The Return-Earnings Relation When Analyst Forecasts Are Used as a Proxy for Investor Expectations*

Orie E. Barron
Pennsylvania State University

Oliver Kim
University of Maryland

Douglas E. Stevens
Syracuse University

February 2002

Abstract

In this paper we model return-earnings regressions in a setting in which analysts' forecasts are used to proxy for investors' expectations. We show how disparities between investors' and analysts' information sets could affect the regression coefficients and the R^2. Our findings provide new and untested implications as well as providing insights into recent return-earnings evidence in the empirical literature.

* We thank Pamela Stuerke and workshop participants at the University of Pennsylvania and Syracuse University for their helpful comments.

1 Introduction

In this paper we theoretically examine the implications of using earnings forecasts as proxies for investors' beliefs in explaining stock returns. This study is motivated by the large number of empirical capital market studies in accounting that have used analysts' earnings forecasts when attempting to understand the relation between stock returns and accounting earnings. Analysts' earnings forecasts are often used to empirically investigate the return-earnings relation because researchers can not directly observe investors' expectations of the forthcoming earnings. In the typical empirical study, stock returns are regressed on the error in analysts' earnings forecasts. Thus, we explicitly model regressions of stock price change on earnings surprise based on earnings forecasts that reflect the underlying information environment of investors and analysts. In particular, we use the precision of public and private information available to investors and analysts to derive expressions for four commonly used measures of the price reaction to an earnings disclosure, namely, the variance of price change, the earnings response coefficient, the reverse-regression response coefficient, and the R^2 of the return-earnings regression.

We model a financial market that is similar to that in Kim and Verrecchia (1991, hereafter KV). As in KV, we derive an expression for price change that is based on the relative importance of the earnings information to the total information about the liquidating dividend. The information structures of investors and analysts in our model are similar to the model in Barron, Kim, Lim and Stevens (1998, hereafter BKLS). That is, each agent observes public and private signals about the forthcoming earnings. The advantage of using this information structure is that the results of our model can be tied to the observable forecast variables suggested by BKLS.

The only other study in the literature that explicitly relates investors' beliefs to characteristics of analysts' forecasts is Abarbanell, Lanen and Verrecchia (1995, hereafter ALV). ALV model this relation by assuming that investors observe earnings forecasts that are in the form of exogenous signals with common and idiosyncratic error terms. Investors observe the signals and, as a result, their beliefs incorporate the forecasts. This study makes contributions distinct from ALV in three important directions. First, our model incorporates analysts' earnings forecasts that are endogenously determined by the underlying information environment. Second, while investors observe the analysts' forecasts as in ALV, they may not infer all of the information possessed by analysts. Thus, we analyze how investors learn from analysts' forecasts and the effect of such learning on the return-earnings relation. Third, as mentioned above, this study utilizes the information structure in BKLS so that the results can be used to generate testable hypotheses regarding how analyst forecast measures are related to the return-earnings relation. The explicit analysis of four measures of the market reaction to an earnings disclosure used in the empirical literature also contributes to the specificity of these testable hypotheses.

We obtain two major results. First, we show in Observation 2 that using the mean analyst earnings surprise causes both the ERC and the regression R^2 to be amplified by a factor determined by the degree to which idiosyncratic surprises are not accounted for because they are diversified away from the mean forecast (i.e., because of using multiple forecasts). Second, we also show that a regression based on the mean analyst earnings surprise generates ERC and R^2 smaller than the true informational importance of earnings to investors under the assumption that there is certain information observed by investors but not reflected in earnings forecasts. If such information asymmetry between investors and analysts is significant, it provides another reason why both the ERC and the R^2 are small in most return-earnings regressions (e.g., Easton and Zmijewski 1989, Lipe 1990, and Kothari and Sloan 1992). Moreover, we argue that this asymmetry-related factor can cause non-linearities in the price reaction to the magnitude of earnings surprise (e.g., see Freeman and Tse 1992).

We begin the analysis in Section 2 by describing the model and the equilibrium market prices. In Section 3 we consider different measures of earnings surprise and their properties. In Section 4 we examine the regression metrics generated by different measures of earnings surprise and, in particular, analyze the effect of using forecast proxies on the ERC and R^2. In Section 5 we provide empirical implications of our results, and conclude in Section 6.

2 The Model and Market Equilibrium

We use a model of the financial market similar to that of Kim and Verrecchia (1991). In this section we describe our model and highlight the features of the rational expectations equilibrium we exploit in our analysis of the return-earnings relation. All proofs are in the appendix.

In our model there are three time periods and two assets, a risky asset traced to the terminal value of a firm and a riskless bond. One unit of the riskless bond pays off one unit of consumption good in period 3. The value of the risky asset is determined by the terminal value of the firm, $\widetilde{u}$, which is a random variable. Trading occurs in periods 1 and 2 and the realization of $\widetilde{u}$ occurs in period 3 along with final consumption. During period 2 the firm discloses an earnings report, $\widetilde{y}$, which is about the terminal value of the firm, $\widetilde{u}$. In this economy there are I investors and N

financial analysts, where $I = \infty$ and $N < \infty$. The timeline of events in our model appears in Table 1 and a summary of the variables is presented in Table 2.

[Place Tables 1 & 2 About Here]

In our model information about the return, $\widetilde{u}$, of the risky asset is obtained in the form of information about earnings, $\widetilde{y} = \widetilde{u} + \widetilde{\eta}$, where $\widetilde{\eta}$ is a random variable representing the noise in the earnings. Five events occur in period 1. First, investor i is endowed with E_i riskless bond and x_i risky asset. The average risky endowment in period 1, $\widetilde{x}_1$, is a random variable that is normally distributed with mean 0 and precision t_1. This randomness in the risky asset supply prevents the prices of the risky asset from fully revealing all private information in the market. Second, analyst $a = 1, 2, ..., N$, observes three signals about earnings, $\widetilde{w}_c = \widetilde{y} + \widetilde{\nu}_c$, $\widetilde{w}' = \widetilde{y} + \widetilde{\nu}'$, and $\widetilde{z}_a = \widetilde{y} + \widetilde{\varepsilon}_a$. Signal $\widetilde{w}_c$ is observed by all analysts and all investors, signal $\widetilde{w}'$ is observed by all analysts but not by investors, and signal $\widetilde{z}_a$ is observed privately by analyst a. Third, each analyst announces his forecast of earnings, denoted by $\widetilde{F}_a$, based on his knowledge. Fourth, investor i also observes three signals about the forthcoming earnings, namely, $\widetilde{w}_c = \widetilde{y} + \widetilde{\nu}_c$, $\widetilde{w} = \widetilde{y} + \widetilde{\nu}$, $\widetilde{z}_i = \widetilde{y} + \widetilde{\varepsilon}_i$, as well as the N earnings forecasts $\widetilde{F}_a$'s. Signal $\widetilde{w}_c$ is observed by all analysts and all investors, signal $\widetilde{w}$ is observed by all investors but not analysts, and signal $\widetilde{z}_i$ is observed privately by investor i. However, individual investors extract some of the analysts' exclusive information, $\widetilde{w}'$ and $\widetilde{z}_a$'s by observing their earnings forecasts. We assume for simplicity that the liquidating dividend, $\widetilde{u}$, the earnings noise, $\widetilde{\eta}$, and the error terms of the above signals, $\widetilde{\nu}_c$, $\widetilde{\nu}'$, $\widetilde{\nu}$, $\widetilde{\varepsilon}_a$, and $\widetilde{\varepsilon}_i$, are all mutually independent and normally distributed with mean $(\overline{u}, 0, 0, 0, 0, 0, 0)$ and precision $(h, n, m_c, m', m, s_a, s)$. Fifth, the market opens and investors trade at competitive market price $\widetilde{P}_1$, which provides another public signal to investors ($\widetilde{q}_1$ with precision $\frac{t_1}{B^2}$).

Two events occur in period 2. First, the firm publicly discloses an earnings report $\widetilde{y} = \widetilde{u} + \widetilde{\eta}$. Second, the market opens again and investors engage in trade. The average risky asset supply in period 2 is $\widetilde{x}_2$, which is normally distributed with mean 0 and precision t_2 and is independent of $\widetilde{x}_1$ and all other random variables.[1] In period 3, $\widetilde{u}$ is realized, the firm is liquidated, and consumption occurs.

