George Li (USA), Donglin Li (USA), Ming Li (USA)

Information quality and analyst forecast accuracy

Abstract

This paper presents a stochastic model of earnings to study and test how the precision of information that analysts have about the unobservable expected earnings growth rates of firms affects the accuracy of analyst earnings forecasts. The article develops a maximum likelihood procedure to estimate the precision of information that analysts have about expected earnings growth rates. Using the I/B/E/S and COMPUSTAT data sets, the authors find that earnings forecast accuracy is positively associated with information precision. This empirical finding helps explain why analyst earnings forecasts are more accurate than the forecasts that are based only on historical earnings data. It implies that investor may be better off relying on information precision rather than the size of a brokerage house to choose forecasting services.

Keywords: information precision estimation, information precision, analyst earnings forecast accuracy, continuous-time model and bayesian learning.
JEL Classification: G12, G17, G14, C11.

Introduction

The accuracy of analyst earnings forecasts has been long an interesting topic to both financial economists and investors. There is a large body of literature studying how firm characteristics affect the accuracy of earnings forecasts (see, for example, Sinha, Brown and Das 1997, Brown and Rozeff 1980; and Richards 1976). However, despite the extensive study of analyst forecast accuracy, little is known about how the precision of information that analysts have about the unobservable expected earnings growth rates of firms affects the accuracy of analyst earnings forecasts. For example, if analysts have more precise information about unobservable future earnings growth rates, do earnings forecasts become more accurate? The answer helps us resolve one misconception in Wall Street: a large brokerage house or a big-name investor necessarily provides better forecasts than a small one.

These issues have become more relevant in light of the recent financial crisis in the U.S. Although financial analysts' earnings forecasts are the key determinant of stock prices, the effect of them on asset prices are murky at best. For example, several financial analysts, including the famous George Soros and Meredith Whitney, warned unsustainable bubbles in housing markets back in 2004 and 2005. Apparently few investors heeded this message seriously until 2008. What was the reason that the market ignored these warning signs? One would postulate that these warnings were just like a typical noise that overwhelmed in the market, too noisy to be noted. The main issue is that people don't have a ready tool to measure the informativeness of analyst research. The key to a credible signal is the precision of these warnings.

In this paper, we study and test how the precision of information that analysts have about the unobservable expected earnings growth rates of firms affects analyst forecast accuracy. We present a stochastic model of corporate earnings with a time-varying expected growth rate that is unobservable to analysts. To forecast future earnings, analysts have to use historical earnings data and the information they have about the unobservable expected earnings growth rate of a firm to estimate the value of the expected earnings growth rate. Thus, our model captures the notion that analysts use more than historical earnings data to forecasts future earnings.

We use biased earnings forecasts to construct unbiased earnings forecasts, which are then used to estimate the precision of information. This approach is based on the idea that unbiased earnings forecasts are analysts' expectations of future earnings conditional on analysts' information about unobservable expected earnings growth rates. Thus, by using unbiased forecasts, we can extract analysts' information about the expected earnings growth rates of firms. This approach is consistent with the practice in accounting and finance that unbiased earnings forecasts are widely used as analysts' expectation of future earnings (see, for example, Brown, Foster and Noreen 1985, Hughes and Ricks 1987; McNichols 1989; Landsman and Maydew 2002; Frankel et al., 2006).

We develop a maximum likelihood procedure to estimate information precision. Our estimation exercise shows that the precision of information that analysts have about the expected earnings growth rates of firms varies from firm to firm. For some firms, analysts have relatively very precise information; for some firms, the information that analysts have about future earnings growth is relatively noisy.

We use data from I/B/E/S and COMPUSTAT to cross-sectionally analyze how analysts' information precision affects forecast accuracy. By controlling for firm characteristics such as earnings volatility, firm size and earnings level, we find that earnings forecast accuracy is positively associated with the precision of information that analysts have about the expected earnings growth rates of firms. This finding indicates that when analysts have more precise information about expected earnings growth rates, their earnings forecasts are more accurate. This empirical finding helps explain why analysts' earnings forecasts are more accurate than the earnings forecasts based only on past realized earnings (see, for example, Brown, Foster and Noreen, 1985), and suggests that to forecast future earnings more accurately, analysts may have to approach management, use better resources, or employ sharper skills to obtain more precise information about future earnings growth rates.

Our work is related to prior studies that use analyst forecasts to infer analysts' information characteristics. A partial list of works in this area includes Barry and Jennings (1992), Abarbanell, Lanen and Verrecchia (1995), Barron, Kim, Lim and Stevens (1998) and Landsman and Maydew (2002), Frankel et al. (2006). Despite similarity, our work differs from them in the following three aspects. First, this paper presents a stochastic model of corporate earnings with an unobservable time-varying expected growth rate, which must be estimated by using both historical earnings data and the information that analysts have about the expected earnings growth rate. Second, the paper also develops a maximum likelihood procedure to estimate information precision. Finally, the paper documents the empirical evidence that a higher precision of information that analysts have about expected earnings growth rates increases earnings forecast accuracy.

The rest of the paper is organized as follows. Section 1 presents a stochastic model of corporate earnings with an unobservable expected growth rate. Section 2 presents a maximum likelihood procedure to estimate the precision of information that analysts have about the expected earnings growth rates of firms. Section 3 discusses the data used in this study. The empirical results and their discussion are presented in section 4. Finally, in the last section, we make conclusions.

1. The model

In this section, we present a simple continuous-time model of corporate earnings to discuss how to estimate the precision of information that analysts have about the expected earnings growth rate of a firm. While a similar discrete-time model may be used to achieve the same purpose, the continuous-time model here helps provide a much simpler solution procedure.

