ and estimates for $\hat{d}$ and $\hat{p}$

Table 8 reports the probability of strategic non-disclosure ξ_1. In the baseline, this probability is small, close to 11% which is significantly lower than that obtained under the static model. Theoretically, this finding is closely related to the estimated high persistence of the information endowment and the high precision of the manager's private information, which imply that investors quickly learn about the information endowment even conditional on non-disclosure and, once they are well-informed, the model tends to feature less strategic withholding.

We also measure the strategic withholding driven by price motives, ξ_2. As noted in the theoretical section, price motives usually encourage more disclosure because they place greater weight on the (negative) price reaction to withholding. In the full sample price motives increase the unconditional probability of disclosure by about 5% in general.

Table 8 provides additional counterfactuals based on four nested versions of our baseline model. First, we consider the special case in which the manager has perfect information ($\sigma_x = \sigma_e$). Second, we shut down the fixed disclosure benefit by setting $\beta = 0$. Third, we consider the case in which the manager's disclosure preferences are deterministic ($Var(\epsilon_t) = 0$). Fourth, we consider the case in which the manager does not have a price motive ($\alpha = 0$). Based on the baseline estimates, we compute/simulate a number of statistics for these four models.

With regard to the probability of disclosure, the largest impact relative to the baseline is caused by the presence of random shocks. The random shocks decreases the probability of disclosure from .59 to .55. The presence of random shocks weakens the penalty for non disclosure as investors become less skeptical that the non-disclosure is strategic. In all models, average earnings are larger in disclosure periods, relative to the non-disclosure case; perhaps the core prediction of disclosure models. The gap between the average earnings with and without disclosure is largest when the manager is perfectly informed since it allows the manager to more selectively choose which information to disclose.

Seemingly counterintuitive, the price of non disclosure $E(e|ND)$ is lower in the baseline model relative to the model with no disclosure benefit $\beta = 0$, even though a higher disclosure benefit implies that more unfavorable types withhold. However, this result is a direct consequence of the minimum principle (Acharya, Demarzo, and Kremer 2011). The disclosure threshold in a standard Dye (1985) minimizes the price of non-disclosure. Any disclosure benefit (such as β) that perturbs the disclosure threshold away from that in Dye (1985), must necessarily lead to a higher price of non-disclosure. In terms of policy, an indirect consequence of disclosure regulations that boost disclosure benefits (or the cost of non disclosure), is that it will increase the price of non-disclosure, thus protecting naive investors who trade during non disclosure periods. Lastly, observe that the residual variance tends to increase when the manager lacks a price motive consistent with the fact that price motives stimulate disclosures.

The probability of withholding ξ_1 goes up significantly in the absence of a disclosure benefit ($\beta = 0$) from .114 to .24. It also increase, but more moderately, in the absence of price motives. The price-driven probability of withholding ξ_2 is negative in all models, but the model with deterministic preferences ($Var(\epsilon) = 0$). This result reinforces the point that stock price incentives stimulate disclosure.

Table 8. Counterfactuals

	$\Pr(d=1)$	$\frac{E(e\mid d=1)-E(e\mid d=0)}{\sigma_e}$	$\sqrt{\frac{EVar(e\mid d)}{Var(e)}}$	ξ_1	ξ_2
Baseline	0.554	0.133	0.928	0.114	-0.052
Perfect forecast ($\sigma_x = \sigma_e$)	0.541	0.617	0.652	0.109	-0.025
No disclosure benef. ($\beta = 0$)	0.433	0.215	0.939	0.236	-0.065
No noise in pref. (Var(ϵ_t)=0)	0.598	0.090	0.925	0.036	0.036
No price incentive ($\alpha = 0$)	0.517	0	0.950	0.131	0

6 Robustness Checks

6.1 Size and analyst coverage

We have seen in the logit analysis that observable covariates such as size or analyst coverage only modestly explain disclosure behavior relative to past disclosures and, indeed, including these covariates as determinants of the disclosure friction does not alter the main insights. Here, we explore a more general effect of these variables, by assuming that they may have a non-linear effect on all structural variables. Models (1)-(4) are estimated in subsamples HH, LH, HL, LL, where the first (second) component represents the first is above the median, H, or below the median, L, in terms of the number of active analysts (size) when a firm enters the sample. Note that our treatment of these characteristics is descriptive as our structural model does not explain why the number of analysts or size might affect disclosure behavior.