Investors are risk averse (negative exponential utility function) with risk tolerance r. Market prices are linear functions of the average of all information signals of investors at the time of trading and of the supply noise in the period. Traders condition their expectations on the market price of the risky asset as well as on their available information, and make self-fulfilling conjectures about the relation between prices and investors' information.

There are a number of differences between our model and those of KV and ALV that should be noted. First, we assume that investors' and analysts' information is about the forthcoming earnings instead of about the liquidating dividend of the risky asset.[2] Second, we assume that investors' risk tolerance and the precision of their private information as well as the precision of analysts' private information are homogeneous.[3] Third, unlike ALV, we explicitly model how investors learn from earnings forecasts in a very general situation where information asymmetries in opposite directions originally coexist between investors and analysts.

In order to analyze the equilibrium in this model, we first examine analysts'

[1] Assuming correlated aggregate risky supplies between periods 1 and 2 does not qualitatively change the results.

[2] We make this assumption to explicitly relate market prices to observable forecast variables including forecast dispersion and error in the mean forecast. Assuming that there is private information about the liquidating dividends in addition to information about earnings does not qualitatively change the results.

[3] This assumption is made for simplicity, for we do not focus on the asymmetry in investors' risk tolerance or in the quality of investors' or analysts' private information.

earnings forecasts and how investors learn from them. We assume that analyst a's forecast, denoted by $\widetilde{F}_a$, is simply his or her expectation of earnings conditional on the information in their possession. That is,

$$\widetilde{F}_a = E[\widetilde{y}|\widetilde{w}_c, \widetilde{w}', \widetilde{z}_a] = \frac{\frac{hn}{h+n}\overline{u} + m_c\widetilde{w}_c + m'\widetilde{w}' + s_a\widetilde{z}_a}{\frac{hn}{h+n} + m_c + m' + s_a}, \tag{1}$$

where $\overline{u}$ and $\frac{hn}{h+n} = \frac{1}{h^{-1}+n^{-1}}$ are the prior mean and the precision of earnings $\widetilde{y}$, respectively. For simplicity, we assume that all N analysts announce their forecasts simultaneously and thus their forecasts are not affected by information obtained from the observation of other analysts' forecasts.[4]

The mean analyst forecast, $\widetilde{F}_{\Sigma a}$, can therefore be expressed as

$$\widetilde{F}_{\Sigma a} = \frac{1}{N}\sum_{a=1}^{N}\widetilde{F}_a = \frac{\frac{hn}{h+n}\overline{u} + m_c\widetilde{w}_c + m'\widetilde{w}' + \frac{1}{N}\sum_{a=1}^{N} s_a\widetilde{z}_a}{\frac{hn}{h+n} + m_c + m' + s_a}. \tag{2}$$

We show in the appendix that investor i's information in period 1 about the forthcoming earnings can be characterized by the conditional expectation,

$$\begin{aligned}\widetilde{\mu}_{1i} &= E[\widetilde{y}|\widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{q}_1, \widetilde{F}_1, \cdots, \widetilde{F}_N] \\ &= \frac{\frac{hn}{h+n}\overline{u} + m_c\widetilde{w}_c + m\widetilde{w} + s\widetilde{z}_i + \frac{t_1}{B^2}\widetilde{q}_1 + \left(\frac{m'+s_a}{m'+\frac{s_a}{N}}\right)\left[m'\widetilde{w}' + \frac{1}{N}\sum_{a=1}^{N} s_a\widetilde{z}_a\right]}{\frac{hn}{h+n} + m_c + m + s + \frac{t_1}{B^2} + \left(\frac{m'+s_a}{m'+\frac{s_a}{N}}\right)(m'+s_a)},\end{aligned} \tag{3}$$

and the conditional precision,

$$\begin{aligned}K &\equiv Var^{-1}(\widetilde{y}|\widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{q}_1, \widetilde{F}_1, \cdots, \widetilde{F}_N) \\ &= \frac{hn}{h+n} + m_c + m + s + \frac{t_1}{B^2} + \left(\frac{m'+s_a}{m'+\frac{s_a}{N}}\right)(m'+s_a),\end{aligned} \tag{4}$$

[4] Since there are information signals, $\widetilde{w}_c$ and $\widetilde{w}'$, common among analysts, this assumption of simultaneous announcements of earnings forecasts is not overly restrictive.

where $\widetilde{q}_1$ is the normalized price signal in period 1 and B is a positive number determined in a rational expectations equilibrium by $B = \left(\frac{n}{h+n}\right)\frac{1}{rs}$ as shown in the appendix. Equations (3) and (4) show that an individual investor's information in period 1 consists of the prior ($\overline{u}$ with precision $\frac{hn}{h+n}$), directly observed public and private signals ($\widetilde{w}_c$, $\widetilde{w}$, and $\widetilde{z}_i$ with precisions m_c, m, and s, respectively), the price signal ($\widetilde{q}_1$ with precision $\frac{t_1}{B^2}$), and information obtained from the earnings forecasts.

An analysis of these first four equations shows how investors learn from analysts' earnings forecasts. As can be seen from equations (1) and (2), analysts' forecasts contain both information also possessed by investors, namely, the prior $\overline{u}$ and $\widetilde{w}_c$, and information possessed only by analysts, namely, $\widetilde{w}'$ and $\widetilde{z}_a$'s. Investor i processes observed forecasts first by filtering out the components due to $\overline{u}$ and $\widetilde{w}_c$, because investors have direct access to this information. When there exists information common among analysts and also information private to individual analysts, investors are unable to extract the entire information observed by analysts as a whole (i.e., $\widetilde{w}'$ and all N $\widetilde{z}_a$'s) from earnings forecasts alone. In other words, the precision achieved from observing N forecasts is less than the total precision of $\widetilde{w}'$ and all N $\widetilde{z}_a$'s, i.e.,

$$\left(\frac{m' + s_a}{m' + \frac{s_a}{N}}\right)(m' + s_a) - (m' + Ns_a) = \frac{-2(N-1)m's_a}{m' + \frac{s_a}{N}} < 0$$

for $N > 1$. This is due to the fact that an individual analyst's relative weighting of $\widetilde{w}'$ and $\widetilde{z}_a$ by their respective precision m' and s_a persists throughout the processing of earnings forecasts by investors, and this weighting overweights $\widetilde{w}'$ (and thus underweights $\widetilde{z}_a$'s).[5]

The equilibrium price in period 1 is computed in the appendix to be the expectation of the liquidating dividend conditional on the available information perturbed

[5] See Kim, Lim and Shaw (2001) for a detailed discussion of the underweighting of private information in aggregating forecasts.

by supply noise. That is,

$$\widetilde{P}_1 = \frac{h\overline{u} + n\widetilde{\mu}_1}{h+n} - \left(\frac{n}{h+n}\right)^2 \frac{\widetilde{x}_1}{rK},$$

where $\widetilde{\mu}_1$ is the average of $\widetilde{\mu}_{1i}$ across all investors.

After the earnings announcement in period 2 investor i's expectation about the liquidating dividend is

$$\widetilde{\mu}_{2i} \equiv E[\widetilde{u}|\widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{F}_1, \cdots, \widetilde{F}_N, \widetilde{P}_1, \widetilde{y}, \widetilde{P}_2] = E[\widetilde{u}|\widetilde{y}] = \frac{h\overline{u} + n\widetilde{y}}{h+n},$$

because earnings, $\widetilde{y}$, is a sufficient statistic for signals $\widetilde{w}_c$, $\widetilde{w}$, $\widetilde{w}'$, $\widetilde{z}_i$, $\widetilde{z}_a$, $\widetilde{q}_1$, $\widetilde{q}_2$ and the earnings forecasts, $\widetilde{F}_1, \cdots, \widetilde{F}_N$. The equilibrium price in period 2 then becomes

$$\widetilde{P}_2 = \frac{h\overline{u} + n\widetilde{y} - \frac{\widetilde{x}_2}{r}}{h+n},$$

which is again the expectation of the liquidating dividend perturbed by supply noise in period 2.

