Uncertainty, Disagreement and Forecast Dispersion: Empirical Estimates from a Model of Analysts' Strategic Conduct

Youfei Xiao*

January 2016

JOB MARKET PAPER
Current Version at `https://stanford.box.com/Xiao-JMP`

Abstract

This paper provides empirical estimates of uncertainty and disagreement about future earnings underlying analyst forecast dispersion from a theoretical model where analysts are motivated by strategic incentives. A parsimonious model which assumes that analysts' payoffs are jointly determined by forecast error and deviation from consensus reproduces many of the descriptive facts observed about forecast dispersion in the data. The strategic behavior that arises from the model distorts both the levels of forecast dispersion and the sensitivity of the measure with respect to cross-sectional variation in uncertainty. The estimated parameters perform better at predicting forecast dispersion out-of-sample than approaches based solely on associations with firm characteristics. Counterfactual simulations indicate that analysts' strategic incentive together with the sequential forecast setting, plays a first-order role in determining forecast dispersion relative to the firm's information environment. The model-implied estimates of earnings uncertainty exhibit a substantially less negative association with future returns relative to the association generated by forecast dispersion and thus partially reconciles previous evidence with theories about the asset pricing implications of uncertainty and disagreement.

*Doctoral Candidate in Accounting, Graduate School of Business, Stanford University. E-mail: youfei.xiao@stanford.edu. I am enormously grateful to Dave Larcker and Peter Reiss for extensive guidance and support. I thank Mary Barth, Bill Beaver, Anna Boisits, Ed deHaan, Lindsey Gallo, Will Gornall, Doron Israeli, Steve Karolyi, Mike Klausner, Charles Lee, Becky Lester, Iván Marinovic, Maureen McNichols, Maria Ogneva, Marlene Plumlee and Durgesh Saraph as well as workshop participants at the AAA Doctoral Consortium and Stanford University for helpful comments and discussions. Financial support from the Chartered Accountants Education Foundation and the Social Sciences and Humanities Council of Canada is gratefully acknowledged. All errors and omissions are my own.

1. INTRODUCTION

Analyst forecast dispersion is used extensively as an empirical proxy in accounting and finance. Underlying the measure's broad appeal is its putative relation to uncertainty and disagreement about future earnings - two economic constructs which are otherwise challenging to operationalize on a timely basis. However, a number of studies focusing on forecast dispersion qualify their conclusions with the insight that observed forecast dispersion is driven by both of these constructs as well as analysts' strategic incentives and that the yet to be analyzed role of these incentives could have implications for interpreting the prevailing empirical evidence. Motivated by such an observation, this study quantifies the earnings uncertainty and disagreement implied by analyst forecast dispersion in a setting where analysts' choices of forecasts are influenced by strategic considerations.

To further emphasize the necessity of analyzing the role of incentives in empirical tests based on forecast dispersion, consider the well-known negative association between forecast dispersion and equity returns. Because the direction of this association is inconsistent with that predicted by traditional models in asset pricing theory, follow-up work that further scrutinize this apparent anomaly has proliferated in recent years. In this stream of research, tests of whether uncertainty or disagreement is reflected in equity returns are actually joint tests of pricing and the appropriateness of using forecast dispersion as a measure in the presence of strategic conduct. Thus, conclusions from this literature require establishing the construct validity of forecast dispersion.

The approach that I use to separate earnings uncertainty and disagreement from analysts' strategic actions and to explore the associated implications is broadly divided into four parts. I begin by documenting stylized empirical regularities about analyst forecast dispersion. Guided by these stylized results, I then develop the theoretical model in the second part of the study. The theory discussion is divided into three intermediate steps where I first relate the model parameters capturing the information environment to analysts' earnings predictions, then link analysts' earnings predictions together with their strategic incentives to their choice of forecasts and finally map the information environment and strategic incentives to forecast dispersion. The third part describes the statistical assumptions and procedures used to transform the theoretical model into estimable form that can be taken to the data. The fourth and final part discusses the results, which are comprised of three components: the empirical estimates of uncertainty and strategic incentives, the assessment of out-of-sample model fit and a re-examination of a well-known empirical result based on forecast dispersion using my parameter estimates.

In the first part of my study, I highlight two descriptive results about analyst forecast dispersion not generally known in this literature. These results are based on a sample comprised of 151,728 two-year ahead earnings per share forecasts issued in the 30 day period after firms' annual earnings announcements. Analyst forecast dispersion appears to increase after each analyst announces his forecast. This increase exists even after accounting for the mechanical impact of analyst coverage. The descriptive evidence also indicates that forecast dispersion is highly correlated with the magnitude of the lead analyst's forecast error. It appears that a formal theoretical model could help inform the interpretation of these facts.

Having identified some stylized facts in the data which I hope to reproduce with theory, the second part of my paper describes a sequential forecasting game after an earnings announcement event. The public earnings news, observed by all analysts, can be aggregated together into an earnings prediction with a certain amount of variance. This variance can be thought of as the earnings uncertainty parameter. Each analyst's proprietary forecasting process produces a noisy but statistically unbiased signal that only he observes. In addition, analysts also observe peer forecasts that have already been announced. In the first intermediate step of my theoretical analysis, I show that each analyst's prediction about earnings is a weighted average of his inferences about the preceding analyst's earnings prediction and his private signal, where the weighting is determined by earnings uncertainty and the variance of the private signal. Thus, both earnings uncertainty and the noise in analysts' private information give rise to disagreement amongst analysts about future earnings. Clearly, if analysts' payoffs are determined solely by their forecast error, then they would forecast their earnings prediction and the resulting dispersion over forecasts can be interpreted as disagreement about future earnings.

My model assumes instead that analysts' payoffs are determined by a weighted sum of the magnitude of his forecast error and the magnitude of his difference from consensus forecast. The second intermediate step of my theoretical analysis characterizes the equilibrium resulting from analysts optimizing over their payoffs. In this equilibrium, the analyst chooses to forecast closer to (farther from) consensus, relative to his earnings prediction, to the extent that his objective function penalizes (rewards) him for deviations from consensus.

In the final step of my theoretical analysis, I numerically compute the model prediction about the forecast dispersion and forecast error for various numerical assumptions about the model parameters. This prediction depends on a stochastic term not observed by the researcher. Thus I also compute the expected forecast dispersion and forecast error by taking an expectation, using Monte Carlo integration methods, over the assumed distribution of this stochastic term. Through some experimentation, I find that under one specific set of numerical assumptions, the equilibrium from the model produces predictions about expected forecast error and dispersion quantitatively consistent with the increasing pattern of forecast dispersion as well as the correlation between the magnitude of the forecast error and dispersion documented in the descriptive analysis from the first part of the paper. My simulations indicate that analysts' strategic behavior not only distorts the level of dispersion but also the extent to which cross-sectional differences in uncertainty are incorporated into dispersion.

The statistical procedure I use to estimate the theoretical model, outlined in the third part of my paper, formalizes the experimentation- simulation process from the last step in my theoretical analysis. I use a Simulated Method of Moments (SMM) approach which searches for the set of model parameters that produce predictions about the expected forecast dispersion as each analyst forecasts and about forecast error which replicate their empirical counterpart. In my theoretical analysis, I showed that changes in parameter assumptions about the uncertainty in public earnings news and about the variance in the analyst's private information jointly influence the model predictions about the expected magnitude of the forecast error and forecast dispersion while changes in the parameter assumptions about the relative importance of deviations from consensus in analysts' payoffs determine the model predictions about the expected levels of forecast dispersion as well as the expected change in dispersion as each analyst forecasts. Accordingly, the progression of forecast dispersion as each analyst forecasts observed in the data gives rise to empirical identification of the parameter capturing analysts' strategic incentives while the level of forecast dispersion, relative to forecast error, provides identification of the earnings uncertainty parameter.

The parameter estimates, presented in the final part of the paper, perform well at explaining the data. In particular, they exhibit superior ability at predicting dispersion out of sample than approaches based on the association with firm characteristics. I consider the impact of counterfactual policies such as reversing the direction of analysts' peer incentives on observed forecast dispersion as well as the elimination of peer incentives altogether. I also analyze a counterfactual policy which prohibits analysts from using already-announced forecasts in forming their earnings predictions. I find that forecast dispersion is far more sensitive to these two types of policies compared to counterfactual policies which alter drastically the quality of the information environment. Because my results indicate that analysts' incentives and behavior play first order roles in determining observed analyst forecast dispersion, regulatory reforms along these two dimensions would have a greater economic impact than regulation that, for example, increases the level of management-analyst disclosures.

Using my model-implied estimates of earnings uncertainty, I revisit one popular empirical application from previous research involving the use of analyst forecast dispersion. My tests confirm the well-known result that analyst forecast dispersion is negatively associated with future returns. When I substitute my model-implied estimate of earnings uncertainty in place of forecast dispersion in the same test, this association drops by more than half in certain specifications and to zero in others. Such a reduction in the association with returns suggests that the dispersion-return relation is, at a minimum, less anomalous than previously thought. Further, certain asset pricing models may be able to explain the sign of the remaining association absent any assumptions about market inefficiencies. Conclusive determination on whether the pricing of uncertainty and disagreement is consistent with the efficient market hypothesis is beyond the scope of this paper. Rather, the intent is to raise recognition that disentangling strategic conduct from uncertainty and disagreement provides insights into the asset pricing implications of the latter two constructs not possible by studying only the dispersion-return relation.

The remainder of the paper is organized into the following sections. Section 2 summarizes the relevant literature. Section 3 provides an overview of regularities about forecast dispersion. Section 4 outlines the theoretical model and discusses the resulting equilibrium. Section 5 describes the estimation approach along with additional assumptions I use to recover the model parameters. Section 6 reports the estimation results and Section 7 concludes.

2. Related Literature

2.1 The Association Between Dispersion and Equity Market Outcomes

Analyst forecast dispersion is studied in a variety of empirical settings in accounting and finance. Perhaps the most prolific result in this literature is that forecast dispersion is negatively associated with the cross-section of returns (Gebhardt et al., 2001; Diether et al., 2002). A trading strategy based on a long portfolio of firms in the lowest quintile of dispersion and a short portfolio of firms in the highest quintile generates an annual return of 9.48% (Diether et al., 2002). In follow-up work, Johnson (2004) notes "this finding is important in that it directly links asset returns with a quantitative of an economic primitive - information about fundamentals-but the sign of the relationship is apparently wrong."[1] Diether et al.'s (2002) explanation for the result is that the cross-sectional variation in forecast dispersion is driven by differences in opinion between analysts. The paper notes that the difference in opinion between analysts is a proxy for the difference in opinion between investors. In a market microstructure model in which investors have heterogeneous expectations, disagreement about stock valuation lowers future returns (Miller, 1977). In contrast, Johnson (2004) assumes that forecast dispersion is driven by information risk about earnings and proposes an alternate theoretical mechanism in which expected equity returns for a levered firm decreases with such type of idiosyncratic risk. Other papers with evidence on the association between

[1] The same paper notes later on that forecasts dispersion may not be an economic primitive due to the potential for analysts' strategic behavior.

dispersion and returns include Avramov et al. (2009), Berkman et al. (2009),Zhang (2006), Anderson et al. (2009) and Sadka and Scherbina (2007).

More broadly, dispersion is known to be associated with an array of other market outcomes. For example, Ajinkya et al. (1991) documents a significant positive association between forecast dispersion and trading volume. Botosan et al. (2004) derives measures of public and private information precision using forecast dispersion and finds that the precision of public information (private information) is negatively (positively) related to the cost of equity capital. Evidence in Garfinkel and Sokobin (2006) points to a negative relation between forecast dispersion and the post-earnings announcement drift.

2.2 Statistics-Based Measures of Uncertainty and Disagreement

A number of previous studies use the measures of uncertainty and disagreement developed in Barron et al. (1998), hereafter "BKLS". My model is based on an information structure similar to that in BKLS. However, BKLS notes that the validity of its measures is dependent on some stylized assumptions. It assumes that all analysts have identical precision while my model allows for some heterogeneity. It also assumes that analysts forecast independently without observing what other analysts have announced. As such, it rules out the possibility that earlier analysts' forecasts inform later analysts about their private information and more importantly, it does not consider analysts' strategic interaction. My estimates are based on an underlying analytical model that explicitly incorporates these two considerations.

Relatedly, Liu and Natarajan (2012) develops a model in which observed dispersion depends linearly on firm and analyst-characteristics as well as an unobservable term. It hypothesizes that strategic conduct necessarily implies that this unobservable term is truncated from below. Using the fitted values estimated from this model, it computes a measure of dispersion which apparently removes the impact of strategic conduct. My approach is a departure in the sense that the statistical methods are explicitly linked to an underlying economic model of analysts' forecasting problem and it is the resulting equilibrium of this economic model that gives rise to the estimation routine. In fact, I show later in the paper that the *econometric* model proposed in Liu and Natarajan (2012) is not consistent with an *economic* model of analysts' strategic conduct.

3. Descriptive Evidence

Prior to developing the formal theory, I provide some preliminary facts about forecast dispersion that, ideally, a model should also be able to produce. The data I use is taken from all two-year ahead earnings forecasts in the IBES Unadjusted Details File issued in the 30 day window following annual earnings announcements between 1990 and 2012.[2] Although the 30-day cutoff may appear arbitrary, there are no forecasts in the subsequent 30 days for more than 90% of firm-years in the sample and earnings news is likely to be stale in forecast formation beyond the 60 day horizon. According to untabulated summary statistics, the mean and median coverage in the sample is 6 analysts. I restrict my analysis to the 90% of all firm-years covered by less than 11 analysts. I further drop firms covered by only two analysts due the noisiness of using a sample dispersion taken over two analysts. On average, the first forecast is issued two days after the earnings announcement and the first two (last two) forecasts are issued three (seven) days apart. As such, I consider the data to be a reasonable approximation of an economic setting where each analysts can observe previous analysts' forecasts when issuing their own forecasts. Additionally, the timing of forecasts in the data suggests that analysts' information sets have a common component that arises from the earnings announcement while the differences in private information is largely attributable to their subjective interpretation of that earnings news. Throughout my analysis, I use raw EPS amounts without any scaling.[3] The above procedure results in a sample of $151,728$ forecasts pertaining to $27,327$ firm-years.

To provide initial insights into the cross-sectional variation in forecast dis-

[2]While many previous studies of analyst forecast dispersion use forecasts over the one-year or even shorter horizons, the timing of those forecasts does not provide a clear-cut case for whether a sequential or simultaneous move game is the appropriate model to assume. In addition, a number of the previously documented associations between forecast dispersion and asset prices, such as that in Diether et al. (2002), are robust to the use of two-year forecasts.