Consider an earnings process, $X(t)$, which evolves as follows:

$$dX = \mu dt + \sigma_X dW_X, \tag{1}$$

where $\mu(t)$ is the expected earnings growth rate at time t and is unobservable, σ_X is the volatility of the earnings and assumed to be a constant, and $W_X(t)$ is a standard Brownian motion. Moreover, the expected earnings growth rate $\mu(t)$ is time-varying and evolves as follows:

$$d\mu = k(\bar{\mu} - \mu)dt + \sigma_\mu dW_\mu, \tag{2}$$

where $\sigma\mu$ is the volatility of the expected earnings growth rate and assumed to be a constant, k the mean-reverting speed parameter, $\bar{\mu}$ the long-run mean of the expected growth rate, and $W_\mu(t)$ a standard Brownian motion, correlated with $W_X(t)$. In the rest of the paper, we let $\sigma_{\mu X} = \rho_{\mu X}\sigma_\mu\sigma_X$ be the covariance between the expected earnings growth rate and the earnings, where $\rho_{\mu X}$ is the correlation between $W_\mu(t)$ and $W_X(t)$.

The consideration of a mean-reverting expected earnings growth rate in equation (2) captures the notion that in the real world, the expected earnings growth rate of a firm is not a constant but time-varying and related to business cycles (see Kandel and Stambaugh, 1990). Previous authors such as Wang (1993), Veronesi (2000) and Brennan and Xia (2001) also model the expected growth rate of dividends as a mean-reverting process, similar to what we do here.

While analysts cannot observe $\mu(t)$, they are assumed to know the process and its parameters for the true expected earnings growth rate. This assumption is a standard way of achieving analytical tractability in the literature of learning in the financial market (see, for example, Wang 1993, 1994; Brennan and Xia, 2001). In addition, analysts are assumed to obtain a noisy signal as follows:

$$dI = \mu dt + \sigma_I dW_I, \tag{3}$$

where σ_I is the volatility of this signal and assumed to be a constant, and $W_I(t)$ is a standard Brownian motion, which, for simplicity, is assumed to be independent of other Brownian motions. While a general correlation structure among Brownian motions can be considered, it will yield similar results.

The volatility σ_I decides on the precision of information that analysts have. When σ_I is large, information is relatively noisy; when σ_I is small, information is relatively precise. At one extreme, when $\sigma_I = 0$, analysts have perfect information about expected earnings growth rates. At the other extreme, when $\sigma_I \to \infty$, the signal conveys no information and analysts use just historical earnings data to learn about the expected earnings growth rate.

The signal in equation (3) is the continuous time analog of the standard signal $I(t)=\mu(t)+\sigma_I\varepsilon(t)$ in a discrete time model, in which the signal equals fundamentals plus noise, with $\varepsilon(t)$ a standard normal (see Veronesi (2000) for using the same way to model the noisy public signal in a continuous time model).

Since analysts cannot observe the expected earnings growth rate, to forecast future earnings, they have to estimate the value of $\mu(t)$ from information $I(t)$, and his observation of $X(t)$. As shown in Liptser and Shiryayev (1978), the conditional distribution of $\mu(t)$ based on analysts' information set $F_t=\{X(s),I(s),s\le t\}$ at time t is also normal, and the mean $m(t)$ of this conditional distribution evolves according to the following diffusion process, which is derived in the Appendix. The result is summarized in the following lemma.

Lemma 1. *Let* $m(t)=E[\mu(t)\,|\,F_t]$ *be the estimate of the expected earnings growth rate. Then m(t) satisfies the following stochastic differential equation:*

$$dm=k(\bar{\mu}-m)dt+a_1d\tilde{W}_I+a_2d\tilde{W}_X, \tag{4}$$

$$d\tilde{W}_X=\frac{1}{\sigma_X}\left[dX-mdt\right],$$

$$d\tilde{W}_I=\frac{1}{\sigma_I}\left[dI-mdt\right],$$

where a_1 and a_2 are constants, defined in the Appendix. The innovation processes $\tilde{W}_I$ and $\tilde{W}_X$ are standard Brownian motions with respect to $F_t\equiv F^{X,I}(t)$. In fact, the information structure generated by $F^{\tilde{W}}(t)$ is equivalent to that generated by $F^{X,I}(t)$, where $\tilde{W}=[\tilde{W}_I,\tilde{W}_X]^T$.

In equation (4), the estimate of the expected earnings growth rate follows a mean-reverting two-dimension process with a constant volatility. Two Brownian motions, $d\tilde{W}_I$ and $d\tilde{W}_X$, respectively, the normalized innovation processes of the signal and earnings realizations. These two stochastic components convey new information about surprises in signals and earnings. For example, when there is an unexpected high signal $d\tilde{W}_I>0$, the analyst increases the expectation of $\mu(t)$.

When the estimate of the expected earnings growth rate at time t is $m(t)$, earnings evolve as follows:

$$dX=mdt+\sigma_X d\tilde{W}_X. \tag{5}$$

Also, equation (4) can be simplified as:

$$dm(t)=k(\bar{\mu}-m)dt+\sigma_m d\tilde{W}_m, \tag{6}$$

where $\sigma_m^2=a_1^2+a_2^2$ and $\tilde{W}_m(t)$ is a standard Brownian motion. In the following, we let $\sigma_{mX} = \rho_{mX}\sigma_m\sigma_X$ be the covariance between the estimate of the expected earnings growth rate and the earnings growth rate, where ρ_{mX} is the correlation between $\tilde{W}_m(t)$ and $\tilde{W}_X(t)$.

2. Estimation of information precision

In this section, we develop a maximum likelihood procedure to estimate the precision of information that analysts have about the expected earnings growth rate of a firm. We first examine how analysts use the estimate of the unobservable expected earnings growth rate at time, t, $m(t)$, to forecast future earnings $X(s)$, $s > t$. Let $UFS(s)=E\left[X(s)\,|\,F_t\right]$, $s > t$, be the unbiased earnings forecast at time t. Then the following lemma summarizes the relationship between the unbiased analyst forecast of future earnings $X(s)$, $UFS(s)$, $s > t$, and the estimate of the expected earnings growth rate, $m(t)$.