From the proof in the appendix, price change can be written as

$$\begin{aligned}
&\widetilde{P}_2 - \widetilde{P}_1 \\
&= \frac{n}{h+n} \cdot \left[\widetilde{y} - \frac{\frac{hn}{h+n}\overline{u} + m_c\widetilde{w}_c + m\widetilde{w} + \left(s + \frac{t_1}{B^2}\right)\widetilde{q}_1 + \left(\frac{m'+s_a}{m'+\frac{s_a}{N}}\right)\left\{m'\widetilde{w}' + \frac{1}{N}\sum_{a=1}^N s_a\widetilde{z}_a\right\}}{\frac{hn}{h+n} + m_c + m + s + \frac{t_1}{B^2} + \frac{(m'+s_a)^2}{m'+\frac{s_a}{N}}}\right] \\
&\quad - \frac{\widetilde{x}_2}{r(h+n)} \\
&= \frac{n}{h+n} \cdot \left[\frac{\frac{hn}{h+n}(\widetilde{y} - \overline{u}) - m_c\widetilde{\nu}_c - m\widetilde{\nu} + \left(\frac{t_1}{B} + \frac{n}{r(h+n)}\right)\widetilde{x}_1 - \left(\frac{m'+s_a}{m'+\frac{s_a}{N}}\right)\left\{m'\widetilde{\nu}' + \frac{1}{N}\sum_{a=1}^N s_a\widetilde{\varepsilon}_a\right\}}{\frac{hn}{h+n} + m_c + m + s + \frac{t_1}{B^2} + \frac{(m'+s_a)^2}{m'+\frac{s_a}{N}}}\right] \\
&\quad - \frac{\widetilde{x}_2}{r(h+n)}. \qquad (5)
\end{aligned}$$

From equation (5) the variance of price change can be calculated as

$$\begin{aligned} Var(\widetilde{P}_2 - \widetilde{P}_1) &= \left(\frac{n}{h+n}\right)^2 \cdot \frac{1}{K} \cdot \left(\frac{K + s + \frac{n^2}{r^2 t_1 (h+n)^2} + \frac{K^2}{r^2 t_2 n^2}}{K}\right) \\ &= \left(\frac{n}{h+n}\right)^2 \cdot \frac{1}{K} \cdot \Psi \end{aligned} \tag{6}$$

where $\Psi \equiv (K + s + \frac{n^2}{r^2 t_1 (h+n)^2} + \frac{K^2}{r^2 t_2 n^2})/K$, which reflects the noisiness of the market due to the presence of supply noise. Equation (6) shows that the variance of price change can be expressed as the product of three terms: the square of the relative precision of the disclosed earnings $\frac{n}{h+n}$, the overall uncertainty of earnings $\frac{1}{K}$, and the noisiness of the market Ψ. Noisiness is greater than 1 because price variance arises not only from the arrival of earnings information but also from the presence of supply noise in periods 1 and 2.

3 Measures of Earnings Surprise and Their Properties

In this section we analyze four different measures of earnings surprise, including two that are based on the underlying expectations of investors. Despite the fact that investor expectations are not directly unobservable, this analysis is important because it is the effect of earnings information on investor expectations and market outcomes that accounting researchers *intend* to examine in return-earnings studies. By examining earnings surprise measures based on both investors' expectations and analysts' forecasts, we are later able to analyze the effect on the return-earnings regression of using analysts' forecasts as a proxy for investors' expectations.

We define earnings surprise, denoted by UE_i, based on an individual investor's expectation as

$$\begin{aligned}
UE_i &\equiv \widetilde{y} - E[\widetilde{y}|\widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{q}_1, \widetilde{F}_1, \cdots, \widetilde{F}_N] \\
&= \widetilde{y} - \frac{\frac{hn}{h+n}\overline{u} + m_c\widetilde{w}_c + m\widetilde{w} + s\widetilde{z}_i + \frac{t_1}{B^2}\widetilde{q}_1 + \left(\frac{m'+s_a}{m'+\frac{s_a}{N}}\right)\left[m'\widetilde{w}' + \frac{1}{N}\sum_{a=1}^{N} s_a\widetilde{z}_a\right]}{\frac{hn}{h+n} + m_c + m + s + \frac{t_1}{B^2} + \frac{(m'+s_a)^2}{m'+\frac{s_a}{N}}} \\
&= \frac{\frac{hn}{h+n}(\widetilde{y} - \overline{u}) - m_c\widetilde{\nu}_c - m\widetilde{\nu} - s\widetilde{\varepsilon}_i + \frac{t_1}{B}\widetilde{x}_1 - \left(\frac{m'+s_a}{m'+\frac{s_a}{N}}\right)\left[m'\widetilde{\nu}' + \frac{1}{N}\sum_{a=1}^{N} s_a\widetilde{\varepsilon}_a\right]}{\frac{hn}{h+n} + m_c + m + s + \frac{t_1}{B^2} + \frac{(m'+s_a)^2}{m'+\frac{s_a}{N}}}, \qquad (7)
\end{aligned}$$

and also define earnings surprise based on the average investor expectations, denoted by $UE_{\Sigma i}$, as

$$\begin{aligned}
UE_{\Sigma i} &\equiv \widetilde{y} - \lim_{I\to\infty}\frac{1}{I}\sum_{i=1}^{I} E[\widetilde{y}|\widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{q}_1, \widetilde{F}_1, \cdots, \widetilde{F}_N] \\
&= \widetilde{y} - \frac{\frac{hn}{h+n}\overline{u} + m_c\widetilde{w}_c + m\widetilde{w} + s\widetilde{y} + \frac{t_1}{B^2}\widetilde{q}_1 + \left(\frac{m'+s_a}{m'+\frac{s_a}{N}}\right)\left[m'\widetilde{w}' + \frac{1}{N}\sum_{a=1}^{N} s_a\widetilde{z}_a\right]}{\frac{hn}{h+n} + m_c + m + s + \frac{t_1}{B^2} + \frac{(m'+s_a)^2}{m'+\frac{s_a}{N}}} \\
&= \frac{\frac{hn}{h+n}(\widetilde{y} - \overline{u}) - m_c\widetilde{\nu}_c - m\widetilde{\nu} + \frac{t_1}{B}\widetilde{x}_1 - \left(\frac{m'+s_a}{m'+\frac{s_a}{N}}\right)\left[m'\widetilde{\nu}' + \frac{1}{N}\sum_{a=1}^{N} s_a\widetilde{\varepsilon}_a\right]}{\frac{hn}{h+n} + m_c + m + s + \frac{t_1}{B^2} + \frac{(m'+s_a)^2}{m'+\frac{s_a}{N}}}. \qquad (8)
\end{aligned}$$

We now write two measures of the earnings surprise based on analysts' expectations. One based on an individual analyst's earnings forecast, denoted by UE_a, can be expressed as

$$\begin{aligned}
UE_a &\equiv \widetilde{y} - \widetilde{F}_a \\
&= \widetilde{y} - \frac{\frac{hn}{h+n}\overline{u} + m_c\widetilde{w}_c + m'\widetilde{w}' + s_a\widetilde{z}_a}{\frac{hn}{h+n} + m_c + m' + s_a} \\
&= \frac{\frac{hn}{h+n}(\widetilde{y} - \overline{u}) - m_c\widetilde{\nu}_c - m'\widetilde{\nu}' - s_a\widetilde{\varepsilon}_a}{\frac{hn}{h+n} + m_c + m' + s_a}. \qquad (9)
\end{aligned}$$

Also, earnings surprise based on the mean forecast, denoted by $UE_{\Sigma a}$, can be expressed as

$$
\begin{aligned}
UE_{\Sigma a} &\equiv \widetilde{y} - \frac{1}{N}\sum_{a=1}^{N}\widetilde{F}_a \\
&= \widetilde{y} - \frac{\frac{hn}{h+n}\overline{u} + m_c\widetilde{w}_c + m'\widetilde{w}' + \frac{1}{N}\sum_{a=1}^{N} s_a\widetilde{z}_a}{\frac{hn}{h+n} + m_c + m' + s_a} \\
&= \frac{\frac{hn}{h+n}(\widetilde{y} - \overline{u}) - m_c\widetilde{\nu}_c - m'\widetilde{\nu}' - \frac{1}{N}\sum_{a=1}^{N} s_a\widetilde{\varepsilon}_a}{\frac{hn}{h+n} + m_c + m' + s_a}.
\end{aligned} \tag{10}
$$