The probability of information endowment remains very persistent. Further, the standard deviation of forecast surprise varies slightly between subsamples, with point estimates between .42 (group HL) and .45 (group LH), it remains slightly above but close to the standard deviation of observed forecasts. The lower estimate for firms with more analysts may be due to the fact that managers have a lower informational advantage where the analyst consensus is more precise.

The small difference between .43 and .37 suggests that price incentives only play a moderate role, as more truncations would create a greater wedge between the standard deviation of forecasts and the standard deviation of information. Nevertheless, price

Table 9. Subsample estimates

Columns (3)-(6) are estimated in sub-samples where the first (second) component indicates if the firm is above-median, H, or below-median, L, on Nb. of analysts (log MCAP) when entering the sample. All model details are as in Table 9.

	(1)	(2)	(3)	(4)
Group	HH	LH	HL	LL
k_0	2.736	2.787	2.773	1.997
	(0.271)	(0.403)	(0.381)	(0.166)
k_1	-1.044	-0.990	-0.558	-0.973
	(0.098)	(0.149)	(0.158)	(0.087)
σ_x/σ_e	0.445	0.450	0.424	0.446
	(0.012)	(0.020)	(0.021)	(0.012)
α	3.425	3.363	2.384	2.146
	(0.114)	(0.334)	(0.416)	(0.176)
β	0.370	0.349	0.640	0.756
	(0.082)	(0.126)	(0.139)	(0.084)
σ_e	0.664	0.759	0.762	0.774
	(0.010)	(0.018)	(0.020)	(0.011)
Time trend	-0.714	-0.712	-1.271	-0.422
		(0.278)	(0.333)	(0.152)
Nb. of analysts				
log MCAP				
Log Lik	-13.20	-15.96	-15.03	-15.94
Obs.	2,371	893	753	2,519

incentives α are non-zero, lying between 3.425 (group HH) and 2.146 (group LL) across specifications. Comparing columns (1)-(2) to columns (3)-(4), the estimates also suggest that large firms feature stronger price incentives. To interpret these numbers, note that $\alpha\sigma$ is the total price incentive effect of a one standard deviation increase in the manager's private information. By and large, this term is close to one, so that price incentives have the same impact as one standard deviation change in the random shock ϵ_t. Hence, price incentives are as important to explain disclosure patterns as the noise component of the utility function.

Price incentives may also be compared to the point estimate for β, which varies between .37 (group HH) to .76 (group LL), with point estimates for the full sample lying mid-way. The effect of this constant remains sizeable, being about two-thirds of price incentives for one standard deviation in the private information for small firms, and a little over a third of a standard deviation for large firms.

6.2 Pure reporting motives

We examine next the extent to which disclosure theory can explain disclosure behavior in the absence of a preference intercept, that is, setting $\beta = 0$. In this special case, any systematic preference to disclose must be explained by reporting motives.[15]

Table 10 reveals that, in the absence of an intercept the model predicts that price incentives play a much larger role. Intuitively, while the intercept worked to increase the predicted probability of disclosure, this is now achieved via the price mechanism. The manager must care more about price to sustain the levels of disclosure observed in the sample. Specifically, this model implies price incentives that are about a third higher than in the baseline model.

Private information σ is also a channel via which price incentives matter more, as the

[15] We maintain the assumption of the noise ϵ in the utility function as in discrete choice theory. Dropping ϵ is problematic because it implies that the likelihood function will be discrete (in any finite sample) when an observation crosses the disclosure threshold, so that the information matrix will be ill-defined.

Table 10. Without preference intercept

All variables are defined as in table 9, with the restriction to the intercept in the preference set to $\beta = 0$.