Lemma 2. *Let* $UFS(s)=E\left[X(s)\,|\,F_t\right]$ *be the unbiased forecast of earnings* $X(s)$, $s>t$. *Then, we have*

$$UFS(s)=X(t)+\bar{\mu}(s-t)+\frac{(m(t)-\bar{\mu})}{\kappa}\left(1-e^{-\kappa(s-t)}\right)$$

where $\bar{\mu}$ is the long-run mean of the expected earnings growth rate.

Clearly, analysts use more than historical earnings data to forecast future earnings, since $m(t)$ in $UFS(s)$ is the estimation of the unobservable expected growth rate of earnings conditional on the analysts' information set, which includes historical earnings data and information.

To use discrete-time data of analyst earnings forecasts to estimate the precision of information, we first derive the discrete-time versions of equations (5) and (6) as follows:

$$X(t)=X(t-1)+\alpha_1+\varphi_1 m(t-1)+\varepsilon_X(t), \tag{7}$$

$$m(t)=\alpha_2+\varphi_2 m(t-1)+\varepsilon_m(t)\ , \qquad (8)$$

where

$$\alpha_1=\bar{\mu}\left(1-\frac{1-e^{-k}}{\kappa}\right),\ \alpha_2=\bar{\mu}(1-e^{-k}),$$

$$\varphi_1=\frac{1-e^{-k}}{\kappa},\ \varphi_2=e^{-k},$$

$$\varepsilon_X(t)=\int_{t-1}^{t}\sigma_X d\tilde{W}_X(\tau)+\frac{\sigma_m}{\kappa}\int_{t-1}^{t}(1-e^{-k(t-\tau)})d\tilde{W}_m(\tau),$$

$$\varepsilon_m(t)=\sigma_m\int_{t-1}^{t}e^{-k(t-\tau)}d\tilde{W}_m(\tau).$$

Since $\varepsilon_X(t)$ and $\varepsilon_m(t)$ both are functions of Brownian motion, they have a joint normal distribution. Then, using equations (7) and (8), we have the likelihood function as follows:

$$L(\sigma_X,\sigma_\mu,k,\rho_{\mu X},\bar{\mu},\sigma_I)=\prod_{t=1}^{n}f(\sigma_1,\sigma_2,\rho_{12})\exp(Q(t))\ , \qquad (9)$$

$$f(\sigma_1,\sigma_2,\rho_{12})=\left(2\pi\sigma_1\sigma_2\sqrt{1-\rho_{12}^2}\right)^{-1},$$

$$Q(t)=-\frac{1}{2\left(1-\rho_{12}^2\right)}\Lambda(t),$$

where σ_1 is the standard deviation of $\varepsilon_X(t)$ and σ_2 is the standard deviation of $\varepsilon_m(t)$ and ρ_{12} is their correlation. We prove in Lemma 3 and Lemma 4 that σ_1, σ_2 and ρ_{12} themselves depend on the original model parameters. And n is the number of observations in the data sample,

$$\Lambda(t)=\left(\frac{X(t)-\bar{X}(t)}{\sigma_1}\right)^2-2\rho_{12}\frac{X(t)-\bar{X}(t)}{\sigma_1}\frac{m(t)-\bar{m}(t)}{\sigma_2}+\left(\frac{m(t)-\bar{m}(t)}{\sigma_1}\right)^2,$$

$$\bar{X}(t)=X(t-1)+\alpha_1+\varphi_1 m(t-1),$$

and $\bar{m}(t)=\alpha_2+\varphi_2 m(t-1)$.

Thus, if we have time-series data about $X(t)$ and $m(t)$ for a firm, we can maximize the natural logarithm of the likelihood function defined in (9) to estimate $\sigma_X, \sigma_\mu, \kappa, \rho_{\mu X}, \bar{\mu}$ and σ_I. In the next section, we discuss how to use analyst forecast data from I/B/E/S to calculate $X(t)$ and $m(t)$.

3. Variables and data

To examine the association between forecast accuracy and information precision, we model analyst forecast accuracy as a function of information precision, earnings volatility, firm size and earnings level. Omitting firm subscripts, we have the following cross-sectional regression model:

$$FAC=\beta_0+\beta_1\times IP+\beta_2\times EVOL+\beta_3\times SIZE+\beta_4\times \times EARNINGS+\varepsilon.$$

In the rest of this section, we define the regression variables, study their measurement, and discuss how to obtain data for them from I/B/E/S (1985.4-2008.4) and COMPUSTAT (1985.4-2008.4).

3.1. Forecast accuracy (FAC). We define analysts earnings forecast accuracy, FAC, as follows:

$$FAC=1/\ln\left(\frac{\sum_{i=1}^{T}\times FSR(t)}{T}\right),$$

$$FSR(t)=\frac{abs\left[FST(t)-ESP(t)\right]}{P(t)},$$

where ln denotes the natural logarithm function, t indexes quarters, T is the total number of quarters in the data sample, $FTS(T)$ is the consensus analyst forecast of earnings at quarter t, $EPS(t)$ are the actual earnings at quarter t, abs is the absolute value function, and $P(t)$ the stock price at quarter t. Clearly, analyst forecast accuracy defined above is just the inverse of the natural logarithm of the average of the absolute forecast errors that are scaled by the stock price. This measurement of analyst forecast accuracy is consistent with the practice in the analyst literature (see, for example, Hong and Kubik, 2003). The quarterly earnings forecasts and actual quarterly earnings are obtained from I/B/E/S. The quarterly stock price per share is obtained from COMPUSTAT, which is the average of the three end-of-month stock prices in each quarter.