An examination of the above four measures of earnings surprise reveals two kinds of differences among the measures. First and most importantly, the pronounced difference between the hypothetical measures of earnings surprise based on unobservable investors' expectations in equations (7) and (8) and the proxy measures based on observable analysts' forecasts in equations (9) and (10) arises from the asymmetry of information between investors and analysts. This can be seen clearly from the denominators of the respective expressions. The denominators are the total precision achieved by an individual investor or analyst. The difference in the total precision is

$$
\begin{aligned}
K - K_a &= \left(\frac{hn}{h+n} + m_c + m + s + \frac{t_1}{B^2} + \frac{(m'+s_a)^2}{m' + \frac{s_a}{N}}\right) - \left(\frac{hn}{h+n} + m_c + m' + s_a\right) \\
&= m + s + \frac{t_1}{B^2} + \frac{\frac{N-1}{N}s_a(m'+s_a)}{m' + \frac{s_a}{N}},
\end{aligned}
$$

which represents the information available to an individual investor but not reflected in individual analysts' forecasts.

The second difference among the four measures is between surprise based on the expectations of an individual investor/analyst and surprise based on those of the average investor/analyst. This difference is caused by the cancelling of idiosyncratic errors in the formulation of the average. Investors' idiosyncratic errors are completely eliminated in equation (8) due to the large (i.e., infinite) number of investors, while analysts' idiosyncratic errors are only partially cancelled in equation (10) due to the small (i.e., finite) number of analysts.

In order to analyze the regressions of price change on the above four alternative measures of earnings surprise, we first consider the covariance of price change and earnings surprise.

Lemma 1 *The covariance of price change and earnings surprise is the same for all four measures of earnings surprise. That is,*

$$\begin{aligned} Cov(\widetilde{P}_2 - \widetilde{P}_1, UE_i) &= Cov(\widetilde{P}_2 - \widetilde{P}_1, UE_{\Sigma i}) \\ &= Cov(\widetilde{P}_2 - \widetilde{P}_1, UE_a) = Cov(\widetilde{P}_2 - \widetilde{P}_1, UE_{\Sigma a}) = \frac{n}{h+n} \cdot \frac{1}{K}. \end{aligned}$$

In fact, the covariance remains the same for any measure of earnings surprise based on the expectations of a person who observes a subset of the investor's information, and the same for any measure based on the average expectations of such persons possessing information of the same quality. Also, the presence of error in the measure of earnings surprise unrelated to price change does not change the covariance.

In order to understand the reasoning behind Lemma 1, consider the expression for price change in equation (5). Since the supply noise in period 2 is not correlated to earnings surprise, Lemma 1 is obtained if the covariance of the first term of equation (5) (in particular, the numerator of the bracketed term) and earnings surprise is the same for different measures of earnings surprise being considered. This turns out to be true because both price change and a measure of earnings surprise properly weight different information according to its precision. For example, a surprise based entirely on the prior can be thought of as $\widetilde{y} - \overline{u}$, and the covariance of the above expression and the measure is 1 because the weight $\frac{hn}{h+n}$ is the inverse of the variance of $\widetilde{y} - \overline{u}$. This holds for any measure based on expectations, including the individual expectations in equations (7) and (9), using different combinations of different information signals observed by an investor. The same holds for an average measure such as equations (8) and (10), because idiosyncratic errors of an individual are not correlated with price change. The same is also true for measures of earnings surprise that contain other errors or noise unrelated to price change including ones with non-zero means (i.e., biased).

The variances of different earnings surprise measures, however, differ depending on the extent of information impounded and also on the extent of price-unrelated errors contained in the measure. The following lemma states the results.

Lemma 2 *The variances of the four measures of earnings surprise are, respectively,*

$$Var(UE_i) = \frac{1}{K}, \qquad Var(UE_{\Sigma i}) = \frac{1}{K} \cdot \frac{K - s}{K},$$
$$Var(UE_a) = \frac{1}{K_a}, \qquad Var(UE_{\Sigma a}) = \frac{1}{K_a} \cdot \frac{K_a - (1 - \frac{1}{N})s_a}{K_a},$$

where $K_a \equiv \frac{hn}{h+n} + m_c + m' + s_a$ *is the precision of an individual analyst at the time of the announcement of his or her earnings forecast. These expressions show that the variance of a measure based on individual expectations is the inverse of the total precision of the individual, and the variance of a measure based on average expectations is the inverse of the total precision of one individual multiplied by a factor (less than one if the individuals possess idiosyncratic information) that accounts for the fact that idiosyncratic errors are diversified away in the averaging process.*

Lemma 2 says that the variance of the measures of earnings surprise is determined by the quality of the information possessed by those from whose expectations the surprise measure is constructed. Averaging over expectations of individuals possessing information of the same quality reduces the variance because some idiosyncratic errors are diversified away in the process. Of course, the presence of an error in the measure unrelated to price change increases the variance. The next lemma shows that the above variances can be clearly ranked. The proof is in the appendix.

Lemma 3 *The variances of earnings surprise for the four measures of earnings surprise can be ranked as*

$$Var(UE_{\Sigma i}) < Var(UE_i) < Var(UE_{\Sigma a}) < Var(UE_a)$$

for $s > 0$, $s_a > 0$, *and* $N > 1$.

The reason for the first and the third inequalities is the fact that mean expectations are more precise due to the diversification of idiosyncratic errors. The second inequality is true due to the fact that the mean expectation of analysts is still inferior to the expectations of an individual investor, because each investor observes the mean forecast.

Given the above three lemmas, we are ready to analyze the differences among regression metrics using different measures of earnings surprise which we turn to in the next section.

4 Regression Metrics for Different Measures of Earnings Surprise

In this section we consider how the use of different measures of earnings surprise affect regression metrics such as the regression coefficient and the R^2. First, consider regressing price change on earnings surprise. Two regression metrics, namely, *ERC* and the R^2, are now defined as

$$ERC \equiv \frac{Cov(\widetilde{P}_2 - \widetilde{P}_1, UE)}{Var(UE)}$$

and

$$R^2 \equiv \frac{\left[Cov(\widetilde{P}_2 - \widetilde{P}_1, UE)\right]^2}{Var(\widetilde{P}_2 - \widetilde{P}_1) \cdot Var(UE)}.$$

Also, considering the reverse regression to the above (i.e., one regressing earnings surprise on price change), the reverse regression coefficient, denoted by RRC, is defined as

$$RRC \equiv \frac{Cov(\widetilde{P}_2 - \widetilde{P}_1, UE)}{Var(\widetilde{P}_2 - \widetilde{P}_1)}.$$

We use appropriate subscripts for the above three constructs depending on the measure of surprise.

The following observation is obtained immediately from equation (5) and Lemmas 1 and 2.