	(1)	(2)
k_0	2.278	3.968
	(0.155)	(0.300)
k_1	-1.480	0.048
	(0.028)	(0.154)
σ_x/σ_e	0.469	0.470
	(0.007)	(0.007)
α	3.487	3.501
	(0.082)	(0.081)
σ_e	0.732	0.732
	(0.006)	(0.006)
Time trend		-0.797
		(0.107)
Nb. of analysts		0.226
		(0.043)
log MCAP		-0.192
		(0.024)
Log Lik	-15.03	-14.94
Obs.	6,857	6,857

manager is now slightly more informed than in the baseline. However, this extra channel is not large, because σ_x and σ_e are well estimated from the observed forecasts. We also find that the information endowment process must be more persistent. The other insights from the baseline model remain similar, even in terms of magnitudes. One important implication of the simplified model with $\beta = 0$ is that, if we do not assume the existence of a systematic preference to disclose, the proportion of strategic withholding is much greater.

6.3 Extended sample

In this section, we re-estimate the model with the entire sample of annual forecasts including earnings announcements before the first forecast, skipping step 16 in table 1. Unfortunately, we do not know investors' belief about the manager's information endowment before the first forecast is issued so, ideally, one would simulate the distribution of this belief conditional on the sample time-series of earnings and forecasts, and integrate the likelihood over every possible belief. However, doing so is not computationally feasible given that it would require integrating many paths separately for each firm and would slow the estimation by a factor equal to the number of paths. Instead, we initialize investors' belief by assuming the manager was uninformed in the period right before the first observation in the sample. We choose this starting point (over, say, the steady-state probability of being uninformed) because the misspecification of the starting belief is more severe if the firm enters the sample with an extended non-disclosure span, which tends to occur if the firm was initially uninformed.

Columns (1) and (2) in table 8 present the results of the estimation. Most of the estimated structural parameters are similar to those in the baseline model. The preference intercept β is estimated to be greater, at 1.18 in column (1) and 1.06 in column (2). The reason for this is that many firms in the sample initiate their first forecast with moderate news which, if the market belief if that they are very likely to be uninformed, should be unlikely if β were close to zero. The higher estimate for β thus indicate that there may be additional reasons, outside of our model, that may explain the initiation of a forecast. The other parameters, including the estimated precision of the information of manager σ_x and the price motives α remain within similar ranges as those estimated in the baseline model.

7 Conclusion

We develop a simple empirical methodology to structurally estimate and test a model of voluntary disclosure with uncertain information endowment. In the model, market beliefs about the friction are updated over time as a function of a firm's past history of withholding and their realized earnings which, in turn, implies that firms with a higher probability of the friction feature a lower negative market reaction to withholding and a greater probability of strategic withholding in future periods.

Several aspects of our approach are preliminary steps toward a more complete under-

Table 11. Other earnings forecast samples

This table reports results from the baseline model in section 2, with extended samples defined as in table 7. Standard errors are in parenthesis are calculated using the information matrix with each observation being a unique firm.

	(1)	(2)
k_0	1.363	2.031
	(0.017)	(0.078)
k_1	-1.144	-0.998
	(0.026)	(0.085)
σ_x/σ_e	0.434	0.436
	(0.005)	(0.005)
α	2.274	2.206
	(0.097)	(0.090)
β	1.181	1.066
	(0.051)	(0.047))
σ_e	0.742	0.742
	(0.004)	(0.004)
Time trend		1.107
		(0.050)
Nb. of analysts		0.103
		(0.022)
log MCAP		-0.166
		(0.013)
Log Lik	-17.66	-17.43
Obs.	19,060	19,060

standing of the voluntary disclosure decision. In our theoretical model, managers only value the posterior expectation about earnings induced by their disclosure, so that types of disclosures such as ranges are mapped to a posterior expectation. In practice, however, some supplementary information appears as if it is conveyed via the choice of a range forecast - possibly about other moments such as the risk of the forecast. While recent work suggests frameworks in which managers do not only maximize price (Hummel, Morgan, and Stocken 2015), whether we can identify alignment parameters with observational data remains an open question.

