Collection of Information About Publicly Traded Firms: Theory and Evidence

Ravi Bhushan
Sloan School of Management
Massachusetts Institute of Technology

WP# 1877-87

Massachusetts Institute of Technology
50 Memorial Drive
Cambridge, Massachusetts 02139

April, 1987
Collection of Information About Publicly Traded Firms: Theory and Evidence

Ravi Bhushan
Sloan School of Management
Massachusetts Institute of Technology

WP# 1877-87

April, 1987
COLLECTION OF INFORMATION ABOUT PUBLICLY TRADED FIRMS: THEORY AND EVIDENCE

Ravi Bhushan
Sloan School of Management
Massachusetts Institute of Technology
April, 1987

This paper is based on my dissertation completed at the University of Chicago. I am grateful to the members of my dissertation committee, Bob Holthausen, Charlie Jacklin, Shmuel Kandel, Richard Leftwich, Rob Vishny, and especially to Doug Diamond, the chairman of the committee, for all their guidance and useful comments. I would also like to thank Bhagwan Chowdhry, Paul Healy, Gur Huberman, Dick Kazarian, Pat O'Brien, Mike Ryngaert, and Katherine Schipper for their comments and suggestions.
1. INTRODUCTION

The motivation for this paper arose from the recently noted regularity in some empirical studies\(^1\) that security price reaction to earnings announcements is a function of firm size. These studies suggest that the marginal information content of announcements decreases as firm size increases, or, the informativeness of the price system is higher for larger than for smaller firms. The focus of this paper is on the information collection process about publicly traded firms in an economy. I develop a model of information collection that addresses questions like 'In equilibrium, will there be more information collected for larger than for smaller firms?' and 'Is the above-mentioned empirical regularity consistent with the predictions of theoretical models of information collection?'. The previous papers written in this area\(^2\) prove the existence of a noisy rational expectations equilibrium when information is costly. However, in these papers the source of the noise in the price system is defined as the noise in the per-capita supply of the risky asset, which is not modeled. I model this noise as the variability of liquidity-motivated trading in the shares of the firm and generate empirical predictions, that are later tested.

In sections 2 and 3, a single-period model is developed and analyzed. The model shows that supply noise, reflecting the variability of supply of a firm's shares, is positively related to the expected volume of trading in the firm. Firm size and the expected volume of trading in a firm are

\(^1\) See, for example, Richardson (1984), Ro (1984), Atiase (1985), and Freeman (1986).

positively correlated, which means that supply noise increases as firm size increases. Verrecchia shows that the informativeness of price increases as supply noise decreases [See Corollary 3, p. 1425, Verrecchia (1982a)]. This implies that informativeness of price should be inversely related to firm size, which is inconsistent with the empirical findings mentioned above. I show that one way to resolve this inconsistency is that although the informativeness of price is inversely related to the expected volume of trading in a firm, the price system may be more informative for larger than for smaller firms if the marginal cost of information collection is a decreasing function of firm size.

I show that firms with higher prior precision have less private information collected about them and have higher marginal information content of announcements. I also show that firms with higher expected volume of trading have more private information acquired about them and have higher marginal information content of announcements. It is also shown that other firm characteristics like firm size and degree of diversification of a firm influence these quantities through their impact on the cost of information collection.

In section 4, I test the predictions of the theory. I conduct cross-sectional regressions of a measure of marginal information content of earnings announcements made by firms on various firm-specific explanatory variables. I focus on quarterly earnings announcements in this paper. This study is more comprehensive than the previous studies in the following respects. I look at all the four quarterly announcements in a year, while previous studies look at either the annual or the second quarter earnings announcements. Secondly, in earlier studies the focus is either on firm size or on conducting univariate tests whereas I study the influence of the
various explanatory variables in a multiple regression context. A contribution of this study is that I am able to separate the effects that the cost of information acquisition and the benefit from information collection have on the marginal information content of earnings announcements of a firm. This is possible by having both volume of trading and size as explanatory variables in the regressions. The volume of trading measures the benefit side of information collection and the firm size captures the cost side.

My empirical results are generally consistent with theory. I find that the marginal information content of earnings announcements increases with an increase in the total dollar volume of trading. It is also found that the marginal information content of earnings announcements is inversely related to firm size, which is consistent with marginal cost of information collection being a decreasing function of firm size or size acting as a proxy for other omitted variables. The variance of residual return and the degree of diversification turn out to be insignificant in affecting this quantity.

Finally, section 5 presents the conclusions of this research.

2. MODEL

The model that I develop is a single-period model and is based on Verrecchia's (1982a). Consider an economy with one riskless bond and one risky asset (firm), both of which pay off in terms of the single consumption good in the economy at the end of the period. The numeraire in the market is the price of the riskless bond. It is common knowledge that the bond returns one unit of the consumption good at the end of the period. The return on the risky asset is unknown until the end of the period. Denote
the return over this period on the risky asset by $u$, whose realization is denoted $u$.

In addition to the uncertainty about the return of the risky asset, the per-capita supply of the risky asset is also noisy. The noise in the per-capita supply of the risky asset (hereafter, supply noise) results in the equilibrium price being a noisy aggregator of the total available information. I model the supply noise as arising out of the liquidity-motivated trading in that asset. To characterize this idea more explicitly, assume there are two classes of traders: traders in class I who invest in information collection about the risky asset and traders in class II who do not invest in information collection. The traders in class II will not invest in information collection when the cost of collecting information is quite high for these traders or when they are extremely risk averse. Assume that there are $T$ uninformed traders and $N$ informed traders. I assume that only the uninformed traders are endowed with the risky asset and that they do not actively participate in the market except to trade for liquidity reasons.\(^3\) The price of the risky asset is thus primarily determined by the actions of the informed.

Assume that the uninformed trader $j$ supplies $k_j$ shares for liquidity reasons. I assume that the random variables $k_j$'s are independently and normally distributed with expected value zero and variance $\sigma^2$. Negative

\(^3\) These assumptions are, however, not critical to the analysis. The only effect of these assumptions is to make the mean supply of the risky asset zero, so the price of the asset does not depend on its mean supply. This assumption does not lead to any loss in generality since in this paper I am more interested in information collection and less in the pricing of the risky asset.
supply corresponds to excess liquidity; on average uninformed traders' net
liquidity needs are zero. Note that I am assuming that for each uninformed
trader, \( k_j \), the number of shares supplied by him is not a function of price.
This assumption is consistent with uninformed traders placing market orders.
The above assumption considerably simplifies the analysis.

