ELSEVIER

Journal of Banking & Finance 19 (1995) 1025–1046

Journal of BANKING & FINANCE

Production of information, information asymmetry, and the bid-ask spread: Empirical evidence from analysts' forecasts

Kee H. Chung [a,*], Thomas H. McInish [a], Robert A. Wood [a], Donald J. Wyhowski [b]

[a] *Department of Finance, Insurance, and Real Estate, The University of Memphis, Memphis, TN 38152, USA*

[b] *Department of Economics, SUNY at Albany, Albany, NY 12222, USA*

Received November 1993; final version received December 1993

Abstract

In this paper we suggest that market makers deduce the extent of the adverse selection problem associated with a stock (and set up the bid-ask spread accordingly) by observing how many financial analysts are following that stock. Market makers do this based on the belief that more financial analysts would follow a stock with a greater extent of information asymmetry since the value of private information increases with informational asymmetry. Similarly, financial analysts deduce the profit potential of a stock from the size of the spread set up by market makers (based on the expectation that market makers would set up a greater spread for the stock with a greater information asymmetry). This structural view of the process determining the bid-ask spread and analyst following is empirically tested using a simultaneous equations regression analysis. The empirical results are generally consistent with this view. Hence, our study supports the notion that the decisions of two major players in the financial markets (i.e., market makers and financial analysts) are made interactively.

Keywords: Information asymmetry; Bid-ask spread; Analysts' forecasts

JEL classification: G1

* Corresponding author.

0378-4266/95/$09.50
SSDI 0378-4266(94)00068-9

1. Introduction

In recent years, there have been significant advances in two types of equilibrium theories, one pertaining to the effect of information asymmetry on the behavior of market makers (the bid-ask spread approach), and the other pertaining to the effect of information asymmetry on the behavior of financial analysts (the analyst following approach). Although these theories have been developed separately, we suggest in this paper that the behavior of market makers and financial analysts may be examined from a broader perspective in which they (both market makers and financial analysts) extract the extent of information asymmetry about a firm's prospect from the behavior of the other party. Indeed our empirical results show that the effects between the spread and analyst following go in both directions; that is, the interaction between them is simultaneous.

Theories explaining the effect