Another rich potential avenue for the estimation of disclosure models is the strategic timing of information if the same information can be reported over multiple periods or following other market wide or firm disclosures (Dye and Sridhar 1995, Acharya, Demarzo, and Kremer 2011, Guttman, Kremer, and Skrzypacz 2014). In particular, our estimates show that forecasts with different frequency appear to have very different properties but, to our knowledge, there is no existing model to combine low-frequency and high-frequency forecasts. We hope that future research could develop these approaches into fully-dynamic empirical models of disclosure.

Appendix

Appendix A-1: Omitted proofs

Proof of Lemma 2. The price P_t^{nd} is equal to $E_t(x_t|d_t = 0)$ from Bayes rule. Since managers with $\rho_t = 1$ or $\rho_t = 0$ and $\beta + \alpha(P_t(x_t) - P_t^{nd}) + \varepsilon_t < 0$ withhold, the conditional expectation $E_t(x_t|d_t = 0)$ can be expanded to

$$
\begin{aligned}
P_t^{nd} &= \frac{(1-p_t)\int \frac{x_t}{\sigma_x}\phi(\frac{x_t}{\sigma_x})\Pr(\alpha x_t + \epsilon_t < \alpha P_t^{nd} - \beta|x_t)dx_t}{p_t + (1-p_t)\Pr(\alpha x_t + \epsilon_t < \alpha P_t^{nd} - \beta)} \\
&= \frac{(1-p_t)\int \frac{x_t}{\sigma_x}\phi(\frac{x_t}{\sigma_x})\Pr(\epsilon_t < \alpha(P_t^{nd} - x_t) - \beta|x_t)dx_t}{p_t + (1-p_t)\int \frac{1}{\sigma_x}\phi(\frac{x_t}{\sigma_x})\Pr(\epsilon_t < \alpha(P_t^{nd} - x_t) - \beta|x_t)dx_t} \\
p_t P_t^{nd} &= (1-p_t)\int \frac{1}{\sigma_x}\phi(\frac{x_t}{\sigma_x})\Phi(\alpha(P_t^{nd} - x_t) - \beta)(x_t - P_t^{nd})dx_t,
\end{aligned}
$$

so that $\Gamma(P_t^{nd}) = P_t^{nd}$ after a change of variable $x' = x_t - P_t^{nd}$.

Since $|x\Phi(-\alpha x - \beta)\frac{1}{\sigma_x}\phi(\frac{x+y}{\sigma_x})$ is integrable,

$$
\lim_{y\to-\infty}\Gamma(y) = \frac{1-p}{p}\int \lim_{y\to-\infty} x\Phi(-\alpha x - \beta)\frac{1}{\sigma_x}\phi(\frac{x+y}{\sigma_x})dx = 0,
$$

implying that $\Gamma(y) > y$ for y sufficiently small.

Further,

$$
\begin{aligned}
\Gamma(y)-y &= \frac{1-p}{p}\int x\Phi(-\alpha x-\beta)\frac{1}{\sigma}\phi(\frac{x+y}{\sigma})dx-y \\
&< \frac{1-p}{p}\int x\Phi(-\beta)\frac{1}{\sigma}\phi(\frac{x+y}{\sigma})dx-y \\
&= -\left[\frac{1-p}{p}\Phi(-\beta)+1\right]y
\end{aligned}
$$

and, therefore, $\Gamma(y)-y<0$ for any $y\geq 0$ implying that $\Gamma(y)-y$ must have at least one negative root.

We prove uniqueness by contradiction. Define

$$
\begin{aligned}
Q_0(y) &\equiv \int_0^{\infty}(1-p)\,z\frac{1}{\sigma}\Phi(-z\alpha-\beta)\,\phi\left(\frac{z+y}{\sigma}\right)dz && (19)\\
Q_1(y) &\equiv \int_{-\infty}^{0}(1-p)(-z)\,\frac{1}{\sigma}\Phi(-z\alpha-\beta)\,\phi\left(\frac{z+y}{\sigma}\right)dz. && (20)
\end{aligned}
$$

The equilibrium condition can be rewritten as $Q_0(y)-Q_1(y)=py$.