Then \( \tilde{X} \), the aggregate supply of the asset from the uninformed to the
informed is given by:

\[
\tilde{X} = k_1 + k_2 + \ldots + k_T
\]

which implies that \( \tilde{x} \), the net per-capita (where capita refers to
an informed trader in the model) supply of the asset, is given by:

\[
\tilde{x} = \frac{\tilde{X}}{N} = \frac{1}{N} \left[ k_1 + k_2 + \ldots + k_T \right]
\]

The informed traders' net per-capita demands in equilibrium should
also sum to this quantity \( \tilde{x} \). The net per-capita supply of the risky asset,
\( \tilde{x} \), has mean zero and variance \((T/N^2)\sigma^2\). Let \( V \) equal \((T/N^2)\sigma^2\) so that
\( V \) is the variance of the per-capita supply of the risky asset. Each
trader's prior beliefs about \( u \) are that it has a normal distribution
with mean \( y_0 \) and precision \( h_0 \). Before trading starts, some public
information is available to the traders and these prior beliefs include
the effect of publicly available information about the firm. Public
information is defined as the amount of prior public disclosure made by the
firm about its return.\(^4\) The public disclosure is an announcement by the
firm, before traders start their information acquisition activities, of an
observation of a random variable, \( Y_0 \) which communicates the true return \( u \)

\(^4\) In this paper, I model the public information released by the firm
as if it is given exogenously. Diamond (1985) shows that this process is
also endogenous.
perturbed by some noise $\tilde{\epsilon}_0$:

$$\tilde{Y}_0 = \tilde{u} + \tilde{\epsilon}_0$$

where $\tilde{\epsilon}_0$ is a random variable which is normally distributed with mean zero and precision $q > 0$. Thus, $q$ represents the amount or quality of public information. I assume that this public information is available costlessly to traders and that the prior precision, $h_0$, includes the effect of the public information $q$. The informed traders can acquire costly information before trading starts. This information entitles the informed trader $t$ to observe a signal $\tilde{y}_t$ for the risky asset which communicates the true return $\tilde{u}$ perturbed by some noise $\tilde{\epsilon}_t$:

$$\tilde{y}_t = \tilde{u} + \tilde{\epsilon}_t$$

where $\tilde{\epsilon}_t$ is a random variable which has a normal distribution with mean zero and precision $s_t > 0$, and is independent of the perturbations of other traders. Each informed trader uses his observation along with what he can learn from the price to form posterior estimates of $\tilde{u}$.

The cost of acquiring a signal with precision $s_t$ is represented by a continuous function $C(s_t, h_0, Z_1, \ldots, Z_n)$ which is strictly increasing and convex in $s_t$. Here $h_0$ is the prior precision of the firm's return. The other arguments $Z_1, \ldots, Z_n$ of the cost function are the characteristics of the risky asset (firm) which influence the cost of information collection about it. The prior precision of return can influence the cost of information acquisition by affecting the amount of potential information that can be collected about the firm.

Some examples of firm characteristics that may influence cost of information acquisition are firm size and the degree of diversification of the firm. Large firms have many more sources of available information than
small firms. Large firms, on average, make more public releases compared to smaller firms, and are likely to be more dispersed geographically. All these factors may lead to information collection costs being different for large and small firms. Similarly, the degree of diversification of a firm may also influence the costs of acquiring information. As an example, I consider two firms: one in just a single line of business and the other in many lines of business. For the firm in one line of business, information acquisition will be about one line of business only. However, for the firm in many lines of business, collection of information can be for one, two or even all the firm's lines of business. How much information will be collected about a particular line depends on the relative importance of that line for the firm as well as on the relative costs of information collection for the different lines.

The assumption of strictly increasing and convex behavior of $C$ in $s_t$ reflects increasing marginal cost of information collection. As the precision of information, $s_t$, increases, collecting an additional unit of precision requires increased effort and cost. By analogy, $C_s(s_t, h_0, Z_1, \ldots, Z_n)$, the marginal cost of collecting an additional unit of precision, is assumed to be increasing in the level of the prior precision, $h_0$. Thus, $C_{sh_0}(s_t, h_0, Z_1, \ldots, Z_n)$ is also assumed positive.\(^5\)

For each informed trader, the utility for wealth, $w$ is assumed to be given by the negative exponential utility function:

\(^5\) I assume that a firm with high return variability (or low prior precision) has proportionately higher amount of potential information that can be collected about it. Hence, collecting an additional unit of precision is going to be relatively effortless for a firm having high return variability. Thus, the marginal cost of information collection is lower for a firm with high return variability (low prior precision), i.e., $C_{sh_0} > 0$. 

\[ U(w) = -\exp \left(-\frac{w}{r}\right), \]

where \( r \) is trader t's constant level of risk tolerance, assumed to be same for all the informed traders.

Under these assumptions, a rational expectations equilibrium exists where the amount of information each informed trader acquires is endogenously determined and \( s^* \), the optimal level of precision collected by a trader about the risky asset is given by\(^6\):

\[
s^* = \max \left[ 0, s \mid 2 C_s (s, h_0, Z_1, \ldots, Z_n) \left[ s + h_0 + \frac{r s^2}{v} \right] = 1 \right] \tag{1}
\]

Equation (1) shows that the optimal level of precision collected about the risky asset by all traders is same. This happens because all traders have the same degree of risk tolerance and face the same cost function.

Let \( S^* \) be the total precision about the risky asset that a trader acquires in equilibrium. Then \( S^* \) is given by:

\[
S^* = h_0 + s^* + \frac{r s^*}{v} \tag{2}
\]

Equation (2) implies that the total precision that a trader acquires is the sum of the prior precision \( h_0 \), the precision of private information \( s^* \), and an expression which is common to all traders, \( \frac{r s^*}{v} \).

This latter expression, denoted \( A^* \), represents the extent to which each trader benefits from the information acquisition activities of all traders by conditioning his beliefs on price. Thus \( A^* \) is the additional precision each trader gets by using the price system. The total precision that traders obtain in equilibrium is important because as shown in the next section it is related to the marginal information content of announcement.

\(^6\) See Verrecchia (1982a) for the details.
made by the firm.

In the next section I examine the influence of different firm characteristics on the precision of private information collected about it and on the marginal information content of announcement made by it.

3. INFORMATION COLLECTION AND FIRM CHARACTERISTICS

I assume that there is a positive amount of information being collected about the firm (i.e., \( s^* \) is non-zero). Then equation (1), which gives the optimal precision that each trader collects in equilibrium, can be rewritten as:

\[
\begin{align*}
  s^* + h_0 + \frac{r^2 s^*}{2} = \frac{r}{V(s^*, h_0, Z_1, ..., Z_n)}.
\end{align*}
\] (3)

To study the influence of different firm characteristics on the precision of private information collected about the firm, I will do comparative statics on equation (3). I am also interested in a measure of the marginal information content of announcement by the firm, which I develop now. In this single-period model, there is only one announcement and that is the declaration at the end of the period by the firm of the actual realization of the return, \( \tilde{u} \). Thus, the announcement resolves all uncertainty about the return of the asset. I propose that the ratio, \( \text{var}(\tilde{u}|y_t, P)/\text{var}(\tilde{u}) \), denoted by \( M \), can be used as a measure of the marginal information content of the announcement. Here, \( y_t \) and \( P \) refer to the private information of trader \( t \) and the price of the asset respectively. This quantity \( M \) is the fraction of the variance of \( \tilde{u} \) that will be resolved by the announcement and can thus be viewed as a measure of the marginal information content of the announcement. The remaining fraction \( (1-M) \) is already impounded in the price through the information acquisition.
activities of the traders. Also note that \( \frac{s^*/h_0}{h_0} \), the ratio of the total precision that traders acquire in equilibrium to the prior precision, is just the inverse of \( M \).

### 3.1 EFFECT OF PRIOR PRECISION OF RETURN ON INFORMATION COLLECTION

In this section I show that ceteris paribus: (1) the precision of private information decreases as the prior precision of the return increases; and (2) firms with higher prior precisions have higher marginal information content of announcements.