[3]Although scaling forecasts by share price or dispersion by the consensus forecast may be appropriate in other research designs, I am interested in the measure that analysts use in their payoff optimization problem. Following evidence in Cheong and Thomas (2011), I assume that this measure is raw EPS itself.

persion, I document how this measure varies with analyst coverage. I compute dispersion as the sample standard deviation over all individual analyst forecasts in the post-earning announcement window discussed above.[4] In Panel A of Table 1, I summarize the average forecast dispersion by coverage. The t-statistic for the difference in means between consecutive groups is also tabulated in the adjacent column. The results in Table 1 confirm findings in Diether et al. (2002) and Liu and Natarajan (2012) that forecast dispersion is positively correlated with coverage size. In fact, this non-parametric counterpart to the regression coefficients from the earlier papers indicates that dispersion increases monotonically with coverage. This increase is statistically different from zero across firms covered by less than 6 analysts. Overall, forecasts for firms covered by 11 analysts are 21% more dispersed than those covered by only 3 analysts.

A second natural descriptive statistic to consider is the within-firm progression of dispersion as each analyst forecasts. For each set of forecasts corresponding to a firm-year, I identify the order of announcements, which I denote as j, using the ANNDATS variable in the IBES Unadjusted Details File. I then compute a running version of forecast dispersion by taking the standard deviation over the first through j-th forecast. I compute this running version of forecast dispersion for $j = 2, 3, \ldots$ up to the total number of analysts for the firm-year. Panel B of Table 1 reports the average running dispersion for each j where the t-statistic in the adjacent column corresponds to the difference in means from the average running dispersion for $j - 1$. The pattern of dispersion increases as each analyst forecasts. For example, the average dispersion is 0.169 after the 11th forecast and 0.129 after the second forecast. The difference is significant at 5% level (unsigned test, t-statistic=6.65).

Although the results indicate that, on average, dispersion increases with more analysts, the concern remains that this trend is driven mechanically by firms with higher analyst coverage. To remove the effects of heterogeneity due to total number

[4]While the sample variance, computed as $s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$, is an unbiased estimator of its population counterpart, the square root of the sample variance (i.e., sample standard deviation) is biased in small samples. There are no results for unbiased estimators of the population standard deviation except under very specific distributional assumptions about the underlying data. This observation does not change the discussion of the results that follow although the t-statistics in Table 1 could be overstated. The implications of the bias in sample standard deviation for estimating my model parameters are addressed in Section 6.

of analysts, I compute a demeaned version of running dispersion, where I subtract from raw dispersion after the j-th analyst by the average for the corresponding coverage size in Panel A. This quantity is then normalized by adding a constant such that for $j = 2$, average demeaned dispersion is the same as average raw dispersion. As the last two columns of Table 1, Panel B shows, the demeaned running dispersion still increases, although more modestly, with j. The average difference in this variable after the 11th and the 2nd forecast is 0.026 (t-statistic=4.31).

My next set of tests is motivated by evidence from prior papers that the magnitude of analysts' forecast error is highly correlated with dispersion (Diether et al. (2002),BKLS). Confirming results from previous research, I find that the Pearson (Spearman) correlation between the absolute value of the first analyst's forecast error, denoted as $|LEADFE|$, and analyst forecast dispersion is 0.43 (0.46).[5] I extend this analysis further by examining the progression of dispersion conditional on $|LEADFE|$. Panel C of Table 1 reports the progression of running dispersion based on a sort of firm-years into terciles using $|LEADFE|$. In contrast to the previous analysis, I discard the first forecast to remove the mechanical impact that the first analyst's forecast would otherwise have on forecast dispersion. I find that forecasts are most dispersed in the highest tercile of $|LEADFE|$, where the running dispersion ranges from 0.173 after the third forecast to 0.273 after the eleventh forecast. In contrast, for firm-years in the lowest tercile of $|LEADFE|$, dispersion ranges from 0.055 to 0.093. Similar to the pattern in Panel B, dispersion increases almost monotonically with every subsequent analyst for all three subsamples. Collectively, the evidence suggests that the same set of underlying economic primitives drive variation in both $|LEADFE|$ and dispersion. The levels of $|LEADFE|$ and the sequence of dispersion documented in Table 1 as well as the correlations between the two form the basis for the discussion of model calibration and estimation that follows.

In the final analysis of this section, I establish a benchmark for the performance of a descriptive approach, based on the association between forecast dispersion and firm characteristics, at explaining the data. Since my previous analysis suggested that the lead forecast error and forecast dispersion have a shared component, I

[5]The precise motivation for using only the first forecast (rather than the forecast error in consensus) will be discussed in the next several sections.

also estimate the association between the same characteristics and $|LEADFE|$. After dropping observations due to the non-availability of data, my sample is comprised of $122,391$ forecasts and $21,787$ firms-years. I randomly select $19,560$ observations, representing approximately 90% of the sample, for estimation. The remaining observations are retained for assessing out-of-sample fit. The exact model I estimate is:

$$\begin{aligned}Y_{it} =& \gamma_0 + \gamma_1 ASSETS_{it} + \gamma_2 BTM_{it} + \gamma_3 PRC_{it} + \gamma_4 COV_{it} + \gamma_5 SIZE_{it} + \gamma_6 RD_{it} \\ &+ \gamma_7 RDMISS_{it} + \gamma_8 SALES_{it} + \gamma_9 |ACCR|_{it} + \gamma_{10} LOSS_{it} + \gamma_{11} BETADEC_{it} \\ &+ \gamma_{12} LEV_{it} + \gamma_{13} PASTVOL_{it} + \upsilon_{it},\end{aligned} \tag{1}$$

where the dependent variable Y_{it} is either the natural logarithm of analyst forecast dispersion, $DISP_{it}$, or the natural logarithm of the magnitude of the lead analyst's forecast error, $|LEADFE|_{it}$. The cross-sectional determinants of $DISP$ and $|LEADFE|$ I use, all measured at the earnings announcement date that marks the start of the measurement window used for sample selection, are largely similar to those considered in prior studies such as Diether et al. (2002) and Liu and Natarajan (2012). These include total assets ($ASSETS$) and accruals ($|ACCR|$), both on a per share basis, share price (PRC), the book-to-market ratio (BTM), analyst coverage (COV), R&D Expense (RD) which is set to zero if missing, an indicator variable if R&D Expense is missing ($RDMISS$), the leverage ratio (LEV), the CRSP Beta decile ($BETADEC$) and the standard deviation of earnings per share over the previous eight quarters $PASTVOL$. I estimate (1) with both calendar-year fixed effects and industry fixed effects determined using the Fama-French 48 Industry Classifications.

Table 2 presents the results from estimating the determinants model in (1). Across the $DISP$ and $LEADFE$ regressions, the sign and, in many cases, the magnitude of the coefficients are similar. I find that firms with more assets as well as those with higher B/M ratios, analyst coverage, R&D activity, accruals and historical volatility tend to have both higher dispersion and lead forecast error. Also both constructs are negatively associated with market capitalization, sales and

Beta. Loss firms exhibit higher dispersion and lead forecast error than non-loss firms. Finally, the sign of the association with price and leverage ratio are not consistent across the two regressions, although in all but one case the coefficients were not significant. Taken together, these common associations with firm risk proxies provide further support for the notion that both dispersion and the unsigned forecast error for the lead analyst measure similar economic constructs.

4. A Model of Analysts' Beliefs & Strategic Interaction

This section outlines the model. I discuss the assumptions about the information that analysts use to form their beliefs about earnings and I provide expressions for the first two analyst's beliefs to illustrate how they vary with underlying parameters followed by numerical examples of dispersion in beliefs for select assumptions about model parameters. Then I describe the analyst's payoff function. Given the assumed payoffs and information structure, I discuss, using the two-analyst example, the strategic behavior arising from optimization of the payoff function. I also compute numerically the model-predicted dispersion for certain assumptions about analysts' incentives to bias. A side-by-side comparison using results from the previous section show that the average dispersion and forecast error documented in the data are almost identical to those computed from the model.

4.1 Analysts' Information and Beliefs

A. Assumptions

The model considers analysts' use of public and private information to forecast earnings A at some horizon. Conditional on only the public information, A is assumed to be normal with mean μ_0 and variance σ^2. That is, μ_0 is a prediction about future earnings which aggregates all firm and industry earnings announcements, management disclosures and other information that all market participants observe and σ^2 captures the precision of the prediction. In colloquial terms, firms with large (small) σ's are difficult (easy) to forecast using contemporaneous earnings news irrespective of analyst forecasts. As such, it is natural to interpret σ as a measure of fundamental uncertainty or information asymmetry about A. Throughout the remainder of the paper, I refer to μ_0 as prior prediction based on public news and σ as the uncertainty parameter.

There are J analysts who sequentially issue forecasts f_j according to an exogenously determined order. All analysts fully observe public information and I assume there is no new information once a forecasting game as started.[6] In addition, each analyst's research activity produces a private signal $s_j = A + \tau_j \cdot \varepsilon_j$ where ε_j is the unit Normal random variable. Equivalently, conditional on actual earnings A, signals s_j are independently normally distributed with mean A and variance τ_j^2. The analyst forecasting in the j-th order observes his own signal s_j, the $j-1$ preceding forecasts $f_1, \ldots, f_{j-1}$ and the inferred signal from preceding analysts' forecast $\hat{s}_1(f_1), \ldots, \hat{s}_{j-1}(f_{j-1})$. Collectively, the observed information gives rise to posterior beliefs about earnings.

B. Expressions for the First Two Analysts' Beliefs

Prior to specifying the analyst's objective function, I provide expressions for the first two analysts' forecasts as well as the resulting forecast dispersion under the assumption that analysts honestly announce their beliefs and that all analysts signals have the same variance (i.e., $\tau_j = \tau, \forall j$). Since the model assumes that A and s_j are jointly normal, the posterior distribution of each analyst's beliefs is linear in the earnings prediction based on public information μ_0 and the private signal s_j. I use μ_j to denote mean of the posterior beliefs for j.[7] For the first analyst, this is:

$$\mu_1 = \frac{\tau^2}{\sigma^2 + \tau^2}\mu_0 + \frac{\sigma^2}{\sigma^2 + \tau^2}s_1. \tag{2}$$

In Equation (2), the weighting is consistent with familiar intuition in the sense that very noisy signals (i.e., high values of τ) induces the analyst to place a higher weight on the prior prediction from public news. Similarly, to the extent that the firm has high fundamental uncertainty (i.e., high values of σ), the first analyst

[6] I acknowledge that this assumption does not hold exactly in the data to the extent that industry peer firms announce earnings once the first analyst has forecasted.

[7] Following prior studies which consider analyst forecast dispersion as a measure of disagreement or divergence of beliefs, it may be natural to think of this construct as the difference between the μ_j's in my model. Section 5 discusses my assumptions about the cross-sectional heterogeneity in disagreement, and the extent to which this concept differs from the cross-sectional heterogeneity in uncertainty.

places a higher weight on his own private signal. The mean of the second analyst's posterior beliefs can be expressed as:

$$\mu_2 = \frac{\tau^2}{2\sigma^2 + \tau^2}\mu_0 + \frac{\sigma^2}{2\sigma^2 + \tau^2}s_2 + \frac{\sigma^2}{2\sigma^2 + \tau^2}\hat{s}_1(f_1). \tag{3}$$

Equation (3) is similar to (2) in that the magnitude of σ and τ jointly determine the weighting placed on analyst 2's private signal relative to the prior prediction based on public news. However, analyst 2's forecast also places some weight on his inferences about 1's signal upon observing f_1. Since I have assumed that both analysts forecast their analysts truthfully, then $f_1^* = \mu_1$ and (2) can be re-arranged as:

$$s_1 = \frac{f_1(\sigma^2 + \tau^2) - \tau^2\mu_0}{\sigma^2}.$$

Since the second analyst's conjecture must be rational given analysts 1's choice of f_1^*, $\hat{s}_1\left(f_1^*(s_1)\right) = s_1$ and (3) can be written as:

$$\mu_2 = \frac{\tau^2 + \sigma^2}{2\sigma^2 + \tau^2}f_1 + \frac{\sigma^2}{2\sigma^2 + \tau^2}s_2. \tag{4}$$

Finally, since I assume that the second analyst also forecasts his beliefs, with $f_2^* = \mu_2$, then the standard deviation after the second forecast is proportional to the absolute value of the difference between the first two analysts' forecasts:

$$|f_2^* - f_1| = \frac{\sigma^2}{2\sigma^2 + \tau^2}|s_2 - f_1|. \tag{5}$$

The quantity in equation (5) is increasing in σ, consistent with the intuition that holding fixed f_1, high fundamental uncertainty induces the second analyst to disagree with the first analyst.[8] The expression generally declines in τ, reflecting the fact that an imprecise private signal is weighted less in μ_2.[9]

[8]Of course, according to Equation (2), the cross-sectional variation in f_1 also arises endogenously with σ and τ. I discuss how this endogeneity enters into my estimation in the next section.

[9]Since $s_2 = A + \tau\varepsilon_2$, there is also an indirect effect which gives rise to more dispersed s_2's, the weighting effect will always dominate.

C. Numerical Examples of Dispersion in Beliefs

In Panel A of Table 3, I provide the expectation of model-predicted progression of dispersion, still maintaining the honest-forecasting assumption, for certain values of σ and τ. To be consistent with the approach used to describe the data in Section 3, I use only forecasts issued for the second, third and up to the eleventh forecast so that f_1 does not mechanically enter into dispersion. The values provided are an expectation because private signals s_j are unobservable to the researcher. Rather, I use the assumption that $s_j = A + \tau\varepsilon_j$ to compute expected forecast dispersion one should expect to observe after integrating over many draws of ε_j. For example, taking an expectation over Equation (5) gives:

$$\mathbb{E}\left(|f_2^* - f_1|\right) = \frac{\sigma^2}{2\sigma^2+\tau^2}\left[\tau\sqrt{\frac{2}{\pi}}e^{-\frac{(A-f_1)^2}{2\tau^2}} + (A-f_1)\left(1-2\Phi\left(\frac{f_1-A}{\tau}\right)\right)\right],$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

In column (1), I assume that $\sigma = 0.35, \tau = 0.1$, while in column (2) I assume that τ is 1.41 and σ is 0.4. Across both columns, I fix the magnitude of the difference between actual and μ_0 to 0.6.[10] In both cases, the model predicts that, on average, dispersion should be 0.146 after the eleventh analyst has forecasted. According to Panel C of Table 1, this quantity matches exactly the average running dispersion after the eleventh forecast for firms in the middle tercile of $|LEADFE|$. In the adjacent two columns, I report the expected dispersion under the alternative assumptions that $\tau = 1.1, \sigma = 0.56$ and that $\tau = 1.64, \sigma = 0.7$. According to the model, dispersion should be 0.173 after the third forecast under both sets of assumptions. This amount is consistent with average dispersion observed in the data after the third forecast for firms in the highest tercile of $|LEADFE|$ (see last column of Table 1, Panel C).