In what follows, we use the monotone-likelihood-ratio property of the standard normal distribution: specifically, it is well-know that if $y_2>y_1$, then for all $x\geq 0$,

$$
\frac{\phi\left(\frac{y_2+x}{\sigma_x}\right)}{\phi\left(\frac{y_1+x}{\sigma_x}\right)}\leq\frac{\phi\left(\frac{y_2}{\sigma_x}\right)}{\phi\left(\frac{y_1}{\sigma_x}\right)}
$$

and the opposite inequality holds for $x<0$.

Assume there are two negative numbers $y_1<y_2$ that satisfy that satisfy the equilibrium condition.

$$
\begin{aligned}
Q_0(y_2) &= \int_0^{\infty}z\Phi(-\alpha z-\beta)\,\frac{1}{\sigma_x}\frac{\phi\left(\frac{z+y_2}{\sigma_x}\right)}{\phi\left(\frac{z+y_1}{\sigma_x}\right)}\phi\left(\frac{z+y_1}{\sigma_x}\right)dz \\
&\leq \frac{\phi\left(\frac{y_2}{\sigma_x}\right)}{\phi\left(\frac{y_1}{\sigma_x}\right)}\int_0^{\infty}z\Phi(-\alpha z-\beta)\,\frac{1}{\sigma_x}\phi\left(\frac{z+y_1}{\sigma_x}\right)dz=\frac{\phi\left(\frac{y_2}{\sigma_x}\right)}{\phi\left(\frac{y_1}{\sigma_x}\right)}Q_0(y_1) && (21)
\end{aligned}
$$

And, similarly,

$$\begin{aligned} Q_1\left(y_2\right) &= \sigma \int_{-\infty}^{0} \left(-z\right) \Phi\left(-z\alpha - \beta\right) \frac{\phi\left(\frac{z+y_2}{\sigma_x}\right)}{\phi\left(\frac{z+y_2}{\sigma_x}\right)} \frac{1}{\sigma_x} \phi\left(\frac{z+y_1}{\sigma_x}\right) dx \\ &\geq \frac{\phi\left(\frac{y_2}{\sigma_x}\right)}{\phi\left(\frac{y_1}{\sigma_x}\right)} \sigma_x \int_{-\infty}^{0} \left(-z\right) \Phi\left(-z\alpha - \beta\right) \frac{1}{\sigma_x} \phi\left(\frac{z+y_1}{\sigma_x}\right) dx = \frac{\phi\left(\frac{y_2}{\sigma_x}\right)}{\phi\left(\frac{y_1}{\sigma_x}\right)} Q_1\left(y_1\right) \end{aligned} \tag{22}$$

In addition, $0 > y_2 > y_1$ implies that $\frac{\phi\left(\frac{y_2}{\sigma_x}\right)}{\phi\left(\frac{y_1}{\sigma_x}\right)} > 1$, so that

$$py_2 > py_1 > \frac{\phi\left(\frac{y_2}{\sigma_x}\right)}{\phi\left(\frac{y_1}{\sigma_x}\right)} py_1. \tag{23}$$

It then follows from (21)-(23) that

$$Q_0\left(y_2\right) - Q_1\left(y_2\right) - py_2 < \frac{\phi\left(\frac{y_2}{\sigma_x}\right)}{\phi\left(\frac{y_1}{\sigma_x}\right)} Q_0(y_1) - \frac{\phi\left(\frac{y_2}{\sigma_x}\right)}{\phi\left(\frac{y_1}{\sigma_x}\right)} Q_1(y_1) - \frac{\phi\left(\frac{y_2}{\sigma_x}\right)}{\phi\left(\frac{y_1}{\sigma_x}\right)} py_1 = 0,$$

where the last equality follows from the fact that y_1 satisfies the equilibrium condition. This contradicts $Q_0\left(y_2\right) - Q_1\left(y_2\right) - py_2 = 0$. ■

Appendix A-2: Numerical methods

The model estimated is estimated with Matlab 2016a, using default options for each optimization routines. To compute the log likelihood function, the theoretical probability of non-disclosure must be computed for each observation from equation (14). However, this would require to solve a different fixed point for each observation and is computationally impractical. Instead, for each set of parameter values, we approximate the non-disclosure price as a function of the current belief p_t by evaluating the non-disclosure price over a 15-point grid $\{((i+1)/16)^2\}_{i=1}^{15}$ and using cubic spline interpolation (command *interp1*) to recover prices outside of the grid.