3.2. Information precision (IP). We define IP, as follows:

$$IP=\ln\left(\frac{P}{\hat{\sigma}_I}\right),$$

$$P=\frac{\sum_{i=1}^{T}P(t)}{T},$$

where ln denotes the natural logarithm function, T is the total number of quarters in the data sample, $P(t)$ is the stock price at quarter t, P is the time average stock price of a firm in the sample, and $\hat{\sigma}_I$ is the estimate of σ_I, which is defined in equation (3). Thus information precision defined above is just the natural logarithm of the inverse of the signal volatility in equation (3) that is scaled by the time average stock price. We also obtain data for quarterly stock prices from COMPUSTAT. In the following, we address how to use the maximum likelihood estimation procedure discussed in section 2 to estimate σ_I.

To estimate σ_I, we first have to use actual quarterly earnings and analyst forecasts of quarterly earnings reported in I/B/E/S to calculate $X(t)$ and $m(t)$, which are defined in equations (7) and (8), respectively. Previous studies have shown that analyst earnings forecasts are biased. Following Das, Levine and Sivaramakrishnan (1998), we estimate the bias in consensus analyst earnings forecasts for a firm as follows:

$$Bias = \frac{\sum_{i=1}^{T}\left(FST(t) - ESP(t)\right)}{T} \tag{10}$$

where $FST(t)$ is the forecast of the earnings per share at quarter t, $EPS(t)$ are the actual earnings per share at quarter t, and T is the total number of quarters in the data sample.

From equation (10), the unbiased forecast of the earnings at quarter t,$UFS(t)$, is $FTS(t)$ - *Bias*.

According to equation (7), $X(t)$ is the actual quarterly earnings at quarter t. That is, $X(t) = EPS(t)$.

The estimate of the expected quarterly earnings growth rate, $m(t)$, at quarter t, according to Lemma 2, is

$$m(t) = \left(UFS(t+1) - ESP(t) - \bar{\mu}\right)\left(\frac{k}{1-e^{-k}}\right) + \bar{\mu},$$

where $EPS(t)$ is the actual earnings per share at quarter t, $UFS(t+1)$ is the unbiased forecast of the earnings per share at quarter $(t+1)$, $\bar{\mu}$ and k are two of the six parameters to be estimated.

We use the I/B/E/S summary file to obtain the consensus forecasts of quarterly earnings of a firm at each quarter. Specifically, in the summary file, in each month of a quarter, there is usually a consensus forecast of the earnings in that quarter. In this paper, $FST(t)$ in equation (10) is the mean of the monthly consensus forecasts of the earnings at quarter t. Actual quarterly earnings data are extracted from I/B/E/S actual files. The sample starts from the last quarter of 1985 and ends in the last quarter of 2008.

3.3. Earnings volatility (EVOL). We consider the natural logarithm of earnings volatility, denoted by *EVOL*, in our analysis to control for the impact of earnings volatility on forecast accuracy. Earnings volatility can affect forecast accuracy by affecting the precision of information that analysts have about the expected earnings growth rates of firms. When the earnings process becomes more variable, it is more difficult for analysts to obtain precise information about future earnings growth rates. So it is interesting to know whether information precision still affects analyst forecast accuracy after controlling for earnings volatility.

3.4. Firm size (SIZE). Firm size can also affect analyst forecast accuracy, since analysts tend to spend more effort to follow large firms and thus may obtain more precise information about their expected earnings growth rates. For example, Bhushan (1989) and Atiase (1985) show that in the financial market, financial analysts are likely to take more effort to gain precise information for large firms than for small firms, since precise information for large firms are more valuable for investors to generate higher profits than the same precise information for small firms. Thus, analysts tend to have more precise information about the expected earnings growth rates of large firms and thus produce more accurate forecasts. In this paper, we use the log market value of equity (*SIZE*) to control for the impact of firm size on analyst forecast accuracy.

3.5. Earnings level (EARNINGS). Like firm size, earnings level also affects analyst forecast accuracy. For example, Butler and Saraoglu (1999) and Brown (2001) show that forecast accuracy is related to the level of earnings. In our analysis, we use the natural logarithm of the time-average quarterly earnings per share, denoted by *EARNINGS*, to control for the impact of earnings level on forecast accuracy.

4. Empirical results

To estimate the volatility of the signal, σ_I , we maximize the natural logarithm of the likelihood function defined in equation (9). Since there is no closed-form solution, we use Nelder and Mead's (1965) optimization approach to estimate σ_I, which initial value is set at 15% for all the firms in our sample. The estimation exercise shows that the mean and standard deviation of $\hat{\sigma}_I$ are about 22% and 18%, respectively. The relatively noisy signal for the expected earnings growth rate has a volatility of more than 300%, but the relatively precise signal about the expected earnings growth rate of a firm has a volatility of less than 5%. Thus the precision of information that analysts have about unobservable expected earnings growth rates varies from firm to firm.

Table 1 reports the statistics for all the variables of interest in our model. Since the means of these variables are different from their medians, they are generally skewed. Table 2 shows the Pearson correlation coefficients of the regression variables. As shown there, forecast accuracy, denoted by *FAC*, is positively correlated with the information precision variable, *IP*, at the significance level of 0.0001. This empirical finding indicates that when analysts obtain more precise information about the expected earnings growth rates of firms, their earnings forecasts become more accurate. In addition, analyst forecast accuracy (*FAC*) is negatively correlated with earnings volatility, denoted by *EVOL*, also at the significance level of 0.0001. The result indicates that when earnings are more variable, analysts tend to forecast future earnings less accurately. This finding is consistent with the result of Das, Levine and Sivaranakrishnan (1998), who also show that when earnings are more variable, analyst forecasts become less accurate. Analyst forecast accuracy is also found to be positively associated with firm size, with a significance level of 0.0001. This result is expected, since analysts tend to spend more effort on large firms and thus acquire more precise information about expected earnings growth rates, which may lead to more accurate earnings forecasts. Finally, the accuracy of analyst earnings forecasts is negatively associated with the level of earnings, but not significantly.