Note that the differences between column (1) and (5) thru (6) illustrate the effect of changing either σ or τ while fixing the other parameter as well as μ_0 constant. Comparing column (1) with column (5), increasing σ from 0.35 to 0.4 increases dispersion for the entire sequence. This is consistent with the expectation that

[10]The model predictions are only unique up to the difference between the two. For example, the model implied dispersion is the same for $A = 0.2, \mu_0 = 0.8$ and $A = 0$ and $\mu_0 = 0.6$

dispersion should increase with the uncertainty. In column (6), the assumption about τ is higher than that in column (1). This change reduces analysts' weighting on their private signal and results in a decrease in dispersion of beliefs. Column (7) assumes all the same parameters as column (1) except that μ_0 is 0.8 instead of 0.6. Increasing this parameter reduces the dispersion in beliefs at every point in the forecasting sequence.

Although with the assumed parameters in columns (1) and (2), the model can explain dispersion after the eleventh forecast for mid-$LEADFE$ firms, its predictions about dispersion over the earlier forecasts are too high relative to the data counterpart in Table 1. Similarly, in columns (3) and (4), the selected parameters is able to explain dispersion after the third forecast for high $LEADFE$ firms. However, the model predictions about dispersion as more analysts announce is too low relative to what is observed in the data. As alluded to in the previous section, a model that incorporates additional features of the institutional setting could potentially provide predictions that better explains the data.

To summarize, I show that with certain assumptions about μ_0, τ and σ, the model can produce numerical predictions about the dispersion we should except to observe, on average, which fall in the approximate range as descriptive statistics about dispersion in the data. There are multiple sets of parameters that can explain dispersion exactly at a certain point in the progression of forecast announcement. However, it appears that there are no feasible assumptions about σ and τ which allow the model to reproduce the observed dispersion over the sequence of analysts' forecast announcements.

4.2 Analysts' Payoffs

A. Assumptions

My assumption about analysts' payoffs is primarily motivated by previous evidence suggesting that analysts' career outcomes are related to their accuracy as well as the extent they deviate from consensus (Hong et al., 2000; Hong and Kubik, 2003). However, it is important to acknowledge that there is an extensive assortment of other incentives giving rise to forecast bias that is well established in analyst research.[11] Further, managers are known to manage their guidance and subsequent earnings to just-meet forecasts. I consider the importance of deviation from consensus in analysts' payoffs because there is an intuitive link to the second moment of analyst forecasts (i.e., forecast dispersion). Other types of behavior, which primarily impact the first moment of analyst forecasts, could be incorporated into future extensions of this simple model. Additionally, surveys of sell-side analysts (Groysberg et al., 2011; Brown et al., 2015) reveal that their forecasts are determined by a rich confluence of institutional features which give rise to incentives unrelated, at least directly, to forecast accuracy or to deviation from consensus itself. Rather than using a richer, potentially analytically intractable model, I assume that for analysts other than the lead analysts, their forecasting decision can be approximated by an objective function comprised of the absolute value of forecast error and the absolute value of deviation from consensus (i.e., the average over analysts who have already announced):

$$L_j = |f_j - A| + \lambda \left| f_j - \frac{1}{j-1} \sum_{k=1}^{j-1} f_k \right|, \quad j = 2, 3, \ldots, J, \tag{6}$$

where λ weights the importance of deviating from consensus relative to the importance of being accurate. It may be appealing to refer to λ as the herding parameter. Such a use of herding would be loosely comparable to the setting considered in in Scharfstein and Stein's (1990) herding model. That analysts infer private information impounded in preceding analysts' forecasts is itself considered herding behavior in other papers in this literature irrespective of any career concerns. This second type of herding is non-strategic and is closer to the behavior modeled in Banerjee (1992), though neither models can be directly applied to an analyst forecasting setting. My model allows for both kinds of behavior and, as such, throughout the paper I refrain from applying the term to λ.

Although in (6) I assume that forecast accuracy and deviation from consensus are determined based on their absolute values, I do not rule out the possibility that it is the squared forecast error and squared deviation that actually matter to analysts. However, assuming a quadratic objective function gives rise to equilibrium

[11]Some of these include future access to management (Lim, 2001; Francis and Philbrick, 1993) and investment banking relations(Lin and McNichols, 1998).

forecast strategies where λ is not identified separately from σ and τ. That is, absent an assumption about the value of λ, the parameter of interest σ cannot be uniquely estimated from the data. Evidence from Gu and Wu (2003) and Basu and Markov (2004) also supports a linear loss function. In addition, it is also possible that analysts are not only concerned with their deviation from forecasts that have already been announced but also with the forecasts from peers who have yet to announce. I do not make this latter assumption because, relative to (6), the resulting equilibrium does not sufficiently improve my ability to explain forecast dispersion observed in the data to justify the additional computational complexity. Another assumption implicit in (6) is that all analysts have the same preferences (i.e., λ is common to all analysts). In a separate paper (Xiao, 2015), I allow for the possibility that λ is indexed by j and I show that with other sources of variation in the data beyond forecast dispersion, it may be possible to determine whether the objective function is based on absolute value or squared terms and whether analysts are forward looking with respect to peer forecasts.

Since consensus forecast is undefined for the lead analyst,[12] an additional assumption is required about f_1. I assume that the first analyst forecasts the mean of his beliefs, with $f_1^* = \mu_1$ where μ_1 is given by (2). Since the variance operator is invariant to addition or subtraction by a constant, this assumption is robust to the extent that the first analyst includes an upward bias that does not depend on σ or τ perhaps due to conflicts of interests documented in previous research, as long as the remaining analysts also include the same bias. Most notably, however, this assumption would not hold to the extent that there is the type of self-selection in analyst coverage suggested in McNichols and O'Brien (1997) because observed forecasts would be systematically related to the expectation of a truncated normal distribution, which depend on σ and τ.[13]

[12]Because I assume that the start of each forecasting game commences with the earnings announcement and because there is a dearth of forecasting activity immediately preceding the event, I assume that the average over any forecasts already issued at the earnings announcement to be stale and does not enter into the lead analyst's objective function.

[13]See results in Hayes (1998) and Hayes and Levine (2000) for the derivation this relation.

B. Optimal Forecasting Strategy for Second Analyst

To illustrate the effect of λ on forecast dispersion between the first two analysts, the second analyst's forecasting strategy when $\lambda > 0$ is:[14]

$$f_2^* = \begin{cases} \mu_2 + \hat{\sigma}_2 \cdot \Phi^{-1}\left(\frac{1+\lambda}{2}\right) & \text{if } \mu_2 < f_1 - \hat{\sigma}_2 \cdot \Phi^{-1}\left(\frac{1+\lambda}{2}\right) \\ f_1 & \text{if } \mu_2 \in \left[f_1 - \hat{\sigma}_2 \cdot \Phi^{-1}\left(\frac{1+\lambda}{2}\right), f_1 + \hat{\sigma}_2 \cdot \Phi^{-1}\left(\frac{1+\lambda}{2}\right)\right], \\ \mu_2 - \hat{\sigma}_2 \cdot \Phi^{-1}\left(\frac{1+\lambda}{2}\right) & \text{if } \mu_2 > f_1 + \hat{\sigma}_2 \cdot \Phi^{-1}\left(\frac{1+\lambda}{2}\right), \end{cases} \tag{7a}$$

In contrast, when $\lambda < 0$, the second analyst forecasts according to:

$$f_2^* = \begin{cases} \mu_2 + \hat{\sigma}_2 \cdot \Phi^{-1}\left(\frac{1+\lambda}{2}\right) & \text{if } \mu_2 < f_1 \\ \mu_2 & \text{if } \mu_2 = f_1, \\ \mu_2 + \hat{\sigma}_2 \cdot \Phi^{-1}\left(\frac{1-\lambda}{2}\right) & \text{if } \mu_2 > f_1 \end{cases} \tag{7b}$$

where the mean of the second analyst's beliefs μ_2 is given by (4)[15] and where $\hat{\sigma}_2 = \frac{\sigma^2\tau^2}{2\sigma^2+\tau^2}$, used to denote the variance of the second analyst's posterior beliefs, increases with both σ and τ. The $\Phi^{-1}(\cdot)$ refers to the inverse of the standard normal cumulative distribution function. In Equations (7a) and (7b), the optimal strategy f_2^* is a function of μ_2 and consensus, which in this case, is simply the first analyst's forecast f_1. To further illustrate the second analyst's behavior, Figure I includes plots of f_2^* assuming $\lambda < 0$ and $\lambda > 0$. Both plots are centered at f_1 and I vary μ_2 along the horizontal axis. In both cases, the analyst forecasts honestly μ_2 when $f_1 = \mu_2$ (i.e., the consensus coincides with the mean of his beliefs) and the strategy is symmetric around this point. When $\lambda > 0$, the analyst biases[16] away from μ_2 in the same direction as f_1. For values of μ_2 close to f_1, the optimal forecast for analyst 2 is the corner solution of $f_2^* = f_1$. The range of μ_2, relative to f_1, that induces the corner solution increases with λ, σ and τ. For values of μ_0 sufficiently far from f_1, the analyst issues a forecast between f_1 and μ_2, where the distance from μ_2 is a constant quantity, in the sense that it does not depend on μ_2 or f_2, and

[14]The analytical solutions for $j > 2$ and accompany proofs are included in Xiao (2015).

[15]This equality holds because I have assumed that the $f_1^* = \mu_1$.

[16]Note that through the paper I use the term bias to refer to the difference between f_j^* and μ_j. This quantity may or may not be related to the more conventional notion of bias, which is the difference between observed forecasts and actual.

varies with σ, τ (through $\hat{\sigma}_2$ in Equation (7a) and (7b)) and λ. When λ is negative, the analyst always biases from μ_2 whenever $\mu_2 \neq f_1$. If the mean of analyst 2's beliefs μ_2 is greater than f_1, then f_2^* will be exceed μ_2 and if μ_2 is less than f_1, then f_2^* will be less than f_2^* as well. Similar to the case where λ is positive, the magnitude of the bias increases with σ, τ and λ. Visually, Figure I shows that there is a "jump" in the direction of forecast bias around $\mu_2 = f_1$ and that the size of the jump increases with the magnitude of all three parameters.

C. Numerical Examples of Forecast Dispersion Implied by Model Equilibrium

In Panel B of Table 3, I report numerically the model predictions about expected forecast dispersion for specific assumptions about the underlying parameters. The approach is similar to the one used in Section 4.1, except of course that the previous analysis assumed a honest forecasting setting (equivalent to $\lambda = 0$). I use the equilibrium forecast strategies from solving the loss function assumed in (6). Since the expectation of forecast dispersion cannot be computed analytically, I use a numerical integration routine. To provide maximum comparability with the model predictions reported in Panel A, I maintain the assumption that $\tau = 1.1$ across all the computations.

In the first column, I use the assumption that $\sigma = 0.6, \lambda = 0.1$ while μ_0 is assumed to be \$$0.1$ greater or less than actual earnings. The model predicted dispersion ranges from 0.052 after the third forecast to 0.101 after the eleventh forecast. Referring back to Table 1 Panel C, the running dispersion for firms in the lowest tercile of $|LEADFE|$ is 0.055 at $j = 3$ and 0.093 at $j = 11$, which are approximately the same. In fact, a side-by-side comparison reveals that the model predictions are quite close at explaining running dispersion at every j. The analogous comparison can be made between the sequence of running dispersion for the Mid $|LEADFE|$ firms in Table 1 and column (2) of Table 3. If I assume that $\sigma = 0.9, |\mu_0 - A| = 0.1$ and $\lambda = 0.1$, the model implied dispersion ranges from 0.083 to 0.146 which closely tracks the 0.085 to 0.146 range in the data.

At first glance, it may be counter-intuitive that to the extent analysts prefer to be close to consensus, dispersion actually *increases* as each analyst announces his forecast. To clarify the impact of λ, consider first, for example, the difference in model-implied between columns (1) and (3) where the assumptions only differ by λ. At every j between 3 and 11, increasing λ from 0.1 to 0.7 reduces the predicted dispersion. Underlying this reduction is analysts' choice to deviate from beliefs towards consensus which, in turn, is driven by the fact that their objective function rewards them for being close to consensus. However, a different analysis is required for how dispersion varies with j for a *fixed* positive value of λ (as well as σ and τ). As the $\lambda > 0$ curve in Figure I shows, the analyst only forecasts differently from consensus when the mean of his beliefs, determined by his private signal, is sufficiently extreme relative to consensus. As a statistical regularity, the presence of additional analysts will increase the probability that very unlikely signals will be drawn to induce a forecast that differs from consensus. In addition, the "flat" interval in Figure I is decreasing in the marginal cost of being incorrect. Holding fixed the other parameters, this marginal cost is higher for later analysts because their beliefs are more precise.[17] As j grows large, this interval will disappear and the randomness from s_j will always be reflected in forecast dispersion. Put together, these two features of the model produce predictions about dispersion that increase with j.

Finally, λ also has important implications for the sensitivity of dispersion with respect to σ. In each of columns (1), (3) and (5), I assume that $\sigma = 0.3$ while in columns (2), (4) and (6), I assume that $\sigma = 0.6$. As expected, increasing σ increases overall dispersion (although a portion of this increase is attributable to the change in assumption about μ_0). However note that the absolute size of this increase is the largest (approximately 0.08) across columns (5) and (6) where I assume that $\lambda = -0.7$ while the increase is more modest between columns (1) and (2). At the other extreme where I assume $\lambda = 0.7$, increasing σ by 0.3 results in only a 0.004 to 0.012 increase in dispersion. More formally, the cross-partial of the model-implied dispersion with respect to σ and λ is negative. The intuition for the result in the two analyst case, which generalizes for any j, can be seen from the forecasting function in either (7a) or (7b).[18] σ affects f_2^* directly through the bias term as well as indirectly through μ_2. If $\lambda > 0$, μ_2 only matters to the extent that it is far from

[17]For example, if the analysts are beliefs are so imprecise that their prediction will almost always be wrong, then there is little cost to forecasting something other than his prediction.

[18]Since λ does not enter into the first analyst's strategy, f_1 can be held fixed. Thus, the impact of σ on the dispersion between analysts 1 and 2 is the same as its impact on f_2^*

f_1, so the overall impact of σ on f_2^* is, loosely speaking, the product of its effect on μ_2 and the probability that μ_2 exceeds the interval over which 2 imitates 1. Since this interval tends to infinity as λ approaches 1, σ will play a progressively smaller role in f_2^*. In contrast, when $\lambda < 0$, the analyst always biases away from consensus unless $\mu_2 = f_1$[19] where there is a dollar increase in the distance between f_2^* and f_1 for every dollar of magnitude in difference between μ_2 and f_1. In addition to this effect, the second analyst's departure from f_1 also includes the product of $\Phi^{-1}(\frac{1-\lambda}{2})$ and the posterior variance of his beliefs $\hat{\sigma}_2$ (the second term in Equation (7b)). The former term decreases with λ and the latter increases with σ. As a result, the interaction between σ and λ, f_2^* is extremely sensitive to σ when λ is extremely negative. The large difference in model-implied dispersion between columns (5) and (6) reflects this sensitivity.