To see how the prior precision of a firm's return influences \( s^* \), the optimal level of precision of private information collected about the firm, we want to know \( \frac{\partial s}{\partial h_0} \) or \( s^*(h_0) \). The comparative statics result from equation (3) is:

\[
\left[ 1 + \frac{rC_{ss}(s(h_0),h_0,Z_1,\ldots,Z_n)}{s^*(h_0)^2} \right] \frac{\partial s}{\partial h_0} = -1 - \frac{rC_{sh_0}(s^*(h_0),h_0,Z_1,\ldots,Z_n)}{2C_s^2(s^*(h_0),h_0,Z_1,\ldots,Z_n)} \tag{4}
\]

Since the cost function is strictly increasing and convex in \( s \), it follows that the expression in square brackets on the left hand side of equation (4) is positive. Thus, the sign of \( s^*(h_0) \) depends on the sign of the right hand side. The first term on the right hand side of equation (4) is negative and the sign of the second term depends on the sign of \( C_{sh_0}(s^*(h_0),h_0,Z_1,\ldots,Z_n) \) which is positive by assumption. Hence, \( s^*(h_0) \) is negative. Thus, keeping everything else constant, a firm with higher prior precision has lower private precision collected about it. The optimal level of precision that the traders collect is obtained at the point where the marginal cost of collecting additional information is equal to the additional benefit provided by this information. An increase in prior
precision results in an increase in the marginal cost of information collection and a decline in the additional benefit derived from that information (since the additional benefit varies inversely with the total precision) and hence the equilibrium private precision collected declines. Thus, increased availability of public information will lead to a decline in the amount of private information acquired about a security since the two are substitutes.

Another item of interest is the effect of prior precision on the marginal information content of the announcement. Recall that $S^*/h_0$ equals $1/M$. From equation (2):

$$\frac{1}{M} = 1 + \frac{s^* + r s}{h_0}$$

As $h_0$ increases, $s^*$ decreases, so in the second term on the right hand side of equation (5), the numerator decreases and the denominator increases. Thus, $1/M$ decreases as $h_0$ increases. Equivalently, as the prior precision increases, the marginal information content of the announcement increases. As the prior precision increases, there is less private information collected about the firm which leaves more room for the announcement by the firm to have information content.

### 3.2 Effect of Volume of Trading on Information Collection

Here I show that firms with higher expected volume of trading in them have higher precision of private information collected about them and have higher marginal information content of announcements. To do so, first I derive the relation between the unobservable variance of the per-capita supply and the expected volume of trading, which is estimable.
In equilibrium, the demand of the informed trader $t$ for the risky asset, after he receives his private signal $y_t$ and observes the market price $P$, is given by [See Pfleiderer (1982, p.34)]:

$$D_t (P, y_t) = r s^* \varepsilon_t + x$$ (6)

Equation (6) also gives the change in his holdings of the risky asset since the informed traders are assumed to have no endowments of the risky asset. Using the above equation, we can derive the volume of trading occurring in this market. The per-capita volume of trading (Vol) can be obtained by summing the absolute values of the changes in the holdings of all traders, both informed and uninformed, halving that quantity and dividing by $N$:

$$\text{Vol} = \frac{1}{2N} \left[ \sum_{t=1}^{N} \left| r s^* \varepsilon_t + x \right| \right] + \frac{1}{2N} \left[ \sum_{j=1}^{T} |k_j| \right]$$ (7)

The expected per-capita volume, denoted $\overline{\text{Vol}}$, is the expectation of Vol. To derive $\overline{\text{Vol}}$, I make use of the fact that if $Z$ is a normally distributed random variable with expected value zero and variance $\sigma^2$, then the expected value and variance of $|Z|$ are $(2/\pi)^{1/2}$ $\sigma$ and $(1-2/\pi)\sigma^2$. Using the fact that $V = \frac{T}{N^2} \sigma^2$ and doing some algebra:

$$\overline{\text{Vol}} = \frac{1}{(2\pi)^{1/2}} 1/2 (r s^* + V)^{1/2} + \frac{1}{(2\pi)^{1/2}} T^{1/2} \frac{1}{2} V^{1/2}$$ (8)

Equation (8) gives the expected per-capita volume in this market. It shows that the expected per-capita volume increases as $V$, the variance of per-capita supply, increases. The intuition behind this result is that an increase in $V$ is equivalent to more trading by the uninformed investors, which results in a higher expected volume. This increase in trading by

---

7 Pfleiderer (1982) also derives a similar expression for volume of trading. However, he ignores the volume of trading occurring within the uninformed sector which turns out to be important, as is seen later.
uninformed investors makes the price system noisier and hence leads to an increase in trading by the informed traders, again increasing the expected volume.

Using equation (8), the unobservable variance $\nu$ of the per-capita supply can be written in terms of the other quantities that are observable or estimable. Unfortunately, there is no closed-form expression for $\nu$ in terms of these other quantities. However, equation (8) reduces to a simple expression if $T$, the number of uninformed traders in the market, is large. Then the second term on the right hand side of equation (8) will dominate the first one and $\sqrt{\nu}$ will be roughly proportional to the standard deviation of the per-capita supply. Also, if the term $\frac{r^2 s^*}{\nu}$ is much smaller than $\nu$, then again $\sqrt{\nu}$ will be roughly proportional to $\nu^{1/2}$. This will happen if the informed traders are very risk-averse or the precision of private information collected by them is small. I make the simplifying assumption that $\sqrt{\nu}$ is proportional to $\nu^{1/2}$, or

$$\nu = \beta \sqrt{\nu^2}. \quad (9)$$

Thus, the unobservable variance of the per-capita supply of the shares of a firm can be written in terms of the estimable expected volume of trading in the firm. The results derived in this section depend on whether the approximation given in equation (9) is a good one or not. However, it is important to note that only the results of this section (i.e., about the expected volume of trading) are affected by this approximation.

To see how the expected volume of trading in a firm influences $s^*$, I want to know $\frac{\partial s^*}{\partial \sqrt{\nu}}$ or $s^*'(\nu)$. In what follows, I drop the arguments in the cost function because it makes the notation cumbersome. Substituting $\beta \sqrt{\nu^2}$ for $\nu$ in (3) and doing the comparative statics:
\[
\frac{2 \Delta^*}{s^*(Vol)} + \frac{rC_{ss}^{*'}}{(Vol)} = \frac{2 \Delta^*}{Vol} \quad (10)
\]

Since the cost function is strictly increasing and convex in \( s \),

it follows that the terms in the brackets on the left hand side of

equation (10) are positive. The sign of the right hand side is positive

since both \( \Delta^* \) and \( Vol \) are positive quantities. Hence, \( s^*(Vol) \) is positive.

An increase in the expected volume of trading in a firm is equivalent to an

increase in supply noise. This has the effect of decreasing the total

precision attained by the traders or increasing the additional benefit

derived from collecting private information. Since the marginal cost of

information collection remains unaffected and the additional benefit

increases, the equilibrium level of private information collected about the

firm increases.

To study the impact of the expected volume of trading on the

marginal information content of the announcement, recall that \( M = h_0^*/s^* \).