As discussed in the previous section, earnings volatility, firm size and earnings level can affect forecast accuracy by affecting the precision of information that analysts have about the expected earnings growth rates of firms. In Table 2, these three variables are indeed correlated with the information precision variable, *IP*, with signs that are consistent with our intuition. For example, earnings volatility is negatively correlated with information precision, since when earnings become more variable, it is more difficult for analysts to obtain precise information about the expected earnings growth rates of firms.

Since firm characteristic variables are correlated with the information precision variable, it is possible that firm characteristics but not information precision explain the cross-sectional difference of forecast accuracy. To understand whether the precision of information that analysts have about the expected earnings growth rate of a firm affects forecast accuracy, in the following, we consider all the firm characteristic variables in the following multiple regression equation to test how forecast accuracy is affected by information precision.

$$FAC = \beta_0 + \beta_1 \times IP + \beta_2 \times EVOL + \beta_3 \times SIZE + \beta_4 \times \times EARNINGS + \varepsilon,$$

where firm subscripts are omitted for simplicity, and *FAC*, *IP*, *EVOL*, *SIZE*, and *EARNINGS* denote forecast accuracy, information precision, earnings volatility, firm size and earnings level per share, respectively.

Table 3 reports the regression results on how the precision of information that analysts have affects earnings forecast accuracy after firm characteristics are controlled for. As shown there, the estimate of the coefficient for information precision, denoted by *IP*, is positive and significant at the 0.0001 level. To address a possible multi-collinearity problem, the Variance Inflation Factor (VIF) is also reported. The VIF = 3.78, indicates that the regression result is not comprised by multi-collinearity. This empirical result is the same as the previous finding in the univariate analysis that when analysts obtain more precise information about expected earnings growth rates, their earnings forecasts tend to be more accurate. In our model, when analysts have no information ($\sigma \to \infty$) about the expected earnings growth rate of a firm, they just use historical earnings data to forecast future earnings. In this case, our empirical finding indicates that analyst forecasts are less accurate than the forecasts when analysts have information about the expected earnings growth rate. Thus, our finding helps explain why analyst earnings forecasts are more accurate than the earnings forecasts that are based only historical time-series earnings data (see, for example, Brown, Foster and Noreen, 1985). In addition, this empirical result suggests that to forecast earnings more accurately, analysts may have to approach management, use better resources, or employ sharper skills to obtain more precise information about expected earnings growth rates.

The estimates of the coefficients for earnings volatility, firm size and earnings level are all significant. This is an interesting result, since this finding indicates that the accuracy of the analyst earnings forecasts are affected not only by information precision but by firm characteristics, such as earnings volatility, firm size and earnings level.

To get a complete picture, we also perform a similar analysis on individual analysts. Since a company is typically followed by more than one analyst, we randomly select a forecast from the pool of all forecasts, instead of using the consensus forecast. The same procedure is used to estimate the IP variable then. The regression result is reported in Table 4. The coefficients estimates are very close to those in the previous regression with consensus forecasts.

Conclusions

In this paper, we have presented a simple continuous-time model of corporate earnings to examine how the precision of information that analysts have about the unobservable expected earnings growth rate of a firm affects earnings forecast accuracy. To forecast future earnings, analysts have to use historical earnings data and the information they have about the unobservable expected earnings growth rate of a firm to estimate the value of the expected earnings growth rate. Thus, our model captures the notion that in the financial market, analysts use more than historical earnings data to forecast future earnings.

We use biased analyst earnings forecasts to construct unbiased earnings forecasts, and develop a maximum likelihood procedure to estimate the precision of information that analysts have about the unobservable expected earnings growth rate of a firm. Our estimation exercise shows that the precision of information that analysts have varies from firm to firm. For some firms, analysts have relatively very precise information; for some firms, the information that analysts have about future earnings growth is relatively noisy.

We use data from I/B/E/S and COMPUSTAT to cross-sectionally analyze how analysts' information precision affects forecast accuracy. By controlling for firm characteristics such as earnings volatility, firm size and earnings level, we find that earnings forecast accuracy is positively associated with the precision of information that analysts have about the expected earnings growth rates of firms. This finding indicates that when analysts have more precise information about expected earnings growth rates, their earnings forecasts are more accurate. This empirical finding helps explain why analysts' earnings forecasts are more accurate than the earnings forecasts based only on past realized earnings (see, for example, Brown, Foster and Noreen, 1985), and suggests that to forecast future earnings more accurately, analysts may have to approach management, use better resources, or employ sharper skills to obtain more precise information about future earnings growth rates.