To summarize, my empirical model assumes that analysts' payoffs are jointly determined by the absolute forecast error and the absolute deviation from consensus. I parametrize the relative importance of relative importance of the latter measure through the parameter λ in Equation (6). To the extent that $\lambda \neq 0$, the analyst's forecasts are influenced by both their beliefs about earnings and strategic conduct. The model predicts that strategic incentives should manifest in the data in one of three ways. First, strategic conduct arising from positive (negative) values of λ reduces (increases) forecast dispersion for a given number of analysts who have announced. Second, it affects the direction and magnitude of the change in dispersion each analyst announces. Finally, large (small) values of λ attenuates (amplifies) the extent to which cross-sectional differences in σ is reflected in forecast dispersion.

5. ESTIMATION

The model I describe in the previous section has four parameters unobservable to the researcher: the earnings predictions based on public information μ_0, the fundamental uncertainty about earnings σ, the variance of the noise in analysts' private signals τ and the relative weighting of deviation from consensus in analysts' payoff function λ. Whereas μ_0 varies almost certainly by firm and λ is assumed

[19] Since the beliefs are normally distributed, the probability of this event is 0.

to be the same across all analysts, I consider both the possibility that σ and τ are the same across all firms and analysts, respectively, and the possibility that there is cross-sectional heterogeneity. In my discussion of the empirical approach, it is important to emphasize that the intent of the estimation is not to conclusively determine the sign of λ. Rather, I focus on how inferences about σ and τ vary by allowing for the possibility that $\lambda \neq 0$.

I select my sample using the exact procedure described in Section 3. In particular, I use the 19,560 firm-years which were randomly sampled from the 21,787 observations with available data to estimate the determinants model in (1). I merge the IBES data to the Nelson Directory of Investment Research, which contains data on analyst characteristics (e.g., All-Star Status) as well as details of the employing research firm, using the Analyst-Broker Translation File.

5.1 Econometric Assumptions

As shown in the previous section, analysts' forecast strategies depend on the mean of their posterior beliefs about earnings. This mean is a weighted average between their private signals, the earnings prediction based on public information μ_0 ("direct effect") and, in the case of analysts forecasting second and later, inferences about previous analysts' signals which in turn is a function of μ_0. However, the model predicts that this latter function of μ_0 exactly offsets the direct effect of μ_0 in determining beliefs. As such, μ_0 only enters into the first analyst's forecast strategy. This result can be seen, for example, in Equation (4) and generalizes for all $j = 2, 3, \ldots, J$.

Although μ_0 can be estimated from the data using any one of the earnings predictions models from previous research, this process introduces econometric noise that is not informative about the economic primitives of interest. However, since the analytical features of the model allow me to sever the dependence of forecasts on the parameter for all but the first analyst, I do not attempt this estimation. Instead, I assume that the true value of the parameter satisfies $\mu_0 = \frac{f_1(\sigma^2+\tau^2)-s_1\sigma^2}{\tau^2}$, where the expression follows from inverting (2) and the assumption that the first analyst is honest. Clearly there is a trade-off because the actual observed value of f_1 must be used in computing the model-predicted dispersion.[20] To mitigate the concern about the resulting mechanical correlation between the model and data, I only consider dispersion over forecasts starting from the second analyst.

In addition, I assume that in the cross-section, the earnings prediction based on public information for firm i and period t exhibits the following relation to actual earnings:

$$\mu_{0,it} = A_{it} + \sigma_{it} \cdot \xi_{it}, \tag{8}$$

where ξ_{it} is the standard normal variable. That $\mu_{0,it}$ is equal to A_{it} in expectation follows mechanically from how the parameter is defined in Section 4.1. However, an econometric assumption is needed about the extent to which $\mu_{0,it}$ varies relative to A_{it} in the data. Since $\mu_{0,it}$ is an earnings prediction which aggregates, amongst other news, the latest earnings announcement, it is natural to think that this variation is systematically related to the firm's uncertainty itself. Put differently, in the cross-section, firms with high σ are hard to forecast using $\mu_{0,it}$ precisely because past earnings news tends to be far from future earnings. Equation (8) assumes that both sources of uncertainty are exactly equal, although in future extensions I plan to relax this assumption with the inclusion of additional noise terms.

The assumption in Equation (8) along with the assumption that the first analyst is honest can be substituted into (2) so cross-sectionally, the first analyst's forecast error satisfies:

$$f^*_{i1t} = A_{it} + \frac{\sigma^2_{it}\tau_{i1t}}{\sigma^2_{it} + \tau^2_{i1t}}\varepsilon_{i1t} + \frac{\sigma_{it}\tau^2_{i1t}}{\sigma^2_{it} + \tau^2_{i1t}}\xi_{it},$$

where ε_{i1t}, originally defined in Section 4.1 as a component of the first analyst's noise term, is unit normal by assumption and since ξ_{it} is also unit normal, the sum of the last two terms in this expression is also normal with mean 0 and variance equal to the sum of the squared weights on ε_{i1t} and ξ_{it}. Thus I re-write the first analyst's forecast error as:

$$f^*_{i1t} - A_{it} = \frac{\sqrt{\sigma^4_{it}\tau^2_{i1t} + \tau^4_{i1t}\sigma^2_{it}}}{\sigma^2_{it} + \tau^2_{i1t}}\zeta_{it}, \tag{9}$$

[20] Since the model predictions about dispersion are only unique up to the unsigned difference between f_1 and A, knowledge of the magnitude of the first analyst's forecast error is sufficient.

where ζ_{it} is a random unobservable term specific to each firm i and year t and is distributed $N(0, 1)$.

I estimate four models that differ based on assumptions about cross-sectional heterogeneity in σ and τ. I consider the two extreme cases in this heterogeneity. In the "No Heterogeneity" model, I assume that forecast dispersion can be explained by a uncertainty parameter common to all firm-years, with $\sigma_{it} = \sigma$, and that all analysts' private information are equally precise. That is, τ is common across analysts.

In the pairs of models where τ is assumed to vary across analysts, I assume that the parameter is a function of observable analyst characteristics. This choice is motivated by evidence from other studies that analysts' forecast accuracy systematically vary with analysts' ability and resources (Clement, 1999; Mikhail et al., 1997). I select two common proxies used in such studies and specify:

$$\tau_{jt}^2 = \tau_0' + \tau_1' expr_{jt} + \tau_2' expr_{jt}^2 + \tau_3' employersz_{jt}, \tag{10}$$

where $expr$ is the analyst's forecast experience, measured in the number of years where the analyst has issued a forecast in IBES and $employersz$ is the log of the total number of analysts employed at the same brokerage firm as analyst j, both of which are measured at the calendar year end previous to time t. These two measures are rough proxies for the precision of analyst's information, where in future extensions I will incorporate additional and more refined measures. There is no theoretical basis for the functional form assumed in (10). I include $expr^2$ along with $expr$ to allow for declining returns to forecast experience but the appropriateness of this specification is untestable. In addition, the two variables could linearly affect the standard deviation τ_{jt} or precision parameter $\frac{1}{\tau_{jt}}$ rather than the variance itself. I consider the fit of these alternate specifications as a robustness check.

Note that the approaches which assume homogeneity or arbitrary heterogeneity represent two extremes along the bias-variance trade-off, where allowing for heterogeneity reduces bias but increases variance. Although I consider only the two "corner case" assumptions, depending on the particular research setting that estimates of σ is intended for, there could be other assumptions more appropriately positioned on the bias-variance frontier.[21]

In the Firm Heterogeneity model, the cross-sectional heterogeneity in disagreement is driven entirely by the heterogeneity in the uncertainty parameter σ (through its role in determining the mean of beliefs μ_j). That is, a cross-sectional sort of firms using uncertainty produces the same result as a sort using disagreement. In the Complete Heterogeneity model, the cross-sectional variation in disagreement is jointly determined by firm heterogeneity (through its effect on σ) and analyst heterogeneity (through its effect in τ under the assumption in Equation (10)). In particular, while I allow for the possibility that disagreement has a firm-specific and an analyst-specific component, I do not allow for a *firm-analyst* pair specific component (i.e., τ is indexed by i, j and t). As intuition would suggest, identification of heterogeneity at such a granular level would be challenging.

5.2 Identification

In Section 3, I computed the average running dispersion for subsamples of firms sorted based on the magnitude of $|LEADFE|$. For the purposes of this stylized discussion, I assume that the average dispersions reported in Table 1 corresponds to a single firm (or a group of homogenous firms).

A. Identification of σ and τ

For simplicity, first consider Panel A of Table 3 where I assume analysts forecast honestly or, equivalently, λ=0. The table shows that when for a fixed value of μ_0, there are multiple permutations of σ and τ such that the model-computed forecast dispersion will reasonably reproduce the average forecast dispersion in the data at a particular point. For those particular pairs of σ and τ, the model will also generate close to identical dispersion for the entire forecasting sequence. Therefore, the two parameters cannot be identified solely using dispersion data. The intuition is that based on only observing forecast dispersion, either in the cross-section or within firm-period, one cannot distinguish whether the variation is driven by noise in analysts' private information or by uncertainty in the earnings prediction based

[21]For example, one could assume that σ is constant within an industry. Alternatively, if the empirical test requires sorting firms into portfolios based on dispersion, the σ could be estimated as an average for firms in the top quintile.

on public information.[22] This necessitates the use of data on the location of forecasts relative to the earnings prediction itself.

Since the model assumes that the earnings prediction is equal to actual in expectation, the a natural variable to use is the relative distance between forecasts and actual earnings (i.e., the magnitude of the forecast error).[23] The BKLS measure of uncertainty is also constructed using actual forecast data. However, whereas the earlier measure is based on the magnitude of the consensus forecast error, I use the unsigned forecast error for the lead analyst. The motivation for this choice is that Equation (2) provides a closed form expression for the expectation of lead analyst's forecast error (under the assumption that analyst 1 forecasts his beliefs) while the expected forecast error for subsequent analysts have to be simulated numerically. However, this choice does not affect the estimated parameters. Finally, using actual earnings fundamentally changes my estimates of σ and τ into ex-post measures. As noted in Sheng and Thevenot (2012), ex-ante measures of uncertainty may be more appropriate depending on the empirical setting. Certainly, any portfolio tests of expected returns has to be constructed using an ex-ante measure. I discuss in the last section that, notwithstanding my use of ex-post earnings data to estimate the parameters, the resulting measures of σ and τ can be easily converted into an ex-ante measure.

Referring back to Panel A of Table 3, although the progression of model predicted dispersion is essentially identical in the first two columns and last two columns, respectively, there are notable differences in the expected forecast error. By exploiting the observation that the bottom quantity in the first (fourth) column is closer to the average $|LEADFE|$ reported in Panel C of Table 1 for the middle (top) tercile of the variable (1.541 and 0.366, respectively), I conclude that the parameters in these two columns are more plausible than those in the second (third) columns. However, for most plausible parameters of σ and τ, the model-predicted lead forecast errors

[22]It can be shown that only the ratio between σ and τ are identified. There could be a case made that if the researcher is interested in only the cross-sectional variation in σ, a monotonic transformation of the parameter suffices. However, to the extent that τ also differs in the cross-section, then the economic interpretation of this ratio becomes less clear.

[23]Although, the underlying this assumption is that earnings predictions are statistically unbiased, I continue to use bias to refer exclusively to the extent to which analysts' forecast depart from their beliefs. Similarly, the observed forecast error may or may not be related to bias under this latter definition.

do not appear to match the quantities in the data even in approximate terms. For example, the difference between the an average $|LEADFE|$ of 1.541 and the model prediction of 0.677 is quite substantial.

B. Identification of λ

From the discussion in Section 5.1, the observed forecast dispersion at a specific point (for example, after all analysts have forecasted) in the forecasting sequence allows me to estimate one of σ and τ while the magnitude of the forecast error distinguishes between the two uniquely. However, my model also provides predictions about the progression of dispersion after each analyst has forecasted which is also observed in the data. It is this variation which I use to estimate to the weight on deviation from consensus in analysts' payoffs, λ. After extensive simulations of the equilibrium in my model, I find that for most plausible ranges of σ and τ, there are admissible values of λ such that the forecast dispersion will increase, decrease or do both over the forecasting sequence. Further, the magnitude of these increases/decreases appear to be sensitive to even modest changes in λ. Though identification can not be proven theoretically, my analysis strongly suggests that λ can be recovered using the joint distribution of forecast dispersion and forecast error.

To provide additional intuition for how λ can be recovered, consider for example the comparison between the first column in Panel A and the second column in Panel B of Table 3. In Panel A, where $\lambda = 0$ by assumption and where I fixed σ at 0.35 and τ at 1.1, the model-predicted dispersion of 0.146 is exactly consistent with the data at $j = 11$ but too high for all the earlier forecasts. As I gradually increase λ, the progression of dispersion will change to an increasing pattern. However, holding fixed the assumption about τ, a higher value of σ is needed such that dispersion after the 11th forecast is still 0.146. Of course, increasing σ also has an effect on the model-implied lead forecast error. More generally, allowing λ to be positive (or negative) will improve the model's ability to explain observed dispersion at certain points in the forecasting sequence while trading off the fit at other points. This latter deterioration can be compensated by adjusting σ and τ, but the adjustment will in turn affect how well the model explains lead forecast error. In the case of the example I use in Table 3, changing in assumptions about λ and σ produced improvements along both dimensions.

C. Identification of cross-sectional heterogeneity in σ

Much of the preceding discussion is premised on the model-predicted path of dispersion for a fixed value of σ. Of course it is unlikely that the parameter within a dataset which we think exhibits vast across-firm heterogeneity in the quality of the information environment. However, Section 4.2 notes that when $\lambda \neq 0$, then forecast dispersion may exhibit very extreme sensitivity or almost no sensitivity to underlying heterogeneity in σ. It follows that the magnitude of the cross-sectional heterogeneity in σ is not identified separately from λ. That is, for a one unit of cross-sectional heterogeneity in σ, the model can generate anywhere between 0.1 and 10 units of cross-sectional variation in dispersion. Conversely, one unit of cross-sectional variation in forecast dispersion can be rationalized by as low as 0.1 or as high as 10 units of heterogeneity in σ (along with appropriate assumptions about λ).

To further clarify the previous discussion of the negative cross-partial of dispersion with respect to λ and σ, consider again the second analyst's forecast strategy in Figure I. For the $\lambda > 0$ case, the flat section of the $\lambda > 0$ plot implies that there is a range of μ_2's that will generate produce identical forecasts. As such, if the researcher approximates σ_{it} with the sample average ($\bar{\sigma}$), the choice has no effect on the model fit as long as $\bar{\sigma}$ is sufficient close to σ_{it}. In contrast if $\lambda < 0$, then any approximation noise in σ_{it} will be amplified by λ and induce substantial variation in the model prediction.

5.3 Moment Restrictions

The three unobserved parameters are $\theta \equiv (\sigma, \tau, \lambda)$, where following the discussion the previous section μ_0 no longer has to be estimated and where each of σ and τ could be vector that corresponds to each firm-year or analyst in the data. These have to be chosen simultaneously such that, on average, multiple predictions from the model match their data counterparts. Further, all 3 parameters, or sets of parameters, have a highly non-linear effect on the model predictions. As such, it is intractable to estimate σ, τ and λ with an OLS approach. Rather, I use Simulated Method of Moments (SMM), which formalizes the procedure I used to produce Table 3 3 such that they are close to Panel C of Table 1. With the addition assumptions I made in Equation (8) about the cross-sectional variation in $\mu_{0,it}$, it would no longer suffice that the model explains dispersion and forecast error for one particular tercile in the latter table. Instead, the differences in $LEADFE$ across the bottom, middle and top terciles have to be rationalized by the estimated value of σ.