The approximation is nearly indistinguishable from the non-disclosure price solved from the fixed point (Figure 6) and we observe a similar near-perfect match between the two curves for many other parameter values.

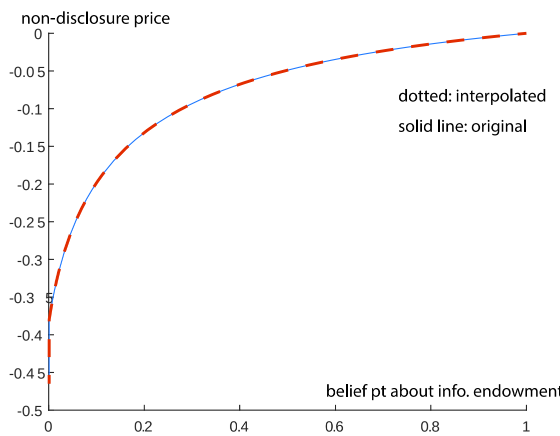


Figure 6. Spline interpolation of (14), $\alpha = \beta = 1$, $\sigma = .5$

To reduce numerical error, we conduct the estimation with standardized covariates, pointwise subtracting the vector $\min(X_t)$ and dividing by the vector $\max(X_t) - \min(X_t)$. We reverse this transformation before reporting estimates and standard-errors. The log-likelihood is maximized using a global search algorithm (command *particleswarm*) followed by a local search algorithm (command *fminsearch*). The information matrix is computed numerically (command *hessian*).

For counter-factual analyses and model moments, we simulate a new sample using the empirical distribution of earnings and forecasts. We construct a simulated sample with 10 times the size of the original sample and use the empirical earnings and management information, when available. For observations in which a forecast is not available in the sample, we simulate a x_t by drawing from $x_t|d_t = 0, e_t$. We then simulate a process for information endowment ρ_t and preference noise ϵ_t, and construct a new sample by assuming that x_t is disclosed as a forecast if and only $\rho_t = 0$ (no friction) and $u_t(d_t = 1|x_t) \geq u_t(d_t = 0|x_t)$.

Appendix A-3: Measuring market expectations

All models are estimated with a median regression to limit the effect of outliers on the prediction model. First, in column (1) of Table 12, we estimate

$$EPS_t = a_0 + a_1 EPS_{t-1} + a_2 \text{ Consensus}_t + v_t, \tag{24}$$

where EPS_t is realized EPS, EPS_{t-1} is the realized prior-period EPS and IBES consensus is the consensus calculated over the window in Figure 1, that is, before the period where forecasts are recorded. The residual of this regression is a measure of the EPS surprise relative to the expectation before a forecast is issued and, assuming that investors use this statistical model to set their initial expectations, corresponds to an observation of e_t. Note that prior EPS and consensus explain a large portion of the cross-sectional variation in EPS, with an adjusted R^2 of 66%.

Then, we run the model including the management forecasts. Since ranges and point forecasts may contain different information, we run the regressions separately for range and interval forecasts, in columns (2) and (3), respectively.

$$EPS_t = a_0 + a_1 \ EPS_{t-1} + a_2 \text{ Consensus}_t + a_3 MF_t + a_4 MF_t^{upper} + v_t, \tag{25}$$

where MF_t is the minimum of the range for range forecasts, or the guidance for open-ended or point forecasts, MF_{upper} is only included in column (2) and is defined as the maximum of the range. For point and open-ended forecasts, the MF appears to be an unbiased estimator of the realized EPS, with a coefficient equal to one. For range forecasts, the estimate of future EPS is about two-thirds of the lower-end and one-third of the upper-end. Management forecasts contain incremental information, increasing the adjusted R^2 from 66% to 71-72%. The forecast surprise x_t is recovered as the predicted value from (25), i.e., the revised expectation about future earnings conditional on the forecast, minus the predicted value from (24), i.e., the expectation prior to the forecast.