References

1. Abarbanell, J.S. (1991). Do Analysts' Earnings Forecasts Incorporate Information In Prior Stock Price Changes? *Journal of Accounting and Economics*, 14, pp. 147-165.
2. Abarbanell, J.S., W.N. Lanen, and Verrecchia, R.E. (1995), Analysts' Forecasts as Proxies for Investor Beliefs in Empirical Research, *Journal of Accounting and Economics*, 20, pp. 31-60.
3. Atiase, R. (1985). Predisclosure Information, Firm Capitalization, And Security Price Reaction Around Earnings Announcements, *Journal of Accounting Research*, 23 (1), pp. 21-36.
4. Bhushan, R. (1989). Firm Characteristics and Analyst Following, *Journal of Accounting and Economics*, 11, pp. 255-274.
5. Barron, O.E., O. Kim, S.C. Lim and Stevens, D.E. (1998). Using Analysts' Forecasts to Measure Properties of Analysts' Information Environment, *The Accounting Review*, 73, pp. 421-433.
6. Barry, C. and Jennings, R. (1992). Information and Diversity of Analyst Opinion, *Journal of Financial and Quantitative Analysis*, 27, pp. 169-183.
7. Brennan, M. and Xia, Y. (2001). Stock Price Volatility, Learning, And The Equity Premium, *Journal of Monetary Economics*, 47, pp. 249-283.
8. Brown, P., G. Foster and Noreen, E. (1985). Security Analyst Multi-Year Earnings Forecasts and the Capital Market, American Accounting Association, Sarasota, FL.
9. Clement, M. B. (1999). Analysts Forecast Accuracy: Do Ability, Resources, and Portfolio Complexity Matter, *Journal of Accounting and Economics*, 27, pp. 285-303.
10. Collins, D., Kothari, S., and Rayburn, J. (1987). Firm Size and The Information Content of Prices with Respect to Earnings, *Journal of Accounting and Economics*, 9, pp. 111-138.
11. Das, S., C.B. Levine, and Sivaramakrishnan, R. (1998). Earnings Predictability and Bias in Analysts' earnings Forecasts, *The Accounting Review*, Vol. 73, pp. 277-294.
12. Frankel, Richard, S.P. Kothari, Joseph Weber (2006). Determinants of the informativeness of analyst research, *Journal of Accounting and Economics*, Volume 41, Issues 1-2, pp. 29-54.
13. Hughes, J., and Ricks, W. (1987). Associations between Forecast Errors and Excess Returns near to Earnings Announcements, *Accounting Review*, 62, pp. 158-175.
14. Hong, H., T. Lim, and J.C. Stein (2000). Bad News Travels Slowly: Size, Analysts Coverage And The Profitability of Momentum Strategies, *Journal of Finance*, 55, pp. 265-296.
15. Hong, H., and J.D. Kubik (2003). Analyzing The Analysts: Career Concerns and Biased Earnings Forecasts, *Journal of Finance*, 58, pp. 313-351.
16. Kandel, S., and R.F. Stambaugh (1990). Expectations and Volatility of Consumption and Asset Returns, *Review of Financial Studies*, 3, pp. 207-232.
17. Landsman, W., Maydew, E. (2002). Has The Information Content of Quarterly Earnings Announcements Declined In The Past Three Decades? *Journal of Accounting Research*, 40, pp. 797-808.
18. Lipster, R.S., and A.N. Shiryayev (1978). *Statistics of Random Processes I &II*, Springer-Verlag, New York.
19. McNichols, M. (1989). Evidence of Informational Asymmetries from Management Earnings Forecasts and Stock Returns, *Accounting Review*, 64, pp. 1-27.
20. Nelder, J.A. and Mead, R. (1965). The Downhill Simplex Method, *Computer Journal*, Vol. 7, pp. 308-313.
21. Stickel, S. (1990). Predicting Individual Analyst Earnings Forecasts, *Journal of Accounting Research*, 28, pp. 409-417.

22. Stickel, S. (1992). Reputation And Performance Among Security Analysts, *Journal of Finance*, 47, pp. 1811-1836.
23. Stickel, S. (1995). The Anatomy Of The Performance Of Buy And Sell Recommendations, *Financial Analysts Journal*, 51, pp. 25-39.
24. Veronesi, Petro (2000). How Does Information Quality Affect Stock Returns? *Journal of Finance*, 55, pp. 807-837.
25. Wang, J. (1993). A Model of Intertemporal Asset Prices under Asymmetric Information, *Review of Economic Studies*, 60, pp. 249-282.
26. Wang, J. (1994). A Model of Competitive Stock Trading Volume, *Journal of Political Economy*, 102, pp. 127-167.

Appendix

Proof of Lemma 1. We follow Brennan and Xia (2001) and Wang (1993) to use Theorem 12.1 in Liptser and Shiryayev (1977) to show Lemma 1. Using the similar notation, we rewrite our problem as follows:

$$ds = \begin{pmatrix} dI \\ dX \end{pmatrix} = \left[a_{s0} + a_{s1}\mu\right]dt + \sigma_s \begin{pmatrix} dW_I \\ dW_X \end{pmatrix},$$

$$d\mu = \left[a_{\mu 0} + a_{\mu 1}\mu\right]dt + \sigma_\mu dW_\mu,$$

where ds is a 2×1 vector signal, which is used by analysts to estimate $\mu(t)$, the state variable. Other parameters are as follows:

$$a_{s0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},\ a_{s1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

$$\sigma_s = \begin{pmatrix} \sigma_I & 0 \\ 0 & \sigma_X \end{pmatrix},$$

$a_{\mu 0} = k\bar{\mu}$ and $a_{\mu 1} = -k$.

We also define

$$q_{ss} = \begin{bmatrix} \sigma_I^2 & 0 \\ 0 & \sigma_X^2 \end{bmatrix},\ q_{\mu\mu} = \sigma_\mu^2 \text{ and } q_{\mu s} = \begin{bmatrix} 0 & \sigma_{\mu X} \end{bmatrix},$$

where $\sigma_{\mu X} = \rho_{\mu X}\sigma_\mu\sigma_X$ and $\rho_{\mu X}$ denotes the correlation between $W_\mu(t)$ and $W_X(t)$.