Using equation (3), \( M \) can be rewritten as:

\[
M = \frac{2C_s(s^*, h_0, Z_1, \ldots, Z_n)h_0}{r}
\]

so that

\[
\frac{r}{2h_0} [M'(Vol)] = C_{ss}^{*'} (Vol) \quad (12)
\]

Equation (17) shows that the sign of \( M'(Vol) \) depends on the sign of

\( s^*(Vol) \), since \( C_{ss} \) is positive by assumption. It was shown earlier

that \( s^*(Vol) \) is positive, so \( M'(Vol) \) is positive. This result implies that

as the expected volume of trading in a firm increases, the marginal

information content of the announcement increases. Ceteris paribus, an
increase in the expected volume of trading in a firm makes the price system noisier, i.e., causes the total precision to decline. This increases the benefit from collecting private information and leads to a higher level of private information collected about the firm. The above result says that this increase in the precision of private information collected about the firm is not enough (because the marginal cost of information collection increases with the level of the precision of private information) to counterbalance the decrease in the total precision caused by the increase in the expected volume of trading. Thus the total precision acquired by the traders in equilibrium decreases, or the marginal information content of the announcement increases.

3.3 EFFECT OF OTHER FIRM CHARACTERISTICS ON INFORMATION COLLECTION

In this section I study the influence of any firm characteristic \( Z_i \) on information collected about the firm and on the marginal information content of the announcement. The firm characteristic \( Z_i \) influences these quantities through its impact on the cost of information collection.

To see how the firm characteristic \( Z_i \) influences \( s^* \), I want to know \( \partial s^*/\partial Z_i \) or \( s^*(Z_i) \). Comparative statics on (3) yields:

\[
\frac{s^*(Z_i)}{2C_s^*(s^*(Z_i), h_0, Z_1, \ldots, Z_n)} \left[ 1 + \frac{rC_{s s}^*(s^*(Z_i), h_0, Z_1, \ldots, Z_n)}{2C_s^*(s^*(Z_i), h_0, Z_1, \ldots, Z_n)} \right] s^*_i (Z_i) = - \frac{rC_{sZ_i}^*(s^*(Z_i), h_0, Z_1, \ldots, Z_n)}{2C_s^*(s^*(Z_i), h_0, Z_1, \ldots, Z_n)}
\]

Equation (13) shows that the sign of \( s^*_i (Z_i) \) is the inverse of the sign of \( C_{sZ_i}^* \). Equation (13) implies that \( s^*_i (Z_i) \) will be positive if \( C_{sZ_i}^* \) is negative, i.e., the marginal cost of collecting a precision level \( s \) decreases as the firm characteristic \( Z_i \) increases. Ceteris paribus, an increase in the firm characteristic \( Z_i \) leaves the
total precision unchanged so that the additional benefit derived from collecting private information remains unchanged. However, the cost of information collection is a function of this firm characteristic. If the marginal cost of collecting additional precision decreases as $Z_i$ increases then in equilibrium more precise signals will be collected for a firm that has a higher value of $Z_i$ compared to another.

Using equation (5), it can be shown that the sign of $M'(Z_i)$ is the opposite that of $s^*(Z_i)$, which in turn is opposite that of $C_{sz_i}$. If $C_{sz_i}$ is negative, i.e., the marginal cost of collecting information decreases as $Z_i$ increases, then more private information is collected about firms with higher $Z_i$. This increase in collection of private information results in the announcements made by them having lower marginal information content as compared to those with smaller values of $Z_i$.

Firm characteristics like size and the degree of diversification are likely to affect the nature of the cost function. The above analysis shows that these characteristics influence both the amount of private information collected and the marginal information content of announcements made by firms through their impact on the cost function governing the information acquisition technology. It is difficult to characterize the dependence of the cost function on any given firm characteristic explicitly. Hence, it becomes an interesting empirical issue to study the influence of different firm characteristics on the amount of information acquired about a firm.

The empirical work is next.

4. **EMPIRICAL WORK:**

4.1 **MOTIVATION**

The model developed and analyzed in sections 2 and 3 was a single-period model and theoretical predictions were obtained in terms of the
precision of private information and the marginal information content of announcements made by firms. The precision of private information is a difficult quantity to measure. However, we can develop a proxy for marginal information content of announcements made by firms. In the single-period model, announcement by a firm resolved all uncertainty about the value of the firm. In the multi-period real world, the counterpart to that announcement is an earnings announcement by a firm since it resolves some, although not all, uncertainty about the value of the firm. Hence, a measure of the marginal information content of the earnings announcements made by a firm can be used as a suitable proxy.

One measure of the marginal information content of an earnings announcement that has been used in the previous studies [See, for example, Beaver (1968)] is the ratio of an estimate of the variance of the security's residual return during the announcement period to that for the non-announcement period. I use a similar measure, which I discuss in section 4.2. The earlier studies in this area [e.g., Ball and Brown (1968), Beaver (1968), and May (1971)] tried to examine if earnings announcements have information content. The later studies focus on studying the cross-sectional differences in price reactions to earnings announcements [See, for example, Atiase (1985), Freeman (1986), Richardson (1984), and Ro (1984)].

My objective is also to explain the cross-sectional variation in price reactions to earnings announcements. This study is more comprehensive in the following respects. First, I look at all four quarterly announcements in a year while those cited above look at either the annual or the second quarter earnings announcements. Secondly, these studies conduct univariate tests on firm size or other variables whereas the present study looks at the various explanatory variables in a multiple regression context. The
explanatory variables I choose are proxies for firm size, the expected volume of trading in the firm, the variance of residual return of the firm and the degree of diversification of the firm. Since size is correlated with all these variables (see, e.g., table 3), it is possible that in these studies size may be capturing the effect of these other variables. Finally, by including proxies for both size and the expected volume of trading, I am able to separate the effects of the costs from the effects of the benefits of information acquisition on the marginal information content of earnings announcements. In the regression, size captures the costs and volume of trading capture the benefits.

4.2 DESCRIPTION OF DATA

For each firm in the sample, quarterly earnings announcement dates were collected from the Compustat tapes for the years 1977 through 1981, for twenty quarterly announcements per firm. To be included in the sample, a firm had to meet the following criteria: (1) be listed on the daily return tapes constructed by CRSP at the University of Chicago for the period of 1976-1982; (2) be on the Compustat tapes; (3) have a December 31 fiscal year end; (4) have the data of the 4-digit SIC codes available in the Standard and Poor's Register of Corporations and the Directory of Corporate Affiliates. The first three criteria were applied to get easy access to returns and appropriate financial data necessary for doing the empirical work. Criterion (4) was used to get data on the degree of diversification of firms. A total of 946 firms met the above criteria and the criterion for the successful estimation of the market model, which is discussed below.

The market model was used to eliminate the marketwide elements of price change:

\[ \tilde{R}_{it} = \alpha_i + \beta_i \tilde{R}_{mt} + \tilde{\epsilon}_{it} \]
where $\tilde{R}_{it}$ is the continuously compounded return for security $i$ on day $t$, $\tilde{R}_{mt}$ is the continuously compounded return for the value-weighted market portfolio on day $t$ and $\epsilon_{it}$ is the stochastic idiosyncratic component of $\tilde{R}_{it}$. For each firm and for every year in the study, I pool all the four quarterly observations in that year for the firm.\textsuperscript{8} Pooling observations for four quarters gives four times as many observations to estimate the announcement period variance than using unpooled data. Thus, the estimate of the announcement period variance is likely to be improved. Pooling also reduces the problem of cross-sectional dependence in the observations across firms induced by clustering of announcement dates for different firms. The problem of cross-firm correlation is not that serious when working with daily data\textsuperscript{9} and process of pooling reduces this correlation even further.