The model generates a set of predictions about the sequence of dispersion after each analyst has forecasted, which I formally define as:

$$\Delta_j(\theta, |LEADFE|, \mathbf{x})$$
$$= \sqrt{\frac{1}{j-1}\sum_{k=2}^{j}\mathbb{E}_{\varepsilon_2,\ldots,\varepsilon_{11}}\left[\left(f_k^*(\theta, |LEADFE|, \mathbf{x}) - \frac{1}{j-1}\sum_{l=2}^{j} f_l^*(\theta, |LEADFE|, \mathbf{x})\right)^2\right]},$$

where $\mathbf{x}$ is a $J \times 2$ matrix of analyst-specific characteristics. In addition, I use D_{ijt} to denote the running dispersion that I compute in Panel C of Table 1. Recall that the variable is defined as the population standard deviation over the 2nd through the j-th forecast. For each j between 3 and 11, the following moment restriction can be used:

$$h_{disp,j}(\theta, |LEADFE|_{it}, \mathbf{x}_{it}) = \sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j\in[3,J_{it}]}\Delta_j(\theta, |LEADFE|_{it}, \mathbf{x}_{it}) - D_{ijt}, \qquad (11)$$

where the inner-most summation is used because, obviously, we do not observe running dispersion for a firm-year in the data beyond the total number of analysts who issued forecasts.

The model also provides predictions about analysts' forecast error. In particular, note that in Equation (9), $|\zeta|_{it}$ follows the standard half-normal distribution, with mean $\sqrt{\frac{2}{\pi}}$. So another moment restriction (or another set of moment restrictions) can be constructed using the observed and model prediction about the first analyst's forecast error:

$$h_{fe}(\theta, |LEADFE|_{it}, \mathbf{x}_{it}) = \sum_{i=1}^{N}\sum_{t=1}^{T} 1 - |LEADFE|_{it}\sqrt{\frac{\pi}{2}}\frac{\sigma_{it}^2 + \tau_{i1t}^2}{\sqrt{\sigma_{it}^4\tau_{i1t}^2 + \tau_{i1t}^4\sigma_{it}^2}}. \quad (12)$$

In the two models where I do not consider firm-heterogeneity, (12) is used as a single additional moment in addition to those based on dispersion. In future iterations of the paper, I plan on using two separate moment restrictions depending on the sign of the $LEADFE$.

In the Firm Heterogeneity and Complete Heterogeneity models, rather than averaging over the entire sample, I require that h_{fe} holds exactly at every i and t over a trivial sample size of 1. Note that in general, doing a numerical search for tens of thousands of σ_{it} is computationally infeasible. Under this approach, for every candidate guess about τ in the estimation, σ_{it} can be solved in closed-form as a function of $LEADFE_{it}$. This is equivalent to imposing an infinitely large weight on h_{fe} relative to $h_{disp,j}$.

Since the sample moments in (11) are computed over vectors of differing lengths, the standard theory on the optimal weighting matrix does not apply. In the No Heterogeneity and Analyst heterogeneity models, I weight each of the moments equally. In the Firm and Complete Heterogeneity Models, I weight the dispersion moments using the length of the moment vector (i.e., the number of firms-years where there was a j-th forecast). As discussed, for these two models, I set the forecast error moments equal exactly. Although these choices are ad-hoc, the weighting matrix only affects efficiency and have no bearing on the consistency of my estimates.

6. Results

6.1 Parameter Estimates

Table 4 contains the parameter estimates for each of the four models with differing assumptions about the cross-sectional heterogeneity. With the No Heterogeneity ("NH") model, I find that the noise in analysts' private signal has standard deviation $\tau_{it} = 1.05$ while the uncertainty in the earnings prediction based on public information is $\sigma_{it} = 1.21$. The economic interpretation of the estimates can be expressed in terms of a well-known property of the normal distribution. In the case of σ, the interpretation is that, on average, actual earnings will deviate from the earnings prediction by less than \$1.21 (\$2.42) approximately two-thirds (95%) of the time. Similarly, $\tau = 1.05$ has the interpretation that there is a two-thirds chance that the subjective component of analyst's information will fall within \$1 of actual earnings. Alternatively, the estimate for τ can be thought of as the standard deviation of analysts' earnings predictions in the absence of any public news. In addition, my estimates of σ and τ can be used to compute the variance of each analyst's posterior beliefs (i.e., $\hat{\sigma}$ in Equations (7a) and (7b)). The model implied reduction in average uncertainty as a result of analysts' forecasting activity is 35%. Finally, the NH model suggests that analysts are rewarded for being close to consensus, with $\lambda = 0.288$.

In columns (2)-(4) of Table 4, I report the estimated parameters assuming that σ_{it} differs for each firm. In the Firm Heterogeneity ("FH") model, the estimates of τ and λ are 1.43 and 0.087, respectively. Since the moment restriction in (12) holds exactly by construction, I can substitute the estimated value of τ into the expression to solve for σ_{it}. In contrast, when I fix σ to be homogeneous but allow τ to vary by analyst, the Analyst Heterogeneity ("AH") column indicates that the variance of analysts' signals τ_{jt} decreases with $expr_{jt}$ (coefficient is -0.18), which is consistent with previous evidence showing that analysts' accuracy increase with experience. The coefficient on $expr^2$ is positive, suggesting that there is a declining return to experience. Also in the AH model, τ increases with the size of the analyst's research firm, although the estimate is not significantly different from zero.[24] The estimated of uncertainty ($\sigma = 1.24$) is similar to that from the NH model. Based on these estimates, on average an additional year of experience improves the analyst's forecasting ability such that there is a 4% reduction in uncertainty. Finally, in the Complete Heterogeneity ("CH") Model, I find that the relation between τ_{jt} and $expr$ is similar to the relation in the AH model. I also find that λ is 0.2 in the CH Model.

[24]For extreme values of $expr$ and $employersz$, the expression for τ_{it} becomes negative. Thus, I use the $[\cdot]^+$ notation to impose a 10^{-4} lower bound on the parameter.

6.2 Assessment of Model Fit

In the bottom row of Table 4, I report the J-Test statistics for all four models. Note that due to lack of theory on the appropriate weighting matrix when the sample moments are constructed from vectors of varying lengths, the statistics are not exact. I compute the J-Test using the length of the longest moment vector (i.e., the total number of firm-years in the sample) and even with this conservative approach, I do not reject the null hypothesis in the J-Test of over-identifying restrictions at the conventional levels for three out of the four models.[25] Since the four models differ by the choice of moments and moment weighting matrix, they cannot be ranked on their fit to the data in a statistical sense. Based on the qualitative observation that the sample and model implied path of dispersion differ by no more 0.08 in all four models, I conclude that all models perform equally well at explaining observed dispersion even though there are substantial differences in assumptions.

To further assess the validity of my model, I consider the out-of-sample performance at predicting forecast dispersion. Using the $2,227$ randomly selected observations from the sample described in Section 3, I compute the model-implied dispersion at the estimated parameter $\hat{\theta}$. I then compare the overall dispersion in the data with the model-prediction over the same number of analysts. For example, if a firm-year in the holdout sample was covered by four analysts over the post-earnings announcement window, I use my model prediction for the dispersion after the fourth analyst. The root mean of the squared prediction error ("RMSE") is reported in Table 5. The RMSE value for the No Heterogeneity model is 0.164. The Firm Heterogeneity model under-performs with the highest RMSE of 0.230. Adding analyst heterogeneity to this model lowers the RMSE slightly to 0.217. These two results are unsurprising considering that in my estimation procedure, I required that the model prediction about forecast error match the data exactly. Inherent in this requirement is a trade-off in how well the model is able to explain dispersion. In contrast, with the assumptions that there is no firm heterogeneity in σ, both forecast error and dispersion are weighted approximately equally in estimating the

[25]Although the p-values I report are based on the usual χ^2 critical values, the asymptotic distribution of the statistic is probably not exact.

parameters. Finally, the model with only analyst heterogeneity exhibits the best out of sample performance with a RMSE of 0.153.

To provide a benchmark of the performance of more descriptive approaches, I use the regression estimates in Table 2 (i.e., the $\hat{\gamma}$'s in Equation (1)) to form out-of-sample predictions.[26] The resulting RMSE of the prediction is 0.228, which exceeds the RMSE for all four model-based approaches. Of course, this may not be a misleading comparison because in the model-based approach requires the use of realized earnings as an input. To mitigate this concern, I also consider the prediction performance of the determinants regression in (1) augmented with consensus forecast error as an additional explanatory variable. The resulting RMSE is 0.219 which is lower than that for the Firm Heterogeneity model but still under-performs relative to the other three model based approaches. Finally, predictions from a regression with only consensus forecast error has a RMSE of 0.237.

6.3 Simulation of Counterfactual Policies

I consider various counterfactual simulations, where I form predictions about average model-implied dispersion from changing one model parameter or assumption while holding fixed the other assumptions and parameters at their estimated values. In each counterfactual, I use the estimates I obtained from Table 4 where there is no heterogeneity in σ and τ across firms and analysts. I first consider a policy which would remove their peer incentives and compensate them only for their forecast accuracy, which is equivalent to assuming that $\lambda = 0$, and a policy which would reverse the direction of peer incentives. That is, whereas my estimation produces $\hat{\lambda} = 0.288$, I assume that $\lambda = -0.288$ instead. Additional counterfactuals policies include doubling or reducing by half the precision of the earnings prediction based on the earnings announcement news σ and the precision of analysts' private information τ. Finally, I impose a policy where all the parameters remain the same as those reported in Table 4 but analysts do not observe previous analysts' forecasts and thus are not able to infer the embedded private information.

The predicted average dispersion after the j-th forecast for the counterfactual

[26] The prediction is appropriately adjusted for the fact that the dependent variable in the regression has been log-transformed

policies are reported in the rows of Table 6, with the actual dispersion observed in the data also reported for reference. Bootstrapped standard errors for the predicted average dispersion appear in parentheses below the point estimates. Note that in some cases, the direction of the counterfactual change in dispersion relative to observed may differ from that relative to the model-implied dispersion at the estimated parameters.[27] The counterfactual policies in rows (1), (2) and (7) exhibit the largest difference relative to observed dispersion. Increasing the precision of public earnings announcement news two-fold, an apparently substantial economic shock to the information environment, would reduce average dispersion after the third analyst's forecast by a mere 0.014, from the observed value of 0.128 to 0.114 while changing analysts' incentives such that they are only concerned about their forecast accuracy would produce a change of 0.318 so that average dispersion after the third analyst's forecast would be 0.446. Reversing the sign of peer incentives or switching to an independent forecasting setting where analysts are unable to infer previous analysts' information would produce similarly large effects. Collectively, the results indicate that forecast dispersion would be relatively insensitive to changes in analysts' information environment such as changes in the amount of public disclosures or the quality of analysts' private research activities. Contemporaneous or independent policy changes to observability of peer analysts' actions or to the incentive environment would effect more pronounced changes.

Table 6 also illustrates the importance of accounting for the number of analysts forecast dispersion is measured over when analyzing the effects of prospective policy changes. In row (5), a counterfactual policy which increases the standard deviation of the noise in analysts' private information τ results in a forecast dispersion of 0.119, 0.143 and 0.158 after the third, fourth and fifth analyst, respectively. These levels are in fact less than the observed levels of 0.128, 0.145 and 0.155. The policy would only increase the average dispersion after the sixth analyst. As discussed in Section 4.2, increasing τ reduces the weight analysts place on their private information as well as increase the variability of the private information

[27]For example, in row (4), the counterfactual dispersion after the 11-th forecast is 0.170. This estimate is greater than its data counterpart of 0.168 but less than the model-implied dispersion at the estimated parameters. In any case, the 0.002 difference between the counterfactual and observed dispersion is within one standard error of the point estimate and thus, it is not possible to reject the null hypothesis that the two estimates are the same at the conventional levels.

itself. Between the two opposing effects, the former dominates when dispersion is measured amongst only early analysts and as more analysts forecast, the latter effect becomes the dominating one. In addition, comparing rows (1) and (2) shows that a counterfactual which rewards analysts for deviations from peer forecast produces lower dispersion relative to one which removes all peer incentives up through the first six analysts but higher dispersion after the seventh analyst forecasts.

6.4 Implications of Estimates for Measuring Uncertainty

Consider the expression for σ_{it} in the CH column of 4 or others based on variants of the model-based approach as a candidate measure of uncertainty for future empirical studies. There are several advantages to the measure. First, since it is consistent with a theoretical model about the economic setting, it allows the researcher to infer the dollar per share variability of the earnings distribution.[28] When combined with the expression for τ_{jt}, the researcher also learn something about the dollar per share change in this variability as analysts announce their forecasts. The firm-level measure can be computed with a variable that can be easily constructed in the data and a straight-forward formula with clear economic intuition. Further, theoretical intuition suggests that inferences from analyst data about uncertainty should be adjusted for the characteristics of the analysts and the comparison of the out-of-sample performance of models with and without analyst heterogeneity in Table 5 confirms this intuition. Equation (10) can be expanded to include arbitrary permutations of analysts characteristics thought to be important for inferences about uncertainty such that the resulting estimates of τ and σ are theoretically consistent.

However, there are two primary drawbacks. First, my model-based approach requires data on realized earnings and thus produces an ex-post measure of uncertainty. As I discuss in Section 4.2, any inferences about uncertainty from analyst data beyond a rough approximation of the ratio of σ and τ necessitates the use of an earnings expectation in some form. If an ex-ante measure is required in a

[28]Of course, historical earnings volatility also provides this interpretation. The premise underlying the use of analyst-based measures is that it is a more contemporaneous measure to the extent that there is time-varying volatility.

particular research setting, an earnings expectation based on, for example, the cross-sectional regression approaches in Hou et al. (2012) and So (2013) can be substituted in place of actual earnings in my proposed measure, although recent work in Gerakos and Gramacy (2013) suggests that a simpler random walk model performs equally as well as more saturated models.[29] Of course, to the extent that the superiority of the model-based measure relative to a descriptive approach that also uses realized earnings as an input demonstrated in Table 5 disappears when the comparison is made using expected earnings, scrutiny of the deficiencies in the earnings prediction model is beyond the scope of this paper.

A second concern is that my firm-level estimates of σ does not exhibit the optimal out-of-sample performance in Table 5. As alluded to earlier, allowing for full heterogeneity in σ increases the variance of estimates. There are computational limitations on searching over tens of thousands of parameters that simultaneously explain forecast error and dispersion. Thus, the choice to prioritize fitting the former moment over the latter also increases the variance. The former source of variance can be reduced by applying some smoothing between observationally similar firms, such as those in similar industries. Regarding the latter source of variance, improvements are theoretically possible with development of more efficient numerical routines. However, it is not clear that they would enhance our inferences about uncertainty in an economically meaningful way.