Table 12. Market Expectations

This table reports results from an OLS regression of EPS on lag EPS, latest I/B/E/S consensus, MF forecast and (if an interval forecast) upper bound of the interval MF_Upper. Model (1) is estimated for all firms in the sample. Model (2) is estimated for all firms with interval MF and model (3) is estimated for all firms with point or open-ended upper MF. Residuals from model (1) are then used to construct the EPS surprise. We compute the MF surprise by subtracting the predicted EPS in models (2) and (3) to the predicted EPS in model (1). Standard errors are in parenthesis with significance at the 1%, 5% and 10% level indicated as *, ** and ***, respectively.

	(1) Pre MF	(2) Range MF	(3) Point or Open-ended MF
Lag EPS	0.19***	0.09***	0.09***
	(0.007)	(0.01)	(0.03)
IBES cons.	0.80***	-0.09***	-0.06
	(0.007)	(0.03)	(0.06)
MF		0.64***	0.97***
		(0.05)	(0.06)
MF_Upper		0.36***	
		(0.05)	
Constant	0.03***	0.07***	0.06**
	(0.008)	(0.01)	(0.03)
Adj. R2	0.66	0.72	0.71
Obs.	8,013	3,293	510

We standardize the variables to a population unit variance in realized EPS, as in (13), by dividing EPS and MF surprises by the standard error of EPS surprises in the entire sample. We also track other firm characteristics that may be associated with disclosure. Assets, liabilities as well as liabilities, $R\&D$, capital expenditures are from Compustat, and we scale the last three by current assets and measure them in the fiscal quarter of lag EPS. We approximate market capitalization by multiplying the number of shares out and the share price one day before the lag EPS earnings announcement, from CRSP. We define a variable nb_analysts as the number of unique (active) analysts issuing a new EPS forecast during the EPS consensus window, from IBES. Lastly, we remove potential outliers by winsorizing EPS surprises, forecast errors (differences between EPS and MF surprises) and characteristics at the 1% level. We then trim the three greatest and three lowest forecast surprises.

References

Acharya, V. V., P. Demarzo, and I. Kremer (2011): "Endogenous Information Flows and the Clustering of Announcements," The American Economic Review, 101(7), 2955–2979.

Baginski, S. P., J. M. Hassell, and M. D. Kimbrough (2002): "The Effect of Legal Environment on Voluntary Disclosure: Evidence from Management Earnings Forecasts Issued in U.S. and Canadian Markets," The Accounting Review, 77(1), 25–50.

Balakrishnan, K., M. B. Billings, B. Kelly, and A. Ljungqvist (2014): "Shaping liquidity: On the causal effects of voluntary disclosure," the Journal of Finance, 69(5), 2237–2278.

Bertomeu, J., A. Beyer, and R. A. Dye (2011): "Capital Structure, Cost of Capital, and Voluntary Disclosures," The Accounting Review, 86(3), 857–886.

Bertomeu, J., A. Beyer, and D. J. Taylor (2015): "From Casual to Causal Inference in Accounting Research: The Need for Theoretical Foundations," Available at SSRN 2694105.

Bertomeu, J., I. Marinovic, S. J. Terry, and F. Varas (2017): "The Dynamics of Concealment," Discussion paper, Working Paper.

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

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

Breuer, M., and D. Windisch (2018): "Investment Dynamics and Earnings-Return Properties: A Structural Approach," .

Caskey, J., G.-L. Gayle, and C. Li (2017): "The role of accounting in a dynamic model of CEO pay and turnover," .

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

Cheong, F. S., and J. Thomas (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., and M. Liu-Watts (2015): "A back-of-the-envelope estimator of disclosure costs," Discussion paper, Columbia Business School Working Paper.