Let $F_t = \{X(\tau),\ I(\tau):\ \tau \le t\}$ be the analyst's information set at time t. Suppose that the prior is $\mu(0) \sim N(m(0), v(0))$. Then, according to Liptser and Shiryayev (1977), the posterior mean of $\mu(t)$, $m(t) \equiv E\left[\mu \mid F_t\right]$, and the posterior variance of $\mu(t)$, $v(t) \equiv E\left[(m-\mu)(m-\mu)^T \mid F_t\right]$, are given, respectively, by the following stochastic differential equations:

$$dm(t) = \left(a_{\mu 0} + a_{\mu 1}m(t)\right)dt + \left(v(t)a_{s1}^T + q_{\mu s}\right)q_{ss}^{-1}\sigma_s d\tilde{W}_s \tag{A1}$$

$$\frac{dv(t)}{dt} = -2\kappa v(t) + \sigma_\mu^2 - \left(v(t)a_{s1}^T + q_{\mu s}\right)q_{ss}^{-1}\left(v(t)a_{s1}^T + q_{\mu s}\right)^T, \tag{A2}$$

where the innovation process, $\tilde{W}_s(t) = [\tilde{W}_I, \tilde{W}_X]^T$, defined by

$d\tilde{W}_s = \sigma_s^{-1}\left[ds - \left(a_{s0} + a_{s1}m\right)dt\right]$, is a vector of Brownian motions.

The earnings process then becomes

$$dX(t) = m(t)dt + \sigma_X d\tilde{W}_X \tag{A3}$$

The solution to the Ricatti equation in (A2) is given by:

$$v(t)=\frac{v_2-v_1\frac{v(0)-v_2}{v(0)-v_1}e^{-\varpi t}}{1-\frac{v(0)-v_2}{v(0)-v_1}e^{-\varpi t}},$$

where

$$\varpi=\sqrt{b_2^2-4b_1b_3},\ v_1=\frac{-b_2-\varpi}{2b_1},\ v_2=\frac{-b_2+\varpi}{2b_1},$$

$$b_1=\frac{1}{\sigma_X^2}+\frac{1}{\sigma_I^2},$$

$$b_2=2\left(k+\sigma_{\mu X}/\sigma_X^2\right)\ \text{and}\ b_3=\sigma_{\mu X}^2/\sigma_X^2-\sigma_\mu^2.$$

Since $\rho_{\mu X}^2\le 1, b_1>0$ and $b_3=\sigma_{\mu X}^2/\sigma_X^2-\sigma_\mu^2=\sigma_\mu^2(\rho_{\mu X}^2-1)\le 0$, hence $b_2^2-4b_1b_3\ge 0$ and there is always a solution to the Ricatti equation.

In this paper, we follow Wang (1993 and 1994) to consider the steady-state solution, in which estimation errors do not change over time. If the economy starts at time zero, when t is very large, the convergence of learning to the steady state is guaranteed, since $\varpi\ge 0$. When learning reaches the steady state, $dv(t)/dt=0$. Let v be the solution to the Ricatti equation in the steady state. Then $v = v2$.

In the steady state, we have:

$$dm=k(\bar{\mu}-m)dt+a_1d\tilde{W}_I+a_2d\tilde{W}_X, \tag{A4}$$

where

$$a_1=\frac{v}{\sigma_I}\ \text{and}\ a_2=\frac{v+\sigma_{\mu X}}{\sigma_X}.$$

Equation (A4) can also be simplified as:

$$dm=k(\bar{\mu}-m)dt+\sigma_m d\tilde{W}_m, \tag{A5}$$

where $\tilde{W}_m$ is a standard Brownian motion and

$$\sigma_m^2=a_1^2+a_2^2 \tag{A6}$$

Proof of Lemma 2. The solution to stochastic equation (A5) is:

$$m(s)=\bar{\mu}+e^{-k(s-t)}(m(t)-\bar{\mu})+\int_t^s e^{-k(s-\tau)}\sigma_m d\tilde{W}_m(\tau), \tag{A7}$$

where $s > t$.

The solution to equation (A3) is:

$$X(s)=X(t)+\int_t^s m(\tau)d\tau+\int_t^s \sigma_X d\tilde{W}_X(\tau) \tag{A8}$$

Substituting $m(\tau)$ from equation (A7) into equation (A8) yields

$$X(s)=X(t)+\bar{\mu}(s-t)+\frac{(m(t)-\bar{\mu})}{k}\left(1-e^{-k(s-t)}\right)+\frac{\sigma_m}{k}\int_t^s\left(1-e^{-k(s-\tau)}\right)d\tilde{W}_m(\tau)+\int_t^s\sigma_X d\tilde{W}_X(\tau). \tag{A9}$$

Since the conditional expectation of the last two terms in the left hand of equation (A9) are zero, we have:

$$UST(s)=E\left[X(s)\,|\,F_t\right]=X(t)+\bar{\mu}(s-t)+\frac{(m(t)-\bar{\mu}))}{k}\left(1-e^{-k(s-t)}\right).$$

Lemma 3. Let $\varepsilon_X(t)$ and $\varepsilon_m(t)$ be defined as above. Let σ_1 and σ_2 be the standard deviations of $\varepsilon_X(t)$ and $\varepsilon_m(t)$, respectively. Let $\sigma_{12}=\rho_{12}\sigma_1\sigma_2$ be the covariance of these two random variables, where ρ_{12} is the correlation between these two variables. Define $\beta(i)=(1-e^{-i\kappa})/(ik)$, where i is a non-negative integer and k is the parameter to be estimated. Then $\varepsilon_X(t)$ and $\varepsilon_m(t)$ are both normal, with the following conditional moments:

$$E\left[\varepsilon_m(t)\,|\,F_{t-1}\right]=E\left[\varepsilon_X(t)\,|\,F_{t-1}\right]=0,$$

$$\sigma_1^2=Var\left[\varepsilon_X(t)\,|\,F_{t-1}\right]=\sigma_X^2+\frac{\sigma_m^2}{k^2}\left[1-2\beta(1)+\beta(2)\right]+\frac{2\sigma_{mX}}{k}\left[1-\beta(1)\right]\ , \tag{A10}$$

$$\sigma_2^2=Var\left[\varepsilon_m(t)\,|\,F_{t-1}\right]=\sigma_m^2\beta(2)\ , \tag{A11}$$

$$\sigma_{12}=Cov\left[\left(\varepsilon_X(t),\varepsilon_m(t)\right)|\,F_{t-1}\right]=\beta(1)\sigma_{mX}+\frac{\sigma_m^2}{k}\left[\beta(1)-\beta(2)\right]\ . \tag{A12}$$

Proof of Lemma 3. According to the normal property of a Brownian motion, straight calculations lead to the result. The calculation of σ_m^2 and σ_{mX} is summarized in the following lemma.