Let $q^1_0$, $q^2_0$, $q^3_0$, and $q^4_0$ denote the announcement dates of earnings for the first, second, third, and the fourth quarters respectively for a firm for a given year. Then the following two non-overlapping periods were defined. Period (1): Days $q^1_0-39$, $q^1_0-37$, $q^1_0-35$, $q^4_0+17$, $q^4_0+19$; and Period (2): Days $q^1_0-38$, $q^1_0-36$, $q^1_0-34$, $q^4_0-18$, $q^4_0+20$. This gives us a total of about 125-130 non-overlapping trading days for each period. A firm with less than 100 days for either period was dropped from pooling for a longer time period may be undesirable because then one can run into problems arising due to non-stationarity.

\textsuperscript{8} Bernard (1986) reports the mean intra-industry cross-sectional correlation of 0.04 and mean inter-industry correlation of 0.013 for daily data.
the sample. Let $a_i$ and $b_i$ be the Scholes-Williams\textsuperscript{10} estimates of $\alpha_i$ and $\beta_i$ respectively over period 1. The estimated coefficients $a_i$ and $b_i$ were used to compute $e_{it}$'s, the daily residuals, over both periods as follows:

$$e_{it} = R_{it} - (a_i + b_i R_{mt})$$ \hspace{1cm} (14)

Let $\sigma_{i1}^2$ and $\sigma_{i2}^2$ be the estimates of $\sigma_i^2$, the variance of $e_{it}$'s, over periods 1 and 2 respectively. Then since period 1 and period 2 do not have any days in common, the errors in $\sigma_{i1}^2$ and $\sigma_{i2}^2$ are uncorrelated and hence $\sigma_{i1}^2$ and $\sigma_{i2}^2$ are independent estimates of $\sigma_i^2$. Let $\sigma_{i1}^{*2}$ be the estimate of the variance of a firm's daily residual return during the announcement period. I use the ratio of $\sigma_{i1}^{*2}$ to $\sigma_{i1}^2$ as the dependent variable in the marginal information content regression, and $\sigma_{i2}^2$ as one of the explanatory variables (proxy for residual return variance) in the regression. This procedure eliminates any spurious correlation between the dependent variable and the explanatory variable since $\sigma_{i1}^2$ and $\sigma_{i2}^2$ are independent estimates of the same quantity, the residual return variance. If the announcement period falls outside the estimation period, then any non-stationarity in data can also result in spurious negative correlation between the dependent variable and the independent variable $\sigma_{i2}^2$. Here, by construction, the announcement period as well as the non-announcement period are part of the estimation period, which eliminates this problem.

To estimate $\sigma_{i1}^{*2}$, the announcement period variance, it was assumed that if day 0 refers to the quarterly announcement date as appearing on the

---

\textsuperscript{10} Tests were also conducted by estimating the market parameters without using the Scholes-Williams correction and the results of those tests are very similar to those reported here.
Compustat tape, then the announcement effect is spread over days -1, 0 and +1. This window of ± 1 days around the reported date on the tape is very likely to capture the actual announcement date.\textsuperscript{11} \( \frac{-\hat{\sigma}_i^2}{\hat{\sigma}_{i1}^2} \) was obtained as:

\[
(1/12) \left[ \sum_{t=q_i^1}^{q_i^0-1} e_{it}^2 + \sum_{t=q_i^2}^{q_i^0} e_{it}^2 + \sum_{t=q_i^3}^{q_i^0+1} e_{it}^2 + \sum_{t=q_i^4}^{q_i^0+1} e_{it}^2 \right] (15)
\]

As discussed earlier, the ratio \( \frac{-\hat{\sigma}_i^2}{\hat{\sigma}_{i1}^2} \), denoted M, is a metric for the marginal information content of earnings announcement.\textsuperscript{12} Panel A of Table 1 reports the deciles for the marginal information content, M. \(~\overline{M}\), the mean of M is 1.36. The median value for M is 1.11 and about 55% of the firms in the sample have M greater than 1.00. Panel B of Table 1 provides the descriptive statistics for M for each of the years 1977 through 1981 covered by the study. For each year the mean and the median values for the marginal information content are greater than one and there is no strong pattern evident in the behavior of the marginal information content variable from year to year.

\textsuperscript{11} A comparison of the announcement date as reported on the tape with the Wall street Journal announcement date for a few firms revealed that for most of them, the difference between the two dates was one trading day and in a few cases, the Compustat date preceded the Wall Street Journal date by two trading days. Hence, the window of ± 1 days of the reported date on the tape is very likely to capture the actual announcement date.

\textsuperscript{12} Tests were also conducted using an estimate of the announcement period variance, which was calculated using the estimated mean of the announcement period residuals from the sample of 12 observations available, instead of assuming this mean to equal zero. Results from these tests were similar to those reported.
My main interest is to explain the cross-sectional variation in $M$. The explanatory variables chosen are proxies for firm size, the expected volume of trading in the firm, the variance of residual return, and the degree of diversification of the firm. The proxy for firm size is the market value of the equity outstanding on the last trading day of the previous year, obtained from the Compustat tapes. The proxy for the expected volume of trading in a firm is the dollar volume of trading in the firm in the previous year, obtained by multiplying the year-end closing price by the number of shares traded during that year as reported on the Compustat tapes. The proxy for the residual return variance is $\sigma_{12}^2$, the estimated variance of the residuals during the estimation period (period 2). Three different proxies were chosen for the degree of diversification of a firm: (1) the number of 4-digit SIC codes associated with the firm as appearing in the Standard and Poor's Register of Corporation, (2) the number of 4-digit SIC codes and (3) the number of 3-digit SIC codes associated with the firm as appearing in the Directory of Corporate Affiliates. These alternate proxies were denoted LOB1, LOB2, and LOB3 respectively.

Table 2 provides some relevant descriptive statistics on these variables. The firms in the sample are relatively big: the median firm has $164$ million in equity outstanding and the median yearly dollar trading volume is $45$ million. The LOB statistics reveal that the median number of 4-digit codes associated with a firm is four.

Table 3 provides the simple correlation matrix for the above variables. $\text{Ln(value)}$ and $\text{ln(volume)}$ are highly positively correlated showing that larger firms have higher dollar trading volume. $\text{Ln(value)}$ and $\text{Res. Var.}$ have a correlation of $-0.53$ indicating that larger firms have, in general, lower residual variance of daily return. The correlation between $\text{ln(value)}$
and any of the LOB variables is about 0.25 which shows that on average larger firms are more diversified than smaller ones. The small magnitude of this correlation may be due to none of the LOB variables being a good proxy for the degree of diversification. The correlation between the residual variance and any of the LOB variables is about -0.15 showing that more diversified firms have lower residual variance. The low correlation between LOB1 and LOB2 shows that there is considerable disparity in the ways the Standard and Poor's Register of Corporations and the Directory of Corporate Affiliates arrive at their decisions on the number of 4-digit SIC codes associated with a firm.