Observation 1 *Consider four different sets of regressions in our model, with respect to price change and a different earnings surprise measure: one based on investors' mean expectation of earnings (with subscript Σi), one based on an individual investor's expectation of earnings (with subscript i), one based on the mean analyst forecast (with subscript Σa), and one based on an individual forecast (with subscript a). Then, the theoretical regression coefficients of the standard and the reverse regressions and the R^2s are as follows:*

$$ERC_i = \frac{n}{h+n}, \qquad ERC_{\Sigma i} = \frac{n}{h+n} \cdot \frac{K}{K-s},$$
$$ERC_a = \frac{n}{h+n} \cdot \frac{K_a}{K}, \qquad ERC_{\Sigma a} = \frac{n}{h+n} \cdot \frac{K_a}{K} \cdot \frac{K_a}{K_a - (1-\frac{1}{N})s_a},$$

$$RRC_i = RRC_{\Sigma i} = RRC_a = RRC_{\Sigma a} = \frac{1}{\frac{n}{h+n} \cdot \Psi},$$

and

$$R_i^2 = \frac{1}{\Psi}, \qquad R_{\Sigma i}^2 = \frac{1}{\Psi} \cdot \frac{K}{K - s},$$
$$R_a^2 = \frac{1}{\Psi} \cdot \frac{K_a}{K}, \qquad R_{\Sigma a}^2 = \frac{1}{\Psi} \cdot \frac{K_a}{K} \cdot \frac{K_a}{K_a - (1 - \frac{1}{N}) s_a},$$

where K and K_a are the precision of an individual investor's and an individual analyst's information, respectively, N is the number of analysts forecasting, and

$$\Psi \equiv \frac{K + s + \frac{n^2}{r^2 t_1 (h+n)^2} + \frac{K^2}{r^2 t_2 n^2}}{K},$$

where Ψ reflects the noisiness of the market due to the presence of supply noise.

Observation 1 and equation (6) reveal major differences among measures of the market reaction to an earnings disclosure often used in the literature, namely, the variance of price change, the earnings response coefficient, the R^2 of the return-earnings regression, and the reverse regression coefficient. First, the relative precision of earnings $\frac{n}{h+n}$ directly affects the variance of price change and the *ERC* (and inversely affects the reverse regression coefficients) but does not affect the R^2. This is due to the fact that the *ERC* measures the magnitude of the impact of earnings on price while R^2 is simply the fraction of the variation in price change that is explained by earnings surprise, and is thus not affected by $\frac{n}{h+n}$, a measure of the strength of the earnings impact. Second, the level of earnings uncertainty among investors $\frac{1}{K}$ has a distinctive positive effect on the variance of price change but no similar effect on the *ERC* or the R^2, again due to the fact that both the *ERC* and the R^2 are ratios of terms each of which is affected similarly by earnings uncertainty. Third, noisiness increases the variance of price change and decreases the R^2, but has no apparent effect on the *ERC*. Finally, the diversification of idiosyncratic errors has little relation to the variance of price change but has a distinctive effect on both the *ERC* and the R^2.

The following observation follows from Observation 1 and Lemma 3.

Observation 2 *The theoretical earnings response coefficients and* R^2s *of the regressions based on four measures of earnings surprise can be ranked as*

$$ERC_{\Sigma i} > ERC_i > ERC_{\Sigma a} > ERC_a \tag{11}$$

and

$$R^2_{\Sigma i} > R^2_i > R^2_{\Sigma a} > R^2_a. \tag{12}$$

for $N > 1$. *In particular, the average (aggregate) earnings surprise measures generate greater earnings response coefficients and* R^2s *than those based on individual earnings surprise measures due to the diversification of idiosyncratic errors in the mean expectations. That is,*

$$\frac{ERC_{\Sigma i}}{ERC_i} = \frac{R^2_{\Sigma i}}{R^2_i} = \frac{K}{K - s} > 1 \tag{13}$$

and

$$\frac{ERC_{\Sigma a}}{ERC_a} = \frac{R^2_{\Sigma a}}{R^2_a} = \frac{K_a}{K_a - (1 - \frac{1}{N})s_a} > 1 \tag{14}$$

for $N > 1$. *Also, the effect of using forecast proxies for investor expectations is to decrease the regression coefficient and* R^2 *by the degree to which investors possess information that analysts do not have. That is,*

$$\frac{ERC_i}{ERC_a} = \frac{R^2_i}{R^2_a} = \frac{K}{K_a} > 1. \tag{15}$$

Observations 1 and 2 constitute the main results of this paper. The unambiguous rank ordering of the earnings response coefficients and the R^2 occurs because the quality of the information reflected in the four surprise measures can be clearly ranked in the order of the aggregate investors, an individual investor (who observes the mean forecast), the aggregate analysts, and an individual analyst.

In light of the above observations, we examine in the next section their empirical implications with respect to the most commonly used return-earnings regressions in the literature and to some empirical results closely related to our findings.

5 Empirical Implications

There are several empirical implications of the model presented here. First, the model provides guidance in choosing an efficient proxy for earnings surprise and interpreting regression results based on such proxies. Since we explicitly link earnings forecasts with price change in our model, our results are much more specific and readily applicable than prior studies such as Marshak and Miyasawa (1968) and Patell (1979).[6] For example, we compare a surprise measure constructed from an individual forecast with one constructed from the mean of multiple forecasts made during a certain time period immediately before an earnings announcement. Observation 2 says that the use of mean forecasts results in a greater ERC and a greater R^2 by the ratio of $\frac{K_a}{K_a-(1-\frac{1}{N})s_a}$. This obtains because, while the covariance of price change and surprise is the same between the two measures of surprise, the variance is smaller for the mean proxy resulting in greater ERC and R^2.[7] In order to more fully utilize

[6] Patell (1979), for example, provides a set-theoretic analysis and gives sufficient conditions for which different proxies can be ranked.

[7] One caveat in applying this result is that our model ignores the variance of a surprise measure arising from the heterogeneity of the sample firms in their return/earnings/forecast generating processes. It is thus possible that a measure constructed from mean forecasts has a greater variance than one constructed from an individual forecast if such heterogeneity significantly increases the variance. However, as long as a study is designed to avoid an undue influence of heterogeneity of

our results, we recommend that a researcher examine the covariance and the variance above separately. For example, consider a situation where researchers expect analyst forecasts to be biased. To the extent that the biases are uncorrelated with returns (which happens, for example, if investors recognize and undo the biases), the biases would not affect the covariance. However, the variance may increase if biases are not uniform across the sample and constitute an additional source of noise.

Our model also provides insights into prior empirical work examining the earnings-return relation. For example, a related empirical result is from O'Brien (1988), who regresses forecast errors and time-series based earnings surprise, respectively, on abnormal returns and finds that the regression coefficient is larger for time-series based surprises. This empirical result may appear counter-intuitive. But a time-series based measure of earnings surprise can be thought of in our model as one based on a much smaller subset of investor information. Lemmas 1 and 2 show that the covariance in this case would be the same for both measures, and thus (as in Observation 1) the coefficients of the reverse regressions are the same for the two measures. Though this does not explain O'Brien's result, it points out that there is no clear reason to expect the opposite to occur.[8]

Finally, to more directly link our model to the empirical domain, we rewrite the expressions for $ERC_{\Sigma a}$ and $R^2_{\Sigma a}$ in Observation 1 (from the common return-earnings regression using the mean analyst earnings surprise as the independent variable) in

the sample, the economic factors that we have identified above would mainly determine regression metrics.

[8] Also, in a supplemental analysis not reported here we find that, if there is information impounded in analysts' forecasts but not learned by investors, the reverse regression coefficient is greater for forecast-based surprise measures based on a smaller set of information (e.g., a time-series measure) than for an analyst-based forecast error, which is consistent with O'Brien's (1988) observation. This occurs because the covariance of price change and forecast error is decreasing in the degree of information asymmetry in this direction (i.e., investors failing to fully utilize forecasts), while it is not affected by information asymmetry of the direction analyzed in this paper (Lemma 1).

terms of an individual analyst-based surprise measure, the (expected) across-analyst correlation in forecast errors ρ (or BKLS consensus), and the precision of analysts' information K_a (which BKLS also suggest how to measure):

$$ERC_{\Sigma a} = \frac{n}{h+n} \cdot \frac{Var(UE_i)}{Var(UE_{\Sigma a})} = \frac{n}{h+n} \cdot \frac{K_a}{K} \cdot \frac{1}{\left(\rho + \frac{1-\rho}{N}\right)} = ERC_a \cdot \frac{1}{\left(\rho + \frac{1-\rho}{N}\right)}, \quad (16)$$

and

$$R^2_{\Sigma a} = \frac{1}{\Psi} \cdot \frac{Var(UE_i)}{Var(UE_{\Sigma a})} = \frac{1}{\Psi} \cdot \frac{K_a}{K} \cdot \frac{1}{\left(\rho + \frac{1-\rho}{N}\right)} = R^2_a \cdot \frac{1}{\left(\rho + \frac{1-\rho}{N}\right)}. \quad (17)$$

Equations (16) and (17) identify factors that influence $ERC_{\Sigma a}$ and $R^2_{\Sigma a}$. For example, from the third equalities of the above equations, respectively, $ERC_{\Sigma a}$ and $R^2_{\Sigma a}$ are both negatively (positively) related to BKLS consensus (the number of analysts) if the corresponding earnings-return metrics based on individual forecasts (or ERC_a and R^2_a) are employed as controls.