Lemma 4. Let v be the estimation error in the steady state, as defined in the appendix. Then we have:

$$\sigma_m^2=\frac{v^2}{\sigma_I^2}+\frac{(v+\sigma_{\mu X})^2}{\sigma_X^2}, \tag{A13}$$

$$\sigma_{mX}=\sigma_{\mu X}+v, \tag{A14}$$

where estimation error v is a function of σ_X, σ_μ, κ, $\rho_{\mu X}$, $\bar{\mu}$ and σ_I.

Substituting σ_m^2 and σ_{mX} from equations (A13) and (A14) into equations (A10)-(A12), we can have σ_1, σ_2 and ρ_{12} expressed as the functions of σ_X, σ_μ, k, $\rho_{\mu X}$, $\bar{\mu}$ and σ_I.

Proof of Lemma 4. From the definitions of a_1, a_2 and σ_m^2 above, we have:

$$\sigma_m^2=a_1^2+a_2^2=\frac{v^2}{\sigma_I^2}+\frac{(v+\sigma_{\mu X})^2}{\sigma_X^2},$$

According to equations (A4) and (A3), we have

$$\sigma_{mX}dt=\mathrm{cov}(dm,dX)=a_2\sigma_X dt=\left(\sigma_{\mu X}+v\right)dt\ .$$

Table 1. Descriptive statistics for variables (n = 319)

	FAC	*IP*	*EVOL*	*SIZE*	*EARNINGS*
Mean	-0.15	5.32	-1.52	8.21	-1.28
Median	-0.15	5.40	-1.60	8.13	-1.17
SD	0.03	0.62	0.63	1.31	0.73
Minimum	-0.27	2.60	-2.79	4.43	-5.53
Maximum	-0.11	6.62	0.72	12.00	1.33

Notes: Variable definitions: n = the sample size; *FAC* = The inverse of the log average of the absolute error scaled by the stock price; *IP* = The log of the inverse of the signal volatility scaled by the stock price; *EVOL* = The log earnings volatility; *SIZE* = The log of the average market value of equity; *EARNINGS* = The log of average earnings per share.

Table 2. Pearson correlation coefficients among variables (n = 319)

	FAC	*IP*	*EVOL*	*SIZE*	*EARNINGS*
FAC	1.000				
IP	0.625 **	1.000			
EVOL	-0.681 **	-0.245 **	1.000		

Table 2 (cont.). Pearson correlation coefficients among variables (n = 319)

	FAC	*IP*	*EVOL*	*SIZE*	*EARNINGS*
SIZE	0.441 **	0.537 **	-0.068	1.000	
EARNINGS	-0.023	0.155 *	0.472 **	0.145 *	1.000

Notes: Variable definitions: n = the sample size; *FAC* = The inverse of the log average of the absolute error scaled by the stock price; *IP* = The log of the inverse of the signal volatility scaled by the stock price; *EVOL* = The log earnings volatility; *SIZE* = The log of the average market value of equity; *EARNINGS* = The log of average earnings per share. Significance: ** < 0.0001, * < 0.05.

Table 3. Information precision and consensus forecast accuracy
(*FAC* and *IP* computed from consensus forecasts, n = 319)

$$FAC = \beta_0 + \beta_1 \times IP + \beta_2 \times EVOL + \beta_3 \times SIZE + \beta_4 \times EARNINGS + \varepsilon$$

	IP	*EVOL*	*SIZE*	*EARNING*	Adj. R^2	*VIF*
Predicted signs	+	-	+	+		
Estimated coefficient	0.014	-0.031	0.004	0.009	75%	3.78
t-value	7.32 *	-16.76 *	4.71 *	5.55 *		

Notes: Variable definitions: n = the sample size; *FAC* = The inverse of the log average of the absolute error scaled by the stock price, using consensus forecasts; *IP* = The log of the inverse of the signal volatility scaled by the stock price, using consensus forecasts; *EVOL* = The log earnings volatility; *SIZE* = The log of the average market value of equity; *EARNINGS* = The log of average earnings per share; *VIF* = Variance Inflation Factor. Significance: *<0.0001.

Table 4. Information precision and individual forecast accuracy
(*FAC* and IP computed from random individual forecasts, n = 319)

$$FAC = \beta_0 + \beta_1 \times IP + \beta_2 \times EVOL + \beta_3 \times SIZE + \beta_4 \times EARNINGS + \varepsilon$$

	IP	*EVOL*	*SIZE*	*EARNING*	Adj. R^2	*VIF*
Predicted signs	+	-	+	+		
Estimated coefficient	0.011	-0.046	0.004	0.009	78%	3.02
t-value	5.70 *	-17.03 *	3.94 *	7.58 *		

Notes: Variable definitions: n = the sample size; *FAC* = The inverse of the log average of the absolute error scaled by the stock price, using a randomly selected forecast on each company; *IP* = The log of the inverse of the signal volatility scaled by the stock price, using a randomly selected forecast on each company; *EVOL* = The log earnings volatility; *SIZE* = The log of the average market value of equity; *EARNINGS* = The log of average earnings per share; *VIF* = Variance Inflation Factor. Significance: * < 0.0001.