There may be important industry differences across firms in terms of private information acquisition: the cost functions governing information collection may be different across different industries or there may be systematic differences across industries in the impact the earnings announcements have on the value of a firm. This may lead to differences across industries in the marginal information content of earnings announcements even when the effect of everything else is kept constant. To control for these industry differences, firms were classified into six separate industries on the basis of their primary line of business. The industry groups (2-digit SIC codes in parentheses) were: (1) Mining (10-14), (2) Construction and Manufacturing (15-39), (3) Transportation, Communication and Other Public Utilities (40-49), (4) Wholesale and Retail Trade (50-59), (5) Finance, Insurance, and Real Estate (60-67), and (6) Services (70-96). The number of firms in the six industry groups were 62, 544, 156, 42, 99, and 43 respectively.
4.3 RESULTS

Ordinary Least Squares (OLS) regressions were run with the marginal information content of earnings announcement as a dependent variable on the different explanatory variables. The approach chosen to control for industry differences was to use dummy variables for the six industry groups in the regression. The regression specification used was:

\[ \ln[M(i)] = b_0 + b_1 \ln[(\text{value})(i)] + b_2 \ln[(\text{volume})(i)] + b_3 \sigma_{12}^2 + b_4 \text{Proxy for the degree of diversification (i)} + \sum_{j=1}^{5} \delta_j I_j(i) + e(i), \]

where \( M(i) \) stands for the marginal information content of earnings announcement and \( I_j \)'s are dummy variables for the respective industry categories. The effect of the sixth category, \( I_6 \) (Services) is captured in the constant \( b_0 \). The diagnostic checks of this regression indicated no apparent violations of the OLS regression assumptions. The residuals from the regression appeared to be normally distributed.\(^{13}\)

Table 4 reports the results of this regression. The three columns in the table correspond to the three proxies for the degree of diversification. A comparison of the coefficients in the three columns reveals that the results are not sensitive to the choice of the proxy used for the degree of diversification. The estimated coefficient of \( \ln(\text{value}) \) is significantly negative at the 1% level while the estimated coefficient of the \( \ln(\text{volume}) \) is significantly positive at the 1% level. The estimated coefficient of \( \sigma_{12}^2 \) is negative with a t-statistic of -1.10 and each of the three LOB variables has a t-statistic close to zero.

\(^{13}\) For example, a Kolmogorov-Smirnov goodness of fit test to check the normality of the residuals from the regression generated a statistic of 0.97 with a 2-tailed p-value of 0.30.
The observed sign on the coefficient of $\ln(\text{volume})$ is consistent with the theory. An increase in the total dollar volume of trading is equivalent to an increase in the variability of liquidity trading in the firm. This makes the price system noisier so that the benefit from information acquisition goes up and hence more private information is collected about the asset. However, this increased acquisition of private information does not completely counteract the effect of the increased noise in the price system. Therefore, the marginal information content of the announcement increases as the total dollar volume of trading increases.

The negative sign on the coefficient of $\ln(\text{value})$ is consistent with the story that the marginal cost of information collection decreases as the firm size increases. Another interpretation is that firm size proxies for other omitted variables (e.g., the amount of publicly available information) and the combined effect of all these other variables on the marginal information content of announcements is negative.

Section 3 predicts that the earnings announcements have less information for firms with higher variance of residual return. Higher variance of residual return leads to more private information collected about the firm, resulting in the firm's announcements to have lower marginal information content. Thus a negative sign is predicted on the coefficient of $\hat{\sigma}_{12}^2$. However, the coefficient does not turn out to be significant. One possible reason why the coefficient turns out to be insignificant is that $\hat{\sigma}_{12}^2$ is an estimate of the residual return variance, measured with error. The errors in variable (assuming that the error here is uncorrelated with the other explanatory variables, which is a reasonable assumption) will result in an underestimate of the absolute magnitude of this coefficient, consequently lowering its t-statistic.
The coefficients of the LOB variables turn out to be insignificantly different from zero. Hence, I conclude that the degree of diversification does not significantly affect the marginal information content of earnings announcements. One reason for this may be that the proxies chosen do not adequately measure the extent of diversification of a firm.

A partial F-test to check for the significance of the industry dummy variables as a group yielded a F-statistic of 5.17, which is significant at the 1% level, indicating that the industry dummy variables add to the explanatory power of the model. This indicates that there are differences across industries in the marginal information content of earnings announcements. This could result from the cost functions governing the information collection being different across different industries or there being systematic differences across industries in the impact earnings announcements have on the value of a firm.

In general, with such a large number of observations, any model misspecification can generate statistically significant coefficients. For example, the market model used to compute the residuals may be misspecified and the misspecification systematically related to one or more of the explanatory variables. Then the estimated coefficients can be different from zero under the null hypothesis of no effect of the explanatory variables. Similarly, in the calculation of M, the numerator is estimated with many fewer observations than the denominator. This procedure may introduce a systematic bias in the estimated coefficients if, for example, there are correlated measurement errors in returns. To examine the possibility of any of the observed empirical relations being spurious, the above regressions were repeated for non-announcement period. A procedure similar to that used for calculating the marginal information content during the announcement
period was employed. Squared residuals were averaged across each three-day window starting from day -10 up to day +10 for all the four quarters. Thus, the first window covers days -10 to -8 and the last one from days +8 to +10. Choosing three-day windows during the non-announcement period makes the non-announcement period results directly comparable with those for the announcement period.

Let \( Y(i) \) denote the counterpart of the marginal information content for any window. Table 5 presents for the non-announcement period the results of the regressions of \( \ln[Y(i)] \) on the same explanatory variables as for the marginal information content regressions. The proxy chosen for the degree of diversification is LOBl. For all of the windows, most of the observed coefficients associated with the various explanatory variables are insignificantly different from zero and the significant ones have signs opposite to those observed for the announcement period regression. The results for the non-announcement period provide a benchmark for the validation of the results for the announcement period since the approach adopted for doing the regressions for the non-announcement period is the same as that for the announcement period. The effect of any kind of model misspecification would have been captured in the the non-announcement period results. These regressions demonstrate that the results are not a consequence of any model misspecification. Specifically, the approach adopted to take care of the spurious negative dependence between the dependent variable and the explanatory variable, \( \sigma_{12}^2 \), seems to have been successful. Hence, the observed relations are not spurious.

5. CONCLUSIONS

Some recent studies have documented that the marginal information content of earnings announcements is related to firm size. This motivated
me to examine the information collection process about publicly traded firms theoretically and empirically.

I modelled the noise in the price system as the variability of liquidity-motivated trading in the shares of a firm. I showed that this noise is positively related to the expected volume of trading in the firm. The expected volume of trading in the firm is estimable and can thus be used as a proxy for the unobservable noise. It was shown that firms with higher prior precision have less private information collected about them and have higher marginal information content of announcements. Firms with higher expected volume of trading have more information collected about them and have higher marginal information content of announcements. Other firm characteristics like size and the degree of diversification influence the amount private information acquisition and the marginal information content of the announcements through their effect on the costs of information collection.