The above equations generate additional interesting implications regarding information asymmetry between investors and analysts. In general, investors learn from earnings forecasts and, in addition, possess their own exclusive information not reflected in earnings forecasts. It is likely that information reflected in analysts' forecasts and investors' exclusive information differ in sources and are thus subject to different institutional factors. As a result, if a sample of firms is divided into two subsamples based on forecast characteristics, the two subsamples may generate different $ERC_{\Sigma a}$ and $R^2_{\Sigma a}$ due not only to possible differences in $\frac{n}{h+n}$ and $\frac{1}{\Psi}$, but also to likely differences in $\frac{Var(UE_i)}{Var(UE_{\Sigma a})}$ and/or $\frac{K_a}{K}$. We will discuss these two ratios in the remainder of this section in relation to some reported empirical observations in the extant literature.

First, consider two samples differing in the magnitude of earnings surprise measured as forecast error. Then, from the first equalities of the above equations, it is likely that the sample with large surprises would exhibit smaller $\frac{Var(UE_i)}{Var(UE_{\Sigma a})}$, and thus smaller $ERC_{\Sigma a}$ and $R^2_{\Sigma a}$. This is related to the observed non-linearities in ERCs like those reported by Freeman and Tse (1992). The typical explanation given for these non-linearities is that large earnings surprises are associated with less precise earnings announcements (i.e., lower n). Such an inverse association between n and $Var(UE_{\Sigma a})$ may be present. However, it is important to note that large surprises may also be associated with lower ERCs because large surprises reflect greater *pre-announcement* information asymmetry between investors and analysts.

Second, the measurement error when using a proxy for earnings surprise based on mean forecasts can be expressed in our model as $UE_{\Sigma a} - UE_i$. Not surprisingly, we can show that the variance of the measurement error relative to the variance of $UE_{\Sigma a}$ moves in the same direction as $1 - \frac{K_a}{K}$. This is intuitive because, on average, the relative magnitude of measurement error in a forecast-based surprise proxy is an increasing function of the degree of information asymmetry, i.e., the proportion (in terms of precision) of investors' information not possessed by analysts and thus not reflected in their forecasts. Since both $ERC_{\Sigma a}$ and $R^2_{\Sigma a}$ are increasing in $\frac{K_a}{K}$ by the second equalities of equations (16) and (17), these two regression metrics would also be decreasing functions of the magnitude of the measurement error in using forecast proxies.

Because the error in measuring investors' expectations through the mean analyst forecast is not directly observable, forecast dispersion is sometimes used as a proxy. From BKLS we can write the expected forecast dispersion as $D = (1 - \rho)/K_a$. Since D is decreasing in K_a, our model suggests that $ERC_{\Sigma a}$ may be negatively related to forecast dispersion through K_a due to the presence of measurement error. This argument is consistent with the documented negative association between forecast dispersion and *ERC*s in Imhoff and Lobo (1992) and Bartov et al. (2001) and the interpretation of such findings by Barron and Stuerke (1998).

6 Conclusion

In this paper we model the market price reaction to earnings announcements utilizing an information structure that is flexible in capturing the private and public information available to investors and analysts. This setting helps us analyze how the use of earnings surprise measures constructed from analysts' earnings forecasts as a proxy for investor earnings surprise affects the regression coefficient and the R^2 in the return-earnings regression. The results of our model provide intuition for constructing improved tests of the return-earnings relation and for correctly interpreting previously obtained empirical results. In particular, our model can assist empirical researchers in determining an appropriate metric to use when detecting and measuring the return-earnings relation using analysts' forecasts. For example, our results suggest that the variance of price change and the *ERC* are appropriate measures when attempting to capture the relative importance of earnings announcements to returns. The measures, however, suffer from different shortcomings. The variance of price change is likely to be influenced by investors' preannouncement earnings uncertainty and market noise, whereas the *ERC* is likely to be influenced by the extent to which investors have information other than that provided by analysts. In contrast, our results suggest that the R^2 is not a suitable matric for measuring the relative importance of earnings.

The results presented here provide an alternative explanation for why the *ERC* and R^2 in prior empirical studies of the return-earnings relation are generally low. We find that the *ERC* and the R^2 in the return-earnings regression using the mean analyst forecast to measure earnings surprise are inflated by the diversification of individual analyst errors and deflated by the presence of information asymmetry between investors and analysts. Thus, rather than indicating low precision or usefulness of earnings, low ERC and R^2 results could be reflecting high measurement error due to this information asymmetry between investors and analysts.

The very applied nature of our analysis and results helps formalize the existing intuition and is likely to stimulate further thought by empiricists. Nevertheless, direct applications of our results will require caution. For example, many of our variables, including market reaction variables (price change, regression coefficients and R^2) and forecast variables (forecast dispersion and mean forecasts), are related through more primitive variables in our model (h, n, m_c, m, m', s, s_a, t_1, and N). Thus, in some cases it may be difficult to design tests in which certain information variables are effectively distinguished from others and controlled for.

Appendix A

Calculation of Rational Expectations Equilibrium:

Let the linear conjecture of $\widetilde{P}_1$ and $\widetilde{P}_2$ be written as

$$\widetilde{P}_1 = \alpha_1 \overline{u} + \beta_1 \widetilde{y} + \theta_c \widetilde{w}_c + \theta \widetilde{w} + \theta' \widetilde{w}' + \delta \cdot \frac{1}{N} \sum_{a=1}^{N} s_a \widetilde{z}_a - \gamma_1 \widetilde{x}_1$$

and

$$\widetilde{P}_2 = \alpha_2 \overline{u} + \beta_2 \widetilde{y} - \gamma_2 \widetilde{x}_2.$$

Also, let

$$\widetilde{q}_1 \equiv \frac{1}{\beta_1}(\widetilde{P}_1 - \alpha_1 \overline{u} - \theta_c \widetilde{w}_c - \theta \widetilde{w} - \theta' \widetilde{w}' - \delta \cdot \frac{1}{N} \sum_{a=1}^{N} s_a \widetilde{z}_a) = \widetilde{u} + \widetilde{\eta} - B \widetilde{x}_1,$$

where $B \equiv \frac{\gamma_1}{\beta_1}$, and

$$\widetilde{q}_2 \equiv \frac{1}{\beta_2}(\widetilde{P}_2 - \alpha_2 \overline{u}) = \widetilde{u} + \widetilde{\eta} - B_2 \widetilde{x}_2,$$

where $B_2 \equiv \frac{\gamma_2}{\beta_2}$.

In period 2

$$\begin{aligned} \widetilde{\mu}_{2i} &\equiv E[\widetilde{u} | \widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{F}_1, \cdots, \widetilde{F}_N, \widetilde{P}_1, \widetilde{y}, \widetilde{P}_2] \\ &= E[\widetilde{u} | \widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{F}_1, \cdots, \widetilde{F}_N, \widetilde{q}_1, \widetilde{y}, \widetilde{q}_2] \\ &= E[\widetilde{u} | \widetilde{y}] = \frac{h\overline{u} + n\widetilde{y}}{h+n}, \end{aligned}$$

and

$$K_2 \equiv Var^{-1}[\widetilde{u} | \widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{F}_1, \cdots, \widetilde{F}_N, \widetilde{P}_1, \widetilde{y}, \widetilde{P}_2] = Var^{-1}[\widetilde{u} | \widetilde{y}] = h + n,$$

because the earnings signal released at the beginning of period 2, $\widetilde{y}$, is a sufficient statistic for signals $\widetilde{w}_c$, $\widetilde{w}$, $\widetilde{z}_i$, $\widetilde{q}_1$, and $\widetilde{q}_2$ and also for earnings forecasts, $\widetilde{F}_1, \cdots, \widetilde{F}_N$.