6.5 Impact of Measure on Empirical Applications that Use Dispersion

To explore the implications of my estimates of earnings uncertainty, i.e., σ_{it}, I re-examine results from a well-known empirical test from prior research based on dispersion. Specifically, I perform portfolio return tests similar to those from Diether et al. (2002) using both dispersion and my model estimates of earnings uncertainty.[30] Table 6 presents the average monthly returns, in percentages, based on sorting by either forecast dispersion or my model estimates. RET6, RET12 and

[29] It is acceptable to use actual earnings to estimate the model parameters as long as a separate sample is used for prediction.

[30] In addition to the fact that Diether et al. (2002) is arguably one of the more well-known empirical results amongst studies on forecast dispersion, examining realized returns has the advantage over other outcomes such as the implied cost of capital due to the well-known empirical measurement issues associated with the latter.

RET24 are average raw returns measured over the period starting from the month after the 30 day post-earnings announcement window used for my sample selection and ending six, twelve and twenty four months later, respectively. ARET6, ARET12 and ARET24 are the market-adjusted counterparts of RET6, RET12 and RET24, where the market-adjusted returns are computed by subtracting raw returns by the CRSP value weighted return for the same month.

Panel A summarizes the average future returns for each $DISP$ quintile. The raw one-year ahead monthly return decreases monotonically, on average, across the 5 $DISP$ portfolios. A hedge strategy from buying low $DISP$ firms and selling high $DISP$ firms generates monthly returns of 90 basis points, or 10.8% on an annualized basis. My results are quantitatively similar with those from Diether et al.'s (2002) portfolio return tests despite some notable differences in methodology.[31] Returns over the six month horizon are largely similar, as are the market-adjusted returns. Over the 24 month horizon, the returns from the hedge portfolio drops to 68 basis points a month, or 8.16% per year.

Panels B and C report average returns based on a sort of firms using estimates of earnings uncertainty from the Firm Heterogeneity model, which I denote using σ^{FH} and from the Complete Heterogeneity model, which I denote using σ^{CH}. These estimates were originally obtained using data on actual forecast errors. For the purposes of the portfolio tests, I substitute a measure of expected forecast error in place of actual forecast errors. I use the cross-sectional earnings regression approach from Hou et al. (2012) to compute expected forecast error. Specifically, the approach involves estimating pooled cross-section regressions of two-year ahead earnings on a constant, current earnings, total assets, accruals, the dividend payment, an indicator equal to 1 for dividend payers and 0 otherwise as well as an indicator variable for firms with negative earnings and 0 otherwise using previous ten years of data.[32] I construct the expected forecast error using the coefficients obtained from fitting this model. Hou et al. (2012) finds that, on average, the

[31]Specifically Diether et al. (2002) uses one-year ahead earnings forecasts with monthly rebalancing of portfolios based on forecast dispersion measured over the previous month while I use two-year ahead forecasts based on dispersion after annual earnings announcements. Additionally, the earlier paper uses a version of dispersion scaled by the consensus forecast (where firms with a consensus of $0 are sorted into the highest dispersion quintile) while I use the raw standard deviation without scaling.

[32]See Equation (1) and accompanying discussion in Hou et al. (2012) for the full details.

adjusted R^2 from its earnings regressions is 0.81.

In Panel B, I continue to find that firms in higher quintiles of σ^{FH} earn lower realized returns. However, the magnitude of such a decline is much smaller. The twelve month hedge portfolio from a trading strategy based on σ^{FH} generates only a return of 36 basis points a month (4.32% annually), which is likely smaller than the trading costs required for implementing such a strategy. Comparing the results in Panels A and B, the difference in future returns between highest and lowest quintiles drops by a factor of two-third when I consider the model estimates of uncertainty rather than raw dispersion. This decline is more striking in Panel C where firms are sorted by σ^{CH}. In Panel C, firms with differing levels of earnings uncertainty, as estimated from the Complete Heterogeneity model, earn similar levels of future returns. The returns from the long-short portfolio based on the lowest and highest quintiles of σ^{CH} are not significantly different from zero. Across both Panels B and C, inferences are similar irrespective of the horizon considered or the use of raw versus market-adjusted returns.

Overall, the fact that using a theoretically-based estimate of earnings uncertainty generates a different association with future returns compared to the association from using forecast dispersion underscores the importance of quantifying strategy separately from uncertainty and casts doubt on conclusions from previous research that the latter association is necessarily anomalous. Although my analysis is not intended to be used for a conclusive reconciliation between theory and evidence on the market pricing of uncertainty and disagreement, it is loosely consistent with the Diether et al. (2002) explanation that disagreement between analysts about firm fundamentals proxies for underlying disagreement between investors and that this latter type of disagreement gives rise to a discount. Specifically, my evidence suggests that removing the portion of forecast dispersion attributable to the noise in analysts' private signals (i.e., their subjective interpretations of earnings news) has an attenuating effect on the dispersion-return relation. My finding that earnings uncertainty exhibits either no correlation or modestly negative correlation could still be interpreted as anomalous in the sense that firms with risky assets should earn a risk premium relative to those with safe assets. However, as Johnson (2004) asserts, the cross-sectional variation in earnings uncertainty could itself by driven by both riskiness in underlying assets and the risk arising from incomplete information on the analysts' part. To the extent that this latter kind of risk dominates the former, then the option value model in Johnson (2004) provides an efficient-market based explanation for my result, though the predominant theoretical prediction in the asset pricing literature is that information asymmetry risk should result in a risk premium.[33]

6.6 Additional Analysis and Robustness Checks

I supplement the main empirical estimates with a specification that includes heterogeneity in analysts' objective functions. Specifically, I assume that in Equation (6), λ varies depending on the analyst's employer. I identify the research firm for the analyst corresponding to each forecast and sort the research firms into size terciles, where size is determined using the number of analysts who issued at least one forecast in IBES during the previous calendar year.[34] Further, λ_{small}, λ_{medium} and λ_{large} denote the estimate of λ corresponding to analysts employed in the lowest, middle and highest size terciles of research firms.

The results of estimating the model with heterogeneity in analysts' objective functions are presented in Table 9. In column (1), where I assume that all firms' earnings predictions using public news has the same precision σ and all analysts' private information contain an equal amount of noise τ, I find that $\lambda_{small} = 0.777$, $\lambda_{medium} = 0.232$ and $\lambda_{large} = 0.111$. According to this set of estimates, analysts who work for larger research firms are only modestly penalized for deviating from peers relative to those working for smaller firms. Of course, in a specification where I assume that the precision of private information τ is the same across analysts irrespective of the size of research firm they work for, it is possible that to confound heterogeneity in τ for heterogeneity in λ. To this extent, I also estimate the model with heterogeneity in τ, with the corresponding results reported in column (2). Here, I find that $\lambda_{small} = 0.045$, $\lambda_{medium} = 0.735$ and $\lambda_{large} = 0.470$. Although I find that λ is positive for all 3 groups, λ_{small} is not significantly different from zero. It is no longer clear from these estimates how the strength of analysts' peer incentives

[33]See for example, Easley and O'hara (2004) and Diamond and Verrecchia (1991).

[34]Parameter estimates are similar irrespective of whether this sort is determined based on pooling together all analysts for a calendar year or pooling together all forecast-analyst pairs.

vary with the size of their employers. My estimates of σ are 1.16 and 1.41 and in the two specifications. Further, I find that more experienced analysts and those working for larger firms have smaller estimated values of τ's.

In footnote 4, I raised the concern that, due to Jensen's Inequality, the standard formula for forecast dispersion (i.e., the square root of the sample variance) produces a biased estimate of the population standard deviation. This bias impacts both the theoretical and empirical moments equally, but could potentially introduce distortions in how each of the dispersion moment restrictions $\Delta_{disp,j}(\cdot)$ are weighted in the SMM objective function. A well-known property of the SMM estimator is that consistency holds for any choice of weighting matrix and as such, this specific source of bias impacts only the efficiency of my estimator.

Nevertheless, to formally address the concerns about using the sample standard deviation, I present in Panel A of Table 9 the observed and the model-implied sample variance computed at the estimated parameters from the NH specification in Table 4. I use this procedure because making the conventional degrees of freedom adjustment in computing the sample variance results in an unbiased estimate of the population variance and thus removes any concerns about the impact of bias on my estimates. The results show that the estimates from a SMM objective function constructed from dispersion moments generate predictions about variance moments that are reasonably close to the average variance observed in the data as each analyst j forecasts. For example, in the data the average sample variance after the third (eleventh) forecast is 0.053 (0.061) while the predicted sample variance using the dispersion-based estimates is 0.047 (0.071). The standard errors corresponding to the model-implied predictions, computed by simulating 1,000 draws from the distribution of the unobservable term, are small - between 0.001 and 0.002. Thus the model predictions are significantly different from average variance in the data. However, the two estimates are similar economically.

To further address concerns arising the computation of forecast dispersion, I repeat my estimation by substituting the dispersion moment restrictions with their variance counterpart. That is, instead of using Equation (11) in constructing the SMM objective function, I replace the moment restrictions with ones based on sample variance:

$$h_{disp,j}(\theta, |LEADFE|_{it}, \mathbf{x}_{it}) = \sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j\in[3,J_{it}]}[\Delta_j(\theta, |LEADFE|_{it}, \mathbf{x}_{it})]^2 - D_{ijt}^2,$$

where $\Delta_j(\theta, |LEADFE|_{it}, \mathbf{x}_{it})$ and D_{ijt} are as defined in Section 5 and where I retain the moment restriction based on the first analyst's forecast error as defined in Equation (12). Similar to the procedure outlined in Section 5, I consider four specifications where there is no cross-sectional heterogeneity in the parameters ("NH"), there is analyst-level heterogeneity in τ ("AH"), firm-level heterogeneity in σ ("FH") as well as heterogeneity along both dimensions ("CH").

The estimates from this modified SMM objective function are reported in Panel B of Table 9. Focusing on the NH specification where all firms and analysts are identical, the estimate of σ and τ are 1.26 and 1.02, respectively, which are comparable to the estimate from Table 4 of 1.21 and 1.05. The estimate of λ is 0.31. All three estimated parameters are economically indistinguishable from those obtained from Table 4 and two of them are within the 95% confidence interval obtained from the earlier table. Thus, I conclude that the main results from my study are reasonably robust to the bias arising from the use of sample dispersion.

7. Conclusion

Quantifying the uncertainty about firm earnings and, in particular, its interaction with analysts' forecasting behavior plays a critical role in understanding the implications of information risk for a large assortment of capital market outcomes. The prevailing assumption underlying numerous accounting and finance studies is that analyst forecast dispersion is an appropriate measure of uncertainty or related constructs. Descriptive evidence from previous studies has identified the presence of biased behavior which cast doubt on such an assumption. My descriptive analysis provides various facts about the within-firm progression of dispersion as well its relation with the magnitude of analysts' forecast error, where certain aspects of these facts are also suggestive of biased behavior. To better understand the economic setting that gives rise to observed dispersion, I develop a simple model based on highly stylized assumptions about the information process which shapes analysts' beliefs immediately after earnings announcements and about analysts' indirect utility function. The equilibrium resulting from the model can reproduce many of the descriptive facts in the data that would be difficult to explain absent formal theory.

Although the model predicts an increasing relationship between dispersion and uncertainty for fixed assumptions about the other model parameters, the interaction between the other unobservable parameters and the strength of this relationship is fairly complex. This interaction has important implications for the researcher's ability to infer from observed dispersion the latter absent some assumptions about the former. I proceed to estimating the model parameters under four sets of assumptions which differ on the extent to which firms and analysts differ. Confirming my analytical predictions, even though the estimates themselves may vary substantially depending on assumptions, all four models provide a comparable fit to the observed data. My model-based estimates exhibit superior out-of-sample performance relative to those from atheoretic approaches. Furthermore, tests of associations between my model estimates of earnings uncertainty and future returns alter the inferences from previous empirical papers which examine associations between forecast dispersion and future returns. Collectively, the model-based approach produces measures of uncertainty that could enhance the economic interpretation from future studies about the relation between forecast dispersion and asset prices.

REFERENCES

Ajinkya, B. B., R. K. Atiase, and M. J. Gift. 1991. Volume of trading and the dispersion in financial analysts' earnings forecasts. *The Accounting Review* pp. 389–401.

Anderson, E. W., E. Ghysels, and J. L. Juergens. 2009. The impact of risk and uncertainty on expected returns. *Journal of Financial Economics* 94:233–263.

Avramov, D., T. Chordia, G. Jostova, and A. Philipov. 2009. Dispersion in analysts' earnings forecasts and credit rating. *Journal of Financial Economics* 91:83–101.

Banerjee, A. V. 1992. A simple model of herd behavior. *The Quarterly Journal of Economics* pp. 797–817.

Barron, O. E., O. Kim, S. C. Lim, and D. E. Stevens. 1998. Using analysts' forecasts to measure properties of analysts' information environment. *Accounting Review* pp. 421–433.

Basu, S., and S. Markov. 2004. Loss function assumptions in rational expectations tests on financial analysts' earnings forecasts. *Journal of Accounting and Economics* 38:171–203.

Berkman, H., V. Dimitrov, P. C. Jain, P. D. Koch, and S. Tice. 2009. Sell on the news: Differences of opinion, short-sales constraints, and returns around earnings announcements. *Journal of Financial Economics* 92:376–399.

Botosan, C. A., M. A. Plumlee, and Y. Xie. 2004. The role of information precision in determining the cost of equity capital. *Review of Accounting Studies* 9:233–259.

Brown, L. D., A. C. Call, M. B. Clement, and N. Y. Sharp. 2015. Inside the "Black Box" of Sell-Side Financial Analysts. *Journal of Accounting Research* 53:1–47.

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* 49:359–401.

Clement, M. B. 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27:285–303.

Diamond, D. W., and R. E. Verrecchia. 1991. Disclosure, liquidity, and the cost of capital. *The Journal of Finance* 46:1325–1359.

Diether, K. B., C. J. Malloy, and A. Scherbina. 2002. Differences of opinion and the cross section of stock returns. *Journal of Finance* pp. 2113–2141.

Easley, D., and M. O'hara. 2004. Information and the cost of capital. *The journal of finance* 59:1553–1583.

Francis, J., and D. Philbrick. 1993. Analysts' decisions as products of a multi-task environment. *Journal of Accounting Research* pp. 216–230.

Garfinkel, J. A., and J. Sokobin. 2006. Volume, opinion divergence, and returns: A study of post-earnings announcement drift. *Journal of Accounting Research* pp. 85–112.

Gebhardt, W. R., C. M. Lee, and B. Swaminathan. 2001. Toward an implied cost of capital. *Journal of accounting research* 39:135–176.

Gerakos, J., and R. B. Gramacy. 2013. Regression-based earnings forecasts. *Chicago Booth Research Paper* .