I used both firm size and the volume of trading in the firm along with other explanatory variables to explain the cross-sectional variation in the marginal information content of earnings announcements. The idea behind this approach is that we can separate the effect of cost from the effect of benefit of information collection using firm size as a proxy for the cost and the volume of trading as a proxy for the benefit. This is in contrast to prior studies that use only firm size and thus cannot separate the two effects. It was found that the marginal information content of earnings announcements decreases with an increase in firm size and increases as the dollar volume of trading in the firm increases. The result for volume of trading is consistent with the predictions of the theory. The inverse relation between firm size and the marginal information content of the
earnings announcements is consistent with the marginal cost of information collection being a decreasing function of the firm size or firm size acting as a proxy for other omitted variables whose combined effect on the marginal information content of announcements is negative.
REFERENCES


TABLE 1

DATA ON MARGINAL INFORMATION CONTENT, M*  

PANEL A: DECILES FOR THE MARGINAL INFORMATION CONTENT

<table>
<thead>
<tr>
<th>Decile</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.493</td>
</tr>
<tr>
<td>0.2</td>
<td>0.668</td>
</tr>
<tr>
<td>0.3</td>
<td>0.802</td>
</tr>
<tr>
<td>0.4</td>
<td>0.949</td>
</tr>
<tr>
<td>0.5</td>
<td>1.107</td>
</tr>
<tr>
<td>0.6</td>
<td>1.288</td>
</tr>
<tr>
<td>0.7</td>
<td>1.520</td>
</tr>
<tr>
<td>0.8</td>
<td>1.856</td>
</tr>
<tr>
<td>0.9</td>
<td>2.470</td>
</tr>
</tbody>
</table>

PANEL B: DESCRIPTIVE STATISTICS FOR MARGINAL INFORMATION CONTENT FOR EACH YEAR OVER THE PERIOD 1977-1981

<table>
<thead>
<tr>
<th>YEAR</th>
<th>NOBS.</th>
<th>MEAN</th>
<th>MEDIAN</th>
<th>STD. DEV.</th>
<th>MIN. OBS.</th>
<th>MAX. OBS.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>902</td>
<td>1.51</td>
<td>1.20</td>
<td>1.48</td>
<td>0.11</td>
<td>25.2</td>
</tr>
<tr>
<td>1978</td>
<td>917</td>
<td>1.44</td>
<td>1.22</td>
<td>0.91</td>
<td>0.07</td>
<td>8.00</td>
</tr>
<tr>
<td>1979</td>
<td>925</td>
<td>1.33</td>
<td>1.09</td>
<td>1.02</td>
<td>0.14</td>
<td>12.42</td>
</tr>
<tr>
<td>1980</td>
<td>933</td>
<td>1.21</td>
<td>1.01</td>
<td>0.85</td>
<td>0.09</td>
<td>7.74</td>
</tr>
<tr>
<td>1981</td>
<td>938</td>
<td>1.31</td>
<td>1.04</td>
<td>1.07</td>
<td>0.08</td>
<td>12.70</td>
</tr>
</tbody>
</table>

The marginal information content, M, is defined as the ratio of the estimate of the variance of the daily residuals during the announcement period to the estimate of the variance of the daily residuals during the estimation period, and is thus a measure of the amount of information revealed as a result of the announcement. The total number of observations on M is 4615 (946 firms and up to 5 years of data on each) and the time period spanned by the study is 1977-81.
TABLE 2

DESCRIPTIVE STATISTICS OF THE DATA USED IN THE REGRESSIONS
OF THE MARGINAL INFORMATION CONTENT\(^a\)

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>MEAN</th>
<th>MEDIAN</th>
<th>STD. DEV.</th>
<th>MIN. OBS.</th>
<th>MAX. OBS.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MARGINAL INFO. CONTENT, N</td>
<td>1.36</td>
<td>1.11</td>
<td>1.09</td>
<td>0.07</td>
<td>25.2</td>
</tr>
<tr>
<td>VALUE ($ 000's)</td>
<td>647726.0</td>
<td>163721.2</td>
<td>2075080.0</td>
<td>570.0</td>
<td>43524248.0</td>
</tr>
<tr>
<td>LN(VALUE)</td>
<td>11.88</td>
<td>12.01</td>
<td>1.84</td>
<td>6.35</td>
<td>17.59</td>
</tr>
<tr>
<td>VOLUME ($ 000's)</td>
<td>189165.4</td>
<td>45359.6</td>
<td>463394.3</td>
<td>83.2</td>
<td>9456651.0</td>
</tr>
<tr>
<td>(\sigma_{12}^2)</td>
<td>0.000535</td>
<td>0.000326</td>
<td>0.000630</td>
<td>0.000027</td>
<td>0.00937</td>
</tr>
<tr>
<td>LOB1</td>
<td>5.40</td>
<td>4.00</td>
<td>5.51</td>
<td>1.0</td>
<td>41.0</td>
</tr>
<tr>
<td>LOB2</td>
<td>4.40</td>
<td>4.00</td>
<td>3.27</td>
<td>1.0</td>
<td>22.0</td>
</tr>
<tr>
<td>LOB3</td>
<td>3.87</td>
<td>3.00</td>
<td>2.72</td>
<td>1.0</td>
<td>18.0</td>
</tr>
</tbody>
</table>

\(\sigma_{12}^2\) is an estimate of the residual variance and LOB1, LOB2, and LOB3 are three alternate proxies for the degree of diversification of the firm. The number of observations on each variable is 4615 (946 firms and up to 5 years of data on each) and the time period spanned by the study is 1977-81.
TABLE 3

CORRELATION MATRIX FOR THE VARIABLES INVOLVED IN THE REGRESSIONS OF THE MARGINAL INFORMATION CONTENTa

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>M</th>
<th>LN(VALUE)</th>
<th>LN(VOLUME)</th>
<th>$\sigma_{12}^2$</th>
<th>LOB1</th>
<th>LOB2</th>
<th>LOB3</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LN(VALUE)</td>
<td>-0.010</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LN(VOLUME)</td>
<td>-0.081</td>
<td>0.93</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_{12}^2$</td>
<td>0.038</td>
<td>-0.53</td>
<td>-0.45</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOB1</td>
<td>-0.008</td>
<td>0.23</td>
<td>0.23</td>
<td>-0.14</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.305 )</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOB2</td>
<td>-0.025</td>
<td>0.30</td>
<td>0.29</td>
<td>-0.15</td>
<td>0.48</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.045)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>LOB3</td>
<td>-0.027</td>
<td>0.30</td>
<td>0.29</td>
<td>-0.16</td>
<td>0.46</td>
<td>0.97</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>

a The marginal information content, $M$, is defined as the ratio of the estimate of the variance of the daily residuals during the announcement period to the estimate of the variance of the daily residuals during the estimation period, $VALUE$ is the market value of equity at the end of previous year, $VOLUME$ is the dollar volume of trading in the previous year, $\sigma_{12}^2$ is an estimate of the residual variance and $LOB1$, $LOB2$, and $LOB3$ are three alternate proxies for the degree of diversification of the firm. The number of observations on each variable is 4615 (946 firms and up to 5 years of data on each) and the time period spanned by the study is 1977-81. The numbers in parentheses under the correlation coefficients indicate the 1-tailed significance probabilities.
### TABLE 4