By the convenient properties of the normal distribution and the exponential utility function, investor i's demand in period 2 is

$$D_{2i} = rK_2(\widetilde{\mu}_{2i} - \widetilde{P}_2) = r(h\overline{u} + n\widetilde{y} - K_2\widetilde{P}_2).$$

The market clearing condition gives

$$\widetilde{x}_2 = Avr[D_{2i}] = r(h\overline{u} + n\widetilde{y} - K_2\widetilde{P}_2).$$

This can be rewritten to give us the rational expectations equilibrium price in period 2:

$$\widetilde{P}_2 = \frac{h\overline{u} + n\widetilde{y} - \frac{\widetilde{x}_2}{r}}{h+n}.$$

The period 2 price, $\widetilde{P}_2$, is simply the weighted average of the common prior about the liquidating value of the security, $\overline{u}$, and the common information contained in the earnings report about this value, $\widetilde{y}$ (less the noise term, $\frac{\widetilde{x}_2}{r}/\,(h+n)$). Again, this is because $\widetilde{y}$ is a sufficient statistic for all prior signals received by investors about these earnings.

Regarding the period 1 price, $\widetilde{P}_1$, investor i's problem at time 1 is to maximize the expected utility with respect to D_{1i}:

$$\begin{aligned}
&E[U_i(W_i)|\widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{q}_1, \widetilde{F}_1, \cdots, \widetilde{F}_N] \\
&= E\left[-\exp\left\{-\frac{1}{r}[E_i + (\widetilde{P}_2 - P_1)D_{1i} + (\widetilde{u} - \widetilde{P}_2)D_{2i}]\right\}\middle|\, \widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{q}_1, \widetilde{F}_1, \cdots, \widetilde{F}_N\right] \\
&= E\left[-\exp\left\{-\frac{E_i}{r} + (\widetilde{P_1} - \widetilde{P}_2)\frac{D_{1i}}{r} - \frac{K_2}{2}(\widetilde{\mu}_{2i} - \widetilde{P}_2)^2\right\}\middle|\, \widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{q}_1, \widetilde{F}_1, \cdots, \widetilde{F}_N\right] \\
&= E\left[-\exp\left\{-\frac{E_i}{r} + (\widetilde{P_1} - \widetilde{\mu}_{2i} + \frac{\widetilde{x}_2}{rK_2})\frac{D_{1i}}{r} - \frac{K_2}{2}(\frac{\widetilde{x}_2}{rK_2})^2\right\}\middle|\, \widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{q}_1, \widetilde{F}_1, \cdots, \widetilde{F}_N\right],
\end{aligned}$$

where W_i is investor i's wealth in period 3, $\widetilde{\mu}_{2i} \equiv E(\widetilde{u}|\widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{q}_1, \widetilde{F}_1, \cdots, \widetilde{F}_N)$, and the second equality follows from the law of iterated expectations. This expression has only two random variables, $\widetilde{\mu}_{2i}$ and $\widetilde{x}_2$ (because $\widetilde{P}_1$ has been observed at this point). However, $\widetilde{x}_2$ is independent of all other variables, and we have

$$\begin{aligned} E[\widetilde{\mu}_{2i}|\widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{q}_1, \widetilde{F}_1, \cdots, \widetilde{F}_N] &= \frac{h\overline{u}}{h+n} + \frac{n}{h+n} \cdot E[\widetilde{y}|\widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{q}_1, \widetilde{F}_1, \cdots, \widetilde{F}_N] \\ &= \frac{h\overline{u} + n\widetilde{\mu}_{1i}}{h+n}, \end{aligned}$$

and

$$Var(\widetilde{\mu}_{2i}|\widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{q}_1, \widetilde{F}_1, \cdots, \widetilde{F}_N) = \left(\frac{n}{h+n}\right)^2 \frac{1}{K},$$

where

$$\begin{aligned} K &\equiv Var(\widetilde{y}|\widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{q}_1, \widetilde{F}_1, \cdots, \widetilde{F}_N) \\ &= \frac{hn}{h+n} + m_c + m + s + \frac{t_1}{B^2} + \frac{(m' + s_a)^2}{m' + \frac{s_a}{N}} \end{aligned}$$

is the precision of an individual investor's information about the forthcoming earnings $\widetilde{y}$, and the conditional expectation of $\widetilde{u}$ (or $\widetilde{y}$) is

$$\begin{aligned} \widetilde{\mu}_{1i} &\equiv E(\widetilde{u}|\widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{q}_1, \widetilde{F}_1, \cdots, \widetilde{F}_N) \\ &= E(\widetilde{y}|\widetilde{w}_c, \widetilde{w}, \widetilde{z}_i, \widetilde{q}_1, \widetilde{F}_1, \cdots, \widetilde{F}_N) \\ &= \frac{\frac{hn}{h+n}\overline{u} + m_c\widetilde{w}_c + m\widetilde{w} + s\widetilde{z}_i + \frac{t_1}{B^2}\widetilde{q}_1 + \left(\frac{m'+s_a}{m'+\frac{s_a}{N}}\right)\left[m'\widetilde{w}' + \frac{1}{N}\sum_{a=1}^N s_a\widetilde{z}_a\right]}{\frac{hn}{h+n} + m_c + m + s + \frac{t_1}{B^2} + \frac{(m'+s_a)^2}{m'+\frac{s_a}{N}}}. \end{aligned}$$

We can now write the expected utility as (omitting terms unrelated to D_{1i})

$$-\int \exp\left[\left(\frac{D_{1i}}{r}\right)\left(\widetilde{P}_1 - \widetilde{\mu}_{2i}\right) - \frac{K}{2}\left(\frac{h+n}{n}\right)^2 \cdot \left\{\widetilde{\mu}_{2i} - \frac{h\overline{u} + n\widetilde{\mu}_{1i}}{h+n}\right\}^2\right] d\widetilde{\mu}_{2i},$$

which can be rewritten as

$$\begin{aligned} &-\exp\left[\left(\frac{D_{1i}}{r}\right)\left(\widetilde{P}_1 - \frac{h\overline{u} + n\widetilde{\mu}_{1i}}{h+n}\right) + \left(\frac{D_{1i}}{r}\right)^2 \cdot \frac{\left(\frac{n}{h+n}\right)^2}{2K}\right] \\ &\quad \times \int \exp\left[-\frac{K}{2}\left(\frac{h+n}{n}\right)^2 \cdot \left\{\widetilde{\mu}_{2i} - \frac{h\overline{u} + n\widetilde{\mu}_{1i}}{h+n} + \left(\frac{n}{h+n}\right)^2 \frac{D_{1i}}{rK}\right\}^2\right] d\widetilde{\mu}_{2i}. \end{aligned}$$

The integrand above is a multiple of a cumulative normal density, which is not a function of D_{1i}. By differentiating the exponent with respect to D_{1i} and setting it equal to zero, aggregating over investors, and equating the aggregate demand with the aggregate supply, we have

$$\widetilde{P}_1 = \frac{h\overline{u} + n\widetilde{\mu}_1}{h+n} - \left(\frac{n}{h+n}\right)^2 \frac{\widetilde{x}_1}{rK},$$

where $\widetilde{\mu}_1$ is the average $\widetilde{\mu}_{1i}$ across all investors. Price in period 1 can then be rewritten as

$$\widetilde{P}_1 = \frac{1}{K} \cdot \left(\frac{n}{h+n}\right) \cdot \Bigg[\left(\frac{K}{n} + \frac{n}{h+n}\right) h\overline{u} + \left(s + \frac{t_1}{B^2}\right)\widetilde{y} + m_c\widetilde{w}_c + m\widetilde{w}$$
$$+ \left(\frac{m' + s_a}{m' + \frac{s_a}{N}}\right) \left\{ m'\widetilde{w}' + \frac{1}{N}\sum_{a=1}^{N} s_a\widetilde{z}_a \right\} - \left(\frac{t_1}{B} + \frac{n}{r(h+n)}\right)\widetilde{x}_1 \Bigg].$$