——— (2016): "When Dye Meets Verrecchia: A Structural Estimation of Disclosure Theory," .

Dye, R. A. (1985): "Disclosure of Nonproprietary Information," Journal of Accounting Research, 23(1), 123–145.

Dye, R. A., and J. S. Hughes (2017): "Equilibrium voluntary disclosures, asset pricing, and information transfers," Journal of Accounting and Economics.

Dye, R. A., and S. Sridhar (2004): "Reliability-Relevance Trade-Offs and the Efficiency of Aggregation," Journal of Accounting Research, 42(1), 51–88.

Dye, R. A., and S. S. Sridhar (1995): "Industry-wide disclosure dynamics," Journal of accounting research, 33(1), 157–174.

Einhorn, E., and A. Ziv (2008): "Intertemporal Dynamics of Corporate Voluntary Disclosures," Journal of Accounting Research, 46(3), 567–589.

Fang, V. W., A. H. Huang, and W. Wang (2017): "Imperfect accounting and reporting bias," Journal of Accounting Research, 55(4), 919–962.

Fischer, P. E., and R. E. Verrecchia (2000): "Reporting bias," The Accounting Review, 75(2), 229–245.

Gayle, G.-L., C. Li, and R. A. Miller (2015a): "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.

——— (2015b): "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 A. Kovrijnykh (2013): "Performance shocks and misreporting," Journal of Accounting and Economics, 56(1), 57–72.

Grossman, S. J. (1981): "The Informational Role of Warranties and Private Disclosure about Product Quality," Journal of Law and Economics, 24(3), 461–483.

Guttman, I., I. Kremer, and A. Skrzypacz (2014): "Not only what but also when: A theory of dynamic voluntary disclosure," The American Economic Review, 104(8), 2400–2420.

Hemmer, T., and E. Labro (2015): "Management by the Numbers: A Formal Approach to Deriving Informational and Distributional Properties of ŞUn-ManagedŤ Earnings," Discussion paper, Working paper, Rice University and The University of North Carolina at Chapel Hill.

Hummel, P., J. Morgan, and P. C. Stocken (2015): "Persuasion Under Higher-Order Uncertainty," Available at SSRN 2656144.

Jung, W.-O., and Y. K. Kwon (1988): "Disclosure When the Market Is Unsure of Information Endowment of Managers," Journal of Accounting Research, 26(1), 146–153.

Kartik, N., M. Ottaviani, and F. Squintani (2007): "Credulity, lies, and costly talk," Journal of Economic theory, 134(1), 93–116.

Kasznik, R., and B. Lev (1995): "To Warn or Not to Warn: Management Disclosures in the Face of an Earnings Surprise," The Accounting Review, 70(1), 113–134.

King, R. R., and D. E. Wallin (1991): "Voluntary disclosures when seller's level of information is unknown," Journal of Accounting Research, pp. 96–108.

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

Li, C. (2013): "Essays on the Structural Models of Executive Compensation," Ph.D. thesis, Carnegie Mellon University.

——— (2018): "Are top management teams compensated as teams? A Structural Modeling Approach," .

Liang, P., B. Sun, and X. S. Tam (2016): "The Real Effect of Financial Reporting: A Quantitative Assessment," .

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

McFadden, D. (1973): "Conditional logit analysis of qualitative choice behavior," .

——— (1980): "Econometric models for probabilistic choice among products," Journal of Business, pp. S13–S29.

Penman, S. H. (1980): "An Empirical Investigation of the Voluntary Disclosure of Corporate Earnings Forecasts," Journal of Accounting Research, 18(1), 132–160.

Skinner, D. J. (1997): "Earnings disclosures and stockholder lawsuits," Journal of Accounting and Economics, 23(3), 249–282.

Verrecchia, R. E. (1983): "Discretionary Disclosure," Journal of Accounting and Economics, 5, 179–194.

Viscusi, W. K. (1978): "A note on" lemons" markets with quality certification," The Bell Journal of Economics, pp. 277–279.

Zhou, F. (2016): "Disclosure Dynamics and Investor Learning," .