**REGRESSION RESULTS FOR THE MARGINAL INFORMATION CONTENTa OF EARNINGS ANNOUNCEMENTS**

<table>
<thead>
<tr>
<th>Proxy used for the degree of diversification:</th>
<th>LOB1</th>
<th>LOB2</th>
<th>LOB3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent Variables:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.48</td>
<td>0.48</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>(5.33)</td>
<td>(5.33)</td>
<td>(5.33)</td>
</tr>
<tr>
<td>Ln(Value)</td>
<td>-0.060</td>
<td>-0.060</td>
<td>-0.060</td>
</tr>
<tr>
<td></td>
<td>(2.45)</td>
<td>(2.46)</td>
<td>(2.46)</td>
</tr>
<tr>
<td>$\sigma_i^2$</td>
<td>-20.15</td>
<td>-20.09</td>
<td>-20.08</td>
</tr>
<tr>
<td>Proxy for degree of divers.</td>
<td>-0.0003</td>
<td>-0.0012</td>
<td>-0.0016</td>
</tr>
<tr>
<td></td>
<td>(1.25)</td>
<td>(1.27)</td>
<td>(1.27)</td>
</tr>
<tr>
<td>$I_1$ (Mining)</td>
<td>-0.071</td>
<td>-0.073</td>
<td>-0.073</td>
</tr>
<tr>
<td></td>
<td>(1.01)</td>
<td>(1.01)</td>
<td>(1.01)</td>
</tr>
<tr>
<td>$I_2$ (Const. and Mnfg.)</td>
<td>0.045</td>
<td>0.045</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>(1.07)</td>
<td>(1.07)</td>
<td>(1.07)</td>
</tr>
<tr>
<td>$I_3$ (Utilities)</td>
<td>-0.077</td>
<td>-0.076</td>
<td>-0.078</td>
</tr>
<tr>
<td></td>
<td>(1.17)</td>
<td>(1.15)</td>
<td>(1.15)</td>
</tr>
<tr>
<td>$I_4$ (Wholesale and Retail Trade)</td>
<td>0.072</td>
<td>0.071</td>
<td>0.071</td>
</tr>
<tr>
<td></td>
<td>(1.15)</td>
<td>(1.15)</td>
<td>(1.15)</td>
</tr>
<tr>
<td>$I_5$ (Financial Institutions)b</td>
<td>-0.019</td>
<td>-0.019</td>
<td>-0.019</td>
</tr>
<tr>
<td></td>
<td>(1.00)</td>
<td>(1.00)</td>
<td>(1.00)</td>
</tr>
<tr>
<td>F-stat.</td>
<td>7.80</td>
<td>7.81</td>
<td>7.81</td>
</tr>
</tbody>
</table>

**Note:**

- The regressions are of the form:
  \[
  \ln[M(i)] = b_0 + b_1 \ln[\text{Value}(i)] + b_2 \ln[\text{Volume}(i)] + b_3 \sigma_i^2 + b_4 [\text{Proxy for degree of divers.}(i)] + \sum_{j=1}^{5} \delta_j I_j(i) + e(i),
  \]
  where the number of observations for the regression is 4615 and the t-statistics are in parentheses.

- The effect of the sixth category, $I_6$ (Services) is captured in the constant $b_0$.
<table>
<thead>
<tr>
<th>Window (Days)</th>
<th>-10 to -8</th>
<th>-7 to -5</th>
<th>-4 to -2</th>
<th>+2 to +4</th>
<th>+5 to +7</th>
<th>+8 to +10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.60</td>
<td>-0.35</td>
<td>-0.26</td>
<td>-0.05</td>
<td>-0.28</td>
<td>-0.48</td>
</tr>
<tr>
<td></td>
<td>(-6.90)</td>
<td>(-3.95)</td>
<td>(-2.85)</td>
<td>(-0.50)</td>
<td>(-3.19)</td>
<td>(-5.60)</td>
</tr>
<tr>
<td>Ln(Value)</td>
<td>0.017</td>
<td>0.017</td>
<td>-0.004</td>
<td>0.019</td>
<td>0.023</td>
<td>0.032</td>
</tr>
<tr>
<td></td>
<td>(1.22)</td>
<td>(1.18)</td>
<td>(-0.26)</td>
<td>(1.80)</td>
<td>(1.66)</td>
<td>(2.23)</td>
</tr>
<tr>
<td>Ln(Volume)</td>
<td>0.016</td>
<td>-0.002</td>
<td>0.010</td>
<td>-0.023</td>
<td>-0.011</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(1.30)</td>
<td>(-0.18)</td>
<td>(0.85)</td>
<td>(-1.87)</td>
<td>(-0.91)</td>
<td>(-0.79)</td>
</tr>
<tr>
<td>$\sigma_{i2}$</td>
<td>39.70</td>
<td>8.40</td>
<td>33.93</td>
<td>-27.15</td>
<td>19.92</td>
<td>54.99</td>
</tr>
<tr>
<td></td>
<td>(2.28)</td>
<td>(0.48)</td>
<td>(1.89)</td>
<td>(-1.50)</td>
<td>(1.14)</td>
<td>(3.14)</td>
</tr>
<tr>
<td>LOB1</td>
<td>0.001</td>
<td>-0.0001</td>
<td>0.001</td>
<td>0.003</td>
<td>-0.002</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.34)</td>
<td>(-0.05)</td>
<td>(0.79)</td>
<td>(1.47)</td>
<td>(-0.92)</td>
<td>(-0.90)</td>
</tr>
<tr>
<td>$I_1$ (Mining)</td>
<td>0.070</td>
<td>-0.029</td>
<td>-0.086</td>
<td>0.001</td>
<td>-0.045</td>
<td>-0.054</td>
</tr>
<tr>
<td></td>
<td>(1.28)</td>
<td>(-0.52)</td>
<td>(-1.52)</td>
<td>(0.02)</td>
<td>(-0.83)</td>
<td>(-0.97)</td>
</tr>
<tr>
<td>$I_2$ (Const. and Mnfg.)</td>
<td>-0.010</td>
<td>-0.082</td>
<td>-0.032</td>
<td>0.029</td>
<td>-0.026</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>(-0.25)</td>
<td>(-1.89)</td>
<td>(-0.73)</td>
<td>(0.66)</td>
<td>(-0.60)</td>
<td>(0.33)</td>
</tr>
<tr>
<td>$I_3$ (Utilities)</td>
<td>0.075</td>
<td>-0.008</td>
<td>-0.006</td>
<td>-0.022</td>
<td>-0.021</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td>(1.58)</td>
<td>(-0.16)</td>
<td>(-0.13)</td>
<td>(-0.46)</td>
<td>(-0.45)</td>
<td>(0.70)</td>
</tr>
<tr>
<td>$I_4$ (Wholesale and Retail Trade)</td>
<td>-0.032</td>
<td>-0.016</td>
<td>-0.059</td>
<td>0.109</td>
<td>0.057</td>
<td>0.036</td>
</tr>
<tr>
<td></td>
<td>(-0.56)</td>
<td>(-0.27)</td>
<td>(-0.98)</td>
<td>(1.80)</td>
<td>(0.97)</td>
<td>(0.61)</td>
</tr>
<tr>
<td>$I_5$ (Financial Institutions)</td>
<td>0.018</td>
<td>0.042</td>
<td>0.049</td>
<td>0.005</td>
<td>-0.020</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>(0.36)</td>
<td>(0.84)</td>
<td>(0.94)</td>
<td>(0.10)</td>
<td>(-0.40)</td>
<td>(1.01)</td>
</tr>
<tr>
<td>F-stat.</td>
<td>6.70</td>
<td>3.87</td>
<td>1.58</td>
<td>2.20</td>
<td>1.02</td>
<td>2.40</td>
</tr>
</tbody>
</table>

The regressions are of the form:

$$\ln[Y(i)] = b_0 + b_1 \ln[\text{Value}(i)] + b_2 \ln[\text{Volume}(i)] + b_3 \sigma_{i2}^2 + b_4 \text{LOB1}(i) + e(i),$$

where $Y(i)$ stands for the ratio of the estimated variance for the window during the non-announcement period to be estimated variance, $\sigma_{i2}^2$, for the estimation period and the number of observations for each regression is 4615 and the t-statistics are in parentheses.