Since the conjecture about B has to be true in equilibrium, we have

$$B = \frac{\frac{t_1}{B} + \frac{n}{r(h+n)}}{s + \frac{t_1}{B^2}}.$$

Thus, we obtain

$$B = \left(\frac{n}{h+n}\right)\frac{1}{rs}.$$

Proof of Lemma 3:

It is easy to see $Var(UE_{\Sigma i}) < Var(UE_i)$ and $Var(UE_{\Sigma a}) < Var(UE_a)$ from Lemma 2. The remaining inequality to prove is $Var(UE_i) < Var(UE_{\Sigma a})$.

$$\begin{aligned} Var(UE_i) - Var(UE_{\Sigma a}) &= \frac{1}{K} - \frac{K_a - (1-\frac{1}{N})s_a}{K_a^2} \\ &= \frac{1}{KK_a^2}\left[K_a^2 - KK_a + (1-\frac{1}{N})Ks_a\right] \\ &= \frac{-1}{KK_a^2}\left[K_a(K-K_a) - (1-\frac{1}{N})(K-K_a)s_a - (1-\frac{1}{N})K_a s_a\right]. \end{aligned}$$

The inside of the bracket above can be written as

$$
\begin{aligned}
&(K - K_a)\left(\frac{hn}{h+n} + m_c + m' + \frac{s_a}{N}\right) - (1 - \frac{1}{N})s_a\left(\frac{hn}{h+n} + m_c + m' + s_a\right) \\
&= \left(m + s + \frac{t_1}{B^2} + \frac{(1-\frac{1}{N})s_a(m'+s_a)}{m' + \frac{s_a}{N}}\right)\left(\frac{hn}{h+n} + m_c + m' + \frac{s_a}{N}\right) \\
&\qquad\qquad\qquad\qquad - (1 - \frac{1}{N})s_a\left(\frac{hn}{h+n} + m_c + m' + s_a\right) \\
&= \left(m + s + \frac{t_1}{B^2}\right)\left(\frac{hn}{h+n} + m_c + m' + \frac{s_a}{N}\right) \\
&\qquad + \frac{(1-\frac{1}{N})s_a}{m' + \frac{s_a}{N}}\left[(m' + s_a)\left(\frac{hn}{h+n} + m_c + m' + \frac{s_a}{N}\right)\right. \\
&\qquad\qquad\qquad\qquad \left. -(m' + \frac{s_a}{N})\left(\frac{hn}{h+n} + m_c + m' + s_a\right)\right] \\
&= \left(m + s + \frac{t_1}{B^2}\right)\left(\frac{hn}{h+n} + m_c + m' + \frac{s_a}{N}\right) + \frac{(1-\frac{1}{N})^2 s_a^2}{m' + \frac{s_a}{N}}\left(\frac{hn}{h+n} + m_c\right) \\
&> 0.
\end{aligned}
$$

This completes the proof of Lemma 3.

References

[1] Abarbanell, J.S., W.N. Lanen, and R.E. Verrecchia. 1995. Analysts' forecasts as proxies for investor beliefs in empirical research. *Journal of Accounting and Economics* 20: 31-60.

[2] Barron, O.E., O. Kim, S.C. Lim and D.E. Stevens. 1998. Using analysts' forecasts to measure properties of analysts' information environment. *The Accounting Review* 73 (4): 421-433.

[3] Barron, O.E. and P. Stuerke. 1998. Dispersion in analysts' earnings forecasts as a measure of uncertainty. *Journal of accounting, Auditing and Finance* 14 (Summer): 243-268.

[4] Bartov, E., S. Lynn, and J. Ronen., 2000. Return-earnings regressions: a mismeasured earnings. Working paper, New York University.

[5] Easton, P, D., and M. E. Zmijewski., 1989. Cross-sectional variations in the stock market response to accounting earnings pronouncements. *Journal of Accounting and Economics* 11: 117-141.

[6] Freeman, R.N. and S.Y. Tse. 1992. A nonlinear model of security price reponses to unexpected earnings. *Journal of Accounting Research 30 (2): 185-209.*

[7] Imhoff, E.A. and G.J. Lobo. 1992. The effect of ex ante earnings uncertainty on earnings response. *The Accounting Review* 67 (2): 427-439.

[8] Kothari, S.P., and R. G. Sloan. 1992. Information in prices about future earnings: Implications for earnings response coefficients. *Journal of Accounting and Economics* 15: 143-171.

[9] Kim, O., S. Lim, and K. Shaw. 2001. The inefficiency of the mean forecast as a summary forecast of earnings. *Journal of Accounting Research 39 (2): 329-335.*

[10] Kim, O. and R.E. Verrecchia. 1991. Trading volume and price reactions to public announcements. *Journal of Accounting Research* 29 (2): 302-321.

[11] Lipe, R. 1990. The relation between stock returns and accounting earnings given alternative information. *The Accounting Review* 65: 49-71.

[12] Marshak, J. and K. Miyasawa. 1968. Economic comparability of information systems. *International Economic Review* 9 (2): 137-174.

[13] O'Brien, P. 1988. Analysts' forecasts as earnings expectations. *Journal of Accounting and Economics* 10: 53-83.

[14] Patell, J.M. 1979. The API and the design of experiments. *Journal of Accounting Research 17 (2): 528-549.*

TABLE 1: Timeline of Events in the Model

Period 1:

1. Investor i is endowed with $\widetilde{x}_i$ risky assets, each with an expected payoff of $\overline{u}$.
2. Analysts observe: $\widetilde{w}'$ (only analysts), $\widetilde{w}_c$ (all agents), and $\widetilde{z}_a$'s (private).
3. Analysts announce forecasts F_a's.
4. Investors observe: $\widetilde{w}$ (only investors), $\widetilde{w}_c$ (all agents), $\widetilde{z}_i$'s (private), and F_a's.
5. The market opens and investors trade at price $\widetilde{P}_1$.

Period 2:

1. The firm publicly discloses its earnings report $\widetilde{y}$.
2. The market opens and investors trade at price $\widetilde{P}_2$.

Period 3:

1. The liquidating value of the firm $\widetilde{u}$ is realized.
2. The firm is liquidated.
3. Final consumption occurs.

TABLE 2: Definitions of Variables in the Model

Description of Variables	Symbol	Definition	Noise Term	Precision
Financial and Market Variables				
Terminal value of the firm (prior)	$\widetilde{u}$	(mean $\overline{u}$)		h
Earnings report of the firm	$\widetilde{y}$	$\widetilde{u}+\widetilde{\eta}$	$\widetilde{\eta}$	n
Average endowment of risky asset	$\widetilde{x}_1, \widetilde{x}_2$	(mean 0)		t_1, t_2
Risk tolerance of investors	r			
Information from price ($\widetilde{P}_1$)	$\widetilde{q}_1$			$\frac{t_1}{B^2}$
Number of investors	I	∞		
Number of analysts	N	$< \infty$		
Information Signals				
Private signals observed by investors	$\widetilde{z}_i$	$\widetilde{y}+\widetilde{\varepsilon}_i$	$\widetilde{\varepsilon}_i$	s
Private signals observed by analysts	$\widetilde{z}_a$	$\widetilde{y}+\widetilde{\varepsilon}_a$	$\widetilde{\varepsilon}_a$	s_a
Common signal observed by all agents	$\widetilde{w}_c$	$\widetilde{y}+\widetilde{\nu}_c$	$\widetilde{\nu}_c$	m_c
Common signal observed by investors	$\widetilde{w}$	$\widetilde{y}+\widetilde{\nu}$	$\widetilde{\nu}$	m
Common signal observed by analysts	$\widetilde{w}'$	$\widetilde{y}+\widetilde{\nu}'$	$\widetilde{\nu}'$	m'
Earnings Forecasts	$\widetilde{F}_a$			
Information of individual analysts				K_a
Information of individual investors				K

$$K_a \equiv \frac{hn}{h+n} + m_c + m' + s_a$$

$$K \equiv \frac{hn}{h+n} + m_c + m + s + \frac{t_1}{B^2} + \frac{(m'+s_a)^2}{m'+\frac{s_a}{N}}$$