Groysberg, B., P. M. Healy, and D. A. Maber. 2011. What Drives Sell-Side Analyst Compensation at High-Status Investment Banks? *Journal of Accounting Research* 49:969–1000.

Gu, Z., and J. S. Wu. 2003. Earnings skewness and analyst forecast bias. *Journal of Accounting and Economics* 35:5–29.

Hayes, R. M. 1998. The impact of trading commission incentives on analysts' stock coverage decisions and earnings forecasts. *Journal of Accounting Research* pp. 299–320.

Hayes, R. M., and C. B. Levine. 2000. An Approach to Adjusting Analysts' Consensus Forecasts for Selection Bias. *Contemporary Accounting Research* 17:61–83.

Hong, H., and J. D. Kubik. 2003. Analyzing the analysts: Career concerns and biased earnings forecasts. *Journal of Finance* pp. 313–351.

Hong, H., J. D. Kubik, and A. Solomon. 2000. Security analysts' career concerns and herding of earnings forecasts. *The Rand Journal of Economics* pp. 121–144.

Hou, K., M. A. Van Dijk, and Y. Zhang. 2012. The implied cost of capital: A new approach. *Journal of Accounting and Economics* 53:504–526.

Johnson, T. C. 2004. Forecast dispersion and the cross section of expected returns. *The Journal of Finance* 59:1957–1978.

Lim, T. 2001. Rationality and analysts' forecast bias. *The Journal of Finance* 56:369–385.

Lin, H.-W., and M. F. McNichols. 1998. Underwriting relationships, analysts' earnings forecasts and investment recommendations. *Journal of Accounting and Economics* 25:101–127.

Liu, X. G., and R. Natarajan. 2012. The effect of financial analysts' strategic behavior on analysts' forecast dispersion. *The Accounting Review* 87:2123–2149.

McNichols, M., and P. C. O'Brien. 1997. Self-selection and analyst coverage. *Journal of Accounting Research* pp. 167–199.

Mikhail, M. B., B. R. Walther, and R. H. Willis. 1997. Do security analysts improve their performance with experience? *Journal of Accounting Research* pp. 131–157.

Miller, E. M. 1977. Risk, uncertainty, and divergence of opinion. *The Journal of Finance* 32:1151–1168.

Sadka, R., and A. Scherbina. 2007. Analyst Disagreement, Mispricing, and Liquidity. *The Journal of Finance* 62:2367–2403.

Scharfstein, D. S., and J. C. Stein. 1990. Herd behavior and investment. *The American Economic Review* pp. 465–479.

Sheng, X., and M. Thevenot. 2012. A new measure of earnings forecast uncertainty. *Journal of Accounting and Economics* 53:21–33.

So, E. C. 2013. A new approach to predicting analyst forecast errors: Do investors overweight analyst forecasts? *Journal of Financial Economics* 108:615–640.

Xiao, Y. 2015. An Empirical Investigation of Analysts' Objective Function.

Zhang, X. 2006. Information uncertainty and stock returns. *The Journal of Finance* 61:105–137.

Figure I: Forecast Strategy

This figure plots the equilibrium forecasting strategies for the second analyst, described in Equations (7a) and (7b), which result from solving the analyst's objective function in Equation (6). In this objective function, λ weights the importance of differing from consensus forecast while the other parameters are σ (the uncertainty of the earnings prediction based on public information) and τ (the variance of the analyst's private signal). The horizontal axis is the mean of the second analyst's beliefs about earnings and satisfies (4). The left plot illustrates the forecasting function f_2^* for a fixed assumption about λ, σ and τ while the right plot illustrates the effect of increasing the magnitude of λ, σ and τ.

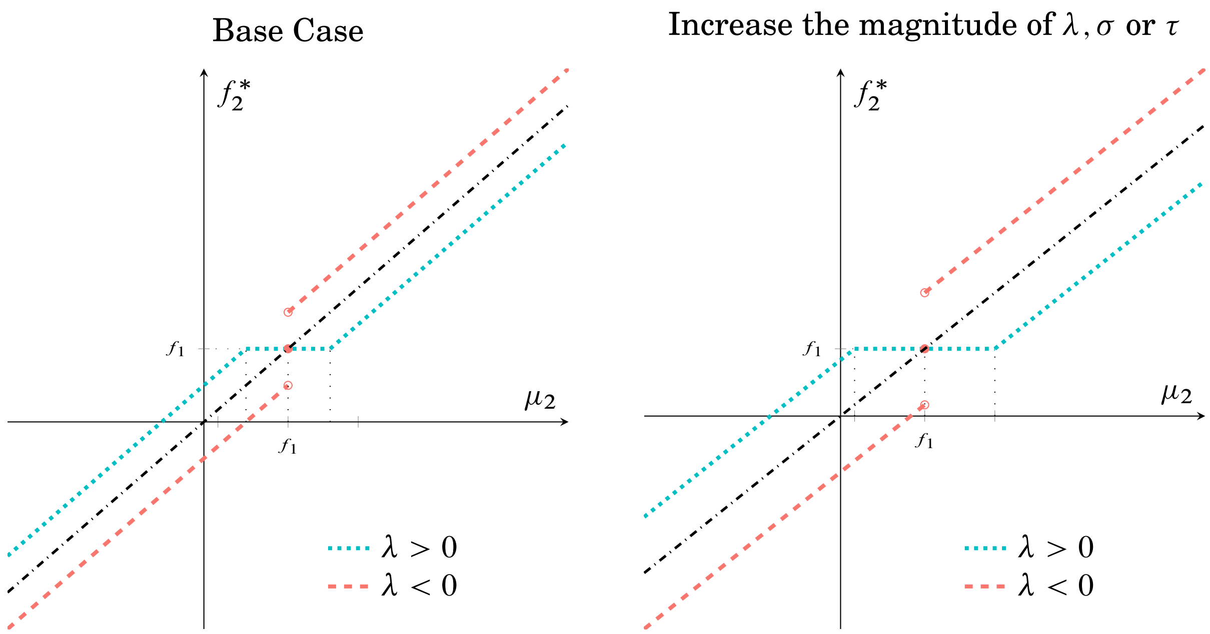

Table 1: Analyst Forecast Dispersion by Coverage and Order

This table reports average analyst forecast dispersion constructed and sorted using several methods. Panel A presents the average forecast dispersion by analyst coverage. Panel B summarizes the average forecast dispersion and forecast dispersion demeaned by the coverage size average after the j-th forecast, where demeaned dispersion has been normalized so that the average raw dispersion for $j = 2$ is the same as the demeaned variable.. In Panels A and B, the t-statistic corresponds to the difference in average dispersion in the row and that in the immediately preceding row and forecast dispersion is the standard deviation over the individual forecasts. Panel C sorts firm-years into terciles based on the magnitude of the lead analyst's forecast error ($|LEADFE|$) and reports the average forecast dispersion, computed as the standard deviation, between the 2nd and j-th forecast. The sample is comprised of 151,728 FY2 forecasts corresponding to 27,327 firm-years, with analyst coverage between 3 and 11, issued in the 30 day window after annual earnings announcements over the 1990-2012 period.

Panel A: Average Dispersion by Coverage

Coverage	Average Dispersion	t-statistic	Observations
3	0.143		6,452
4	0.146	(0.67)	5,127
5	0.158	(3.24)	3,783
6	0.166	(1.79)	3,108
7	0.168	(0.28)	2,401
8	0.173	(0.92)	2,016
9	0.174	(0.18)	1,578
10	0.166	(-1.15)	1,333
11	0.172	(0.73)	1,057
11 less 3	0.028	(4.79)	

Panel B: Average Dispersion After j-th Forecast

	Running Dispersion		Demeaned Dispersion	
j	Average	t-statistic	Average	t-statistic
2	0.129		0.129	
3	0.145	(8.68)	0.145	(8.68)
4	0.153	(4.47)	0.148	(2.05)
5	0.160	(3.41)	0.150	(0.68)
6	0.165	(1.80)	0.152	(0.91)
7	0.168	(0.95)	0.153	(0.35)
8	0.170	(0.58)	0.154	(0.23)
9	0.169	(-0.26)	0.154	(0.03)
10	0.166	(-0.63)	0.155	(0.16)
11	0.169	(0.46)	0.155	(0.01)
11 less 2	0.040	(6.65)	0.026	(4.31)

Panel C: Dispersion after j-th forecast by Terciles of $|LEADFE|$

j	1 (Low $|LEADFE|$)	2 (Mid $|LEADFE|$)	3 (High $|LEADFE|$)
3	0.055	0.085	0.173
4	0.070	0.110	0.223
5	0.079	0.122	0.245
6	0.085	0.129	0.258
7	0.087	0.133	0.267
8	0.089	0.134	0.276
9	0.089	0.142	0.275
10	0.092	0.140	0.267
11	0.093	0.146	0.273
Mean $|LEADFE|$	0.084	0.366	1.541

Table 2: Cross-Sectional Determinants of Analyst Forecast Dispersion

The table reports the regression coefficients from a regression of the log of forecast dispersion or the magnitude of the lead forecast error on assets per share($ASSETS$), the book to market ratio (BTM), analyst coverage (COV), an indicator for loss firms ($LOSS$), the log of share price (PRC), the log of the market value of equity ($SIZE$), R&D expense scaled by sales (RD), an indicator for missing R&D data ($RDMISS$), sales scaled by total assets ($SALES$), the magnitude of accruals per share ($|ACCR|$), the Beta decile assignment ($BETADEC$), the leverage ratio (LEV) and the log standard deviation of earnings per share over the previous eight quarters ($PASTVOL$). Industry fixed effects and calendar time-effects are included in the regressions. Standard errors clustered by industry and year are shown in parentheses. The constant estimate in the regression is not reported. The sample is constructed from two-year ahead forecasts issued within 30 days of annual earnings announcements between 1990 and 2012. All explanatory variables are measured as of the fiscal year-end pertaining to the earnings announcement. The data includes 19,560 firm-years randomly sampled from 21,787 covered by between 3 and 11 analysts. ***,** and * denote significance at the 0.01, 0.05 and 0.10 levels (two-tailed test), respectively.

	Dependent variable:	
	log($DISP$)	log($\|LEADFE\|$)
	(1)	(2)
$ASSETS$	0.004***	0.003***
	(0.0003)	(0.0003)
BTM	0.144***	0.112***
	(0.017)	(0.019)
PRC	−0.001	0.116***
	(0.013)	(0.014)
COV	0.038***	0.036***
	(0.003)	(0.004)
$SIZE$	−0.015**	−0.088***
	(0.007)	(0.008)
RD	1.497***	0.950***
	(0.103)	(0.115)
$RDMISS$	0.193***	0.146***
	(0.018)	(0.020)
$SALES$	−0.089***	−0.060***
	(0.011)	(0.012)
$\|ACCR\|$	0.005*	0.002
	(0.003)	(0.003)
$LOSS$	0.290***	0.152***
	(0.021)	(0.023)
$BETADEC$	−0.024***	−0.037***
	(0.003)	(0.003)
$LEV \times 10^{-3}$	−0.009	0.004
	(0.018)	(0.019)
$PASTVOL$	0.194***	0.197***
	(0.007)	(0.008)
R^2	0.340	0.226

Table 3: Model Predictions about Forecast Dispersion and Forecast Error

This table presents the expected forecast dispersion and forecast error for the lead analyst resulting from the model equilibrium outlined in Section 4 under various assumptions about parameters. The quantities reported in Panel A are based on the assumption that analysts honestly report their beliefs and those reported in Panel B are based on the assumption there is strategic behavior arising from the objective function described in Section 4.2. μ_0 and σ are the mean and standard deviation of the earnings prediction based on public information. τ is the standard deviation in analysts' private signal. The parameter λ weights the importance of deviation from consensus relative to forecast error in determining analyst's payoffs.

Panel A: Model Predictions under Honest Forecasting Assumption

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Parameter Assumptions							
τ	1.1	1.41	1.1	1.64	1.1	1.41	1.1
σ	0.35	0.4	0.56	0.7	0.4	0.35	0.35
$\lvert\mu_0 - A\rvert$	0.6	0.6	0.8	0.8	0.6	0.6	0.8
Model Implied Expected Dispersion After j-th Forecast							
j=3	0.212	0.210	0.173	0.173	0.216	0.207	0.129
j=4	0.207	0.206	0.190	0.192	0.214	0.200	0.133
j=5	0.194	0.194	0.190	0.194	0.203	0.187	0.128
j=6	0.183	0.183	0.186	0.191	0.192	0.175	0.123
j=7	0.173	0.173	0.181	0.187	0.182	0.165	0.118
j=8	0.165	0.165	0.175	0.182	0.174	0.157	0.114
j=9	0.158	0.158	0.170	0.177	0.167	0.150	0.110
j=10	0.151	0.152	0.166	0.173	0.161	0.144	0.106
j=11	0.146	0.146	0.161	0.168	0.155	0.138	0.103
Model Implied Expected Forecast Error for Lead Analyst							
$\mathbb{E}_{\varepsilon_1}(\lvert f_1^* - A\rvert)$	0.545	0.556	0.637	0.677	0.530	0.566	0.727

Panel B: Predictions Assuming Objective Function in Equation (6)

	(1)	(2)	(3)	(4)	(5)	(6)
Parameter Assumptions						
τ	1.1	1.1	1.1	1.1	1.1	1.1
σ	0.6	0.9	0.6	0.9	0.6	0.9
$\lvert\mu_0 - A\rvert$	0.1	0.62	0.1	0.62	0.1	0.62
λ	0.1	0.1	0.7	0.7	-0.7	-0.7
Model Implied Expected Dispersion After j-th Forecast						
j=3	0.052	0.083	0.001	0.005	0.448	0.525
j=4	0.071	0.110	0.001	0.008	0.474	0.575
j=5	0.081	0.124	0.002	0.010	0.475	0.576
j=6	0.088	0.132	0.002	0.011	0.470	0.567
j=7	0.093	0.137	0.002	0.012	0.464	0.557
j=8	0.096	0.141	0.002	0.013	0.458	0.546
j=9	0.098	0.143	0.002	0.013	0.451	0.535
j=10	0.100	0.145	0.002	0.014	0.445	0.525
j=11	0.101	0.146	0.002	0.014	0.438	0.515
Model Implied Expected Forecast Error for Lead Analyst						
$\mathbb{E}_{\varepsilon_1}(\lvert f_1^* - A\rvert)$	0.071	0.371	0.071	0.371	0.071	0.371

Table 4: Parameter Estimates

This table reports the Simulated Method of Moments estimates for the empirical model described in Section 5. σ is the uncertainty of the earnings prediction based on public information. τ is the standard deviation in analysts' private signal. The parameter λ weights the importance of deviation from consensus relative to forecast error in determining analyst's payoffs. NH refers to the No Heterogeneity Model where all firms and analysts are assumed to be identical. FH refers to the Firm-Heterogeneity Model where there is cross-sectional variation in σ_{it}. AH refers to the Analyst Heterogeneity model where there is cross-sectional variation in τ_{jt}. CH is the Complete Heterogeneity Model where there is cross-sectional variation in both dimensions. In the NH and AH models, I use equal-weighted moments. For the FH and CH models, I weight each of the dispersion moments by the length of the sample vector. $LEADFE$ is the first analyst's forecast error for the firm-year, $expr$ is analyst experience, in years, and $employersz$ is the log of total number of analysts in the same research firm who issued a forecast in the preceding calendar year. The $[\cdot]^+$ notation used in reporting the estimate of σ and τ indicates that whenever the quantity inside the square brackets is non-positive, the estimated value is 10^{-4}. Standard errors, where applicable, are reported in parentheses under each point estimate. The sample is constructed from two-year ahead forecasts issued within 30 days of annual earnings announcements between 1990 and 2012. The data includes 19,560 firm-years randomly sampled from 21,787 firm-years covered by between 3 and 11 analysts. The row labeled J-Test includes the test statistic for the J-Test for overidentifying restrictions. The p-value is based on the comparison of the J-Test statistic to the critical values for the χ^2 distribution.

Panel A: NH and AH Specifications

	(1) NH	(2) AH
λ	0.288 (0.01)	0.287 (0.02)
σ_{it}	1.21 (0.09)	1.24 (0.16)
τ_{ijt}	1.05 (0.05)	$\sqrt{\left[\underset{(0.22)}{2.25} - \underset{(0.02)}{0.18} expr_{jt} + \underset{(0.002)}{0.004} expr_{jt}^2 + \underset{(0.001)}{0.001} employersz_{jt}\right]^+}$
J-Test	0.54	0.47
p-value	>0.10	>0.10

Panel B: FH and CH Specifications

	(3) FH	(4) CH
λ	0.087 (0.001)	0.200 (0.003)
σ_{it}	$\frac{\underset{(0.04)}{1.43} \lvert LEADFE \rvert_{it}}{\sqrt{\left[\underset{(0.04)}{1.33} - LEADFE_{it}^2\right]^+}}$	$\frac{\sqrt{\left[\underset{(0.05)}{2.9} - \underset{(0.02)}{0.49} expr_{i1t} + \underset{(0.002)}{0.18} expr_{i1t}^2 - \underset{(0.003)}{0.36} employersz_{i1t}\right]^+} \times \lvert LEADFE \rvert_{it}}{\left(\left[\underset{(0.05)}{1.85} - \underset{(0.12)}{0.32} expr_{i1t} + \underset{(0.002)}{0.11} expr_{i1t}^2 - \underset{(0.003)}{0.23} employersz_{i1t} - LEADFE_{it}^2\right]^+\right)^{\frac{1}{2}}}$
τ_{ijt}	1.43 (0.04)	$\sqrt{\left[\underset{(0.05)}{2.9} - \underset{(0.12)}{0.49} expr_{jt} + \underset{(0.002)}{0.18} expr_{jt}^2 + \underset{(0.003)}{0.36} employersz_{jt}\right]^+}$
J-Test	16.7	2.65
p-value	0.02	>0.10

Table 5: Comparison of Model-Based and Descriptive Approaches

This table reports the predictive performance, measured in the square root of the mean forecast error (RMSE), for the four models I describe in Section 5 and the descriptive based approaches. For each of the No Heterogeneity, Firm Heterogeneity, Analyst Heterogeneity and Complete Heterogeneity models, I use the parameter estimates reported in Table 4 to compute the model-implied sequence of dispersion. In the full determinants regression approach, I apply the coefficients from estimating Equation (1), which are reported in Table 2, to construct a prediction about earnings. I also form a prediction using a regression of log forecast dispersion on only the unsigned consensus forecast error. The last prediction is formed by augmenting the regression in Equation (1) with the consensus forecast error. The RMSE is determined using the difference between actual dispersion and the model-predicted dispersion. In both cases, the sample is comprised of 2,227 randomly selected to be excluded from both the estimation samples.

	# of Estimated Parameters	RMSE
Model-Based Approach		
No Heterogeneity Model	3	0.164
Firm Heterogeneity Model	2	0.230
Analyst Heterogeneity Model	7	0.153
Complete Heterogeneity Model	6	0.217
Descriptive Approach		
Full determinants regression from Table 2	78	0.228
Regression with consensus forecast error	2	0.237
Full determinants regression with consensus forecast error	79	0.219

Table 6: Predicted Forecast Dispersion Under Counterfactuals Policies

This table reports the average forecast dispersion in the observed data after the j-th analyst as well as model-predictions under various counterfactual forecast policies. In row 1, I assume that $\lambda = 0$ in Equation (6). In row 2, I assume that $\lambda = -0.288$, which is the negative of the estimated value of λ obtained in Table 4. In row 3(4), I assume that $\sigma = 2.42$ ($\sigma = 0.60$) which is twice (half) the estimated value of σ obtained in Table 4. In row 5(6), I assume that $\tau = 2.09$ ($\tau = 0.52$) which is twice (half) the estimated value of τ obtained in Table 4. In row 7, I assume a different model in which analysts do not observe previous analysts' forecasts in forming their beliefs about earnings. Bootstrapped standard errors are reported in parentheses below each estimated prediction. All counterfactual policies use the estimates from Table 4, with homogeneous σ and τ, as the remaining parameter inputs. The sample is comprised of 21,787 firms with between 3 and 11 two-year ahead forecasts in the 30 days following annual earnings announcements between 1990 and 2012. Bootstrapped standard errors are tabulated in the parentheses below the point estimates.

	Average Dispersion after the j-th forecast								
	j=3	j=4	j=5	j=6	j=7	j=8	j=9	j=10	j=11
Observed	0.128	0.145	0.155	0.161	0.165	0.167	0.167	0.164	0.168
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.006)
(1) Remove analysts' peer incentives	0.446	0.414	0.382	0.355	0.332	0.315	0.300	0.287	0.276
	(0.005)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
(2) Reverse sign of peer incentives	0.316	0.341	0.346	0.342	0.336	0.330	0.325	0.320	0.315
	(0.005)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
(3) Reduce precision of EA news by half	0.133	0.151	0.161	0.167	0.171	0.174	0.175	0.175	0.176
	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
(4) Double precision of EA news	0.114	0.134	0.146	0.154	0.160	0.164	0.167	0.168	0.170
	(0.004)	(0.003)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
(5) Reduce precision of private signal by half	0.119	0.143	0.158	0.168	0.175	0.181	0.185	0.187	0.189
	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
(6) Double precision of private signal	0.122	0.141	0.152	0.158	0.163	0.165	0.167	0.168	0.168
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
(7) Independent forecasting	0.337	0.385	0.418	0.431	0.437	0.445	0.455	0.462	0.470
	(0.026)	(0.029)	(0.031)	(0.031)	(0.031)	(0.032)	(0.032)	(0.032)	(0.033)

Table 7: Realized Return Tests Based on Dispersion and Model Estimates

This table presents the monthly raw and market-adjusted returns, expressed in percentages, by quintiles of dispersion. The sample is comprised of 21,787 firms with between 3 and 11 two-year ahead forecasts in the 30 days following annual earnings announcements between 1990 and 2012. RET6, RET12 and RET24 (ARET6, ARET12, ARET24) are average raw (market-adjusted) returns measured over 6, 12 and 24 months starting from the month after the end of the 30 day post-earnings announcement window. The market adjustment is computed by subtracting the CRSP value weighted return from the raw returns. In Panel A, firms are sorted into quintiles based on analyst forecast dispersion $DISP$. Panel B (Panel C) sorts firms into quintiles based on σ^{FH} (σ^{CH}), which is the estimate of earnings uncertainty from the Firm Heterogeneity (Complete Heterogeneity) model described in Section 6.1 and Table 4. The expected forecast error, determined using the cross-section earnings regression approach developed in Hou et al. (2012), is used in place of actual forecast error to compute the model estimates.

Panel A: Monthly returns by quintiles of $DISP$

	Raw returns (%)			Market-adjusted returns (%)		
	RET6	RET12	RET24	ARET6	ARET12	ARET24
1 (Low $DISP$)	1.83	1.74	1.61	1.08	0.94	0.80
2	1.27	1.22	1.26	0.52	0.42	0.44
3	1.25	1.21	1.14	0.51	0.40	0.32
4	1.03	0.95	1.06	0.30	0.15	0.24
5 (High $DISP$)	0.83	0.84	0.94	0.08	0.04	0.12
High less Low	-1.00	-0.90	-0.68	-1.00	-0.90	-0.68
t-statistic (High=Low)	(-8.54)	(-10.30)	(-11.05)	(-9.41)	(-11.47)	(-11.91)

Panel B: Monthly returns by quintiles of σ^{FH}

	Raw returns (%)			Market-adjusted returns (%)		
	RET6	RET12	RET24	ARET6	ARET12	ARET24
1 (Low σ^{FH})	1.55	1.49	1.46	0.81	0.69	0.64
2	1.17	1.21	1.20	0.44	0.41	0.39
3	1.17	1.07	1.15	0.42	0.26	0.33
4	1.12	1.04	1.13	0.38	0.24	0.31
5 (High σ^{FH})	1.20	1.13	1.08	0.46	0.33	0.26
High less Low	-0.35	-0.36	-0.38	-0.35	-0.36	-0.38
t-statistic (High=Low)	(-2.88)	(-4.08)	(6.27)	(-3.23)	(-4.62)	(-6.88)

Panel C: Monthly returns by quintiles of σ^{CH}

	Raw returns (%)			Market-adjusted returns (%)		
	RET6	RET12	RET24	ARET6	ARET12	ARET24
1 (Low σ^{CH})	1.39	1.28	1.28	0.65	0.48	0.46
2	1.10	1.04	1.09	0.36	0.24	0.27
3	1.13	1.13	1.18	0.38	0.33	0.36
4	1.22	1.14	1.14	0.48	0.34	0.32
5 (High σ^{CH})	1.37	1.36	1.33	0.64	0.56	0.52
High less Low	-0.02	0.08	0.05	-0.01	0.08	0.06
t-statistic (High=Low)	(-0.16)	(0.90)	(0.91)	(-0.09)	(1.01)	(-1.00)

Table 8: Estimates from Model which Assumes Analysts Have Different Objective Functions

This table reports the Simulated Method of Moments estimates for the empirical model described in Section 5 with the modification that λ, the parameter which weights the importance of deviation from consensus relative to forecast error in determining analyst's payoffs, varies based on the size of the analyst's employer research firm. The value of the parameter corresponding to each research firm size tercile is denoted as λ_{small}, λ_{medium} and λ_{large}. The remaining parameters are σ, the uncertainty of the earnings prediction based on public information and τ, which is the standard deviation in analysts' private signal. In column (1), NH refers to the No Heterogeneity Model where all firms and analysts are assumed to be identical and the AH model in column (2) assumes that the parameter τ varies across individual analysts. $expr$ is analyst experience, in years, and $employersz$ is the log of total number of analysts in the same research firm who issued a forecast in the preceding calendar year. The $[\cdot]^+$ notation used in reporting the estimate of τ indicates that whenever the quantity inside the square brackets is non-positive, the estimated value is 10^{-4}. Standard errors, where applicable, are reported in parentheses under each point estimate. The sample is constructed from two-year ahead forecasts issued within 30 days of annual earnings announcements between 1990 and 2012. The data includes 19,560 firm-years randomly sampled from 21,787 firm-years covered by between 3 and 11 analysts. The row labeled J-Test includes the test statistic for the J-Test for overidentifying restrictions. The p-value is based on the comparison of the J-Test statistic to the critical values for the χ^2 distribution.

	(1) NH	(2) AH
λ_{small}	0.777 (0.13)	0.045 (0.20)
λ_{medium}	0.232 (0.07)	0.735 (0.26)
λ_{large}	0.111 (0.05)	0.470 (0.28)
σ_{it}	1.16 (0.09)	1.41 (0.22)
τ_{ijt}	1.08 (0.09)	$\sqrt{[\underset{(0.48)}{1.63} - \underset{(0.05)}{0.157} \times expr_{jt} + \underset{(0.003)}{0.006} \times expr_{jt}^2 - \underset{(0.001)}{0.002} \times employersz_{jt}]^+}$
J-Test	0.30	0.27
p-value	>0.10	>0.10

Table 9: Estimates Based on Variance Moments

Panel A reports the observed and model-implied average forecast variance after the j-th analyst has forecasted. The sample variance is computed based on the sum of squared deviation from sample mean and averaged over j-1. The model-implied average variance is determined using the estimates from the NH Model reported in Table 4. Standard errors are obtained from computing the model-predicted average sample variance using each simulated draw of the unobservable term and then averaging over 1,000 simulations. In Panel B, I report the parameter estimates from the model described in Section 5, but with the substitution of moment restrictions based on sample variance discussed in 6.6. Refer to Table 4 for a description of the model parameters and the NH, AH, FH and CH specifications. $LEADFE$ is the first analyst's forecast error for the firm-year, $expr$ is analyst experience, in years, and $employersz$ is the log of total number of analysts in the same research firm who issued a forecast in the preceding calendar year. The $[\cdot]^+$ notation used in reporting the estimate of σ and τ indicates that whenever the quantity inside the square brackets is non-positive, the estimated value is 10^{-4}. Standard errors, where applicable, are reported in parentheses under each point estimate. The sample is constructed from two-year ahead forecasts issued within 30 days of annual earnings announcements between 1990 and 2012. The data includes 19,560 firm-years randomly sampled from 21,787 firm-years covered by between 3 and 11 analysts.

Panel A: Comparison of Average Observed and Model Predicted Forecast Variance After the j-th Forecast

j=	Observed	Model	Std. Error
3	0.053	0.047	0.002
4	0.057	0.051	0.001
5	0.059	0.054	0.001
6	0.062	0.058	0.001
7	0.062	0.060	0.001
8	0.063	0.064	0.001
9	0.062	0.065	0.001
10	0.068	0.069	0.001
11	0.061	0.071	0.001

Panel B: Estimates Based on Variance Moment Restrictions

	(1) NH	(2) AH
λ	0.306	0.251
σ_{it}	1.26	1.30
τ_{jt}	1.02	$\sqrt{[2.73 - 0.21expr + 0.003expr^2 + 0.001employersz]^+}$

Panel B (Continued): Estimates Based on Variance Moment Restrictions

	(3) FH	(4) CH
λ	0.101	0.262
σ_{it}	$\frac{1.53\|LEADFE\|_{it}}{\sqrt{[1.49-LEADFE_{it}^2]^+}}$	$\frac{\sqrt{[2.92-0.38expr_{i1t}+0.18expr_{i1t}^2-0.26employersz_{i1t}]^+}\times\|LEADFE\|_{it}}{\left([1.86-0.24expr_{i1t}+0.11expr_{i1t}^2-0.17employersz_{i1t}-LEADFE_{it}^2]^+\right)^{\frac{1}{2}}}$
τ_{ijt}	1.53	$\sqrt{\left[2.92 - 0.38expr_{jt} + 0.18expr_{jt}^2 + 0.26employersz_{jt}\right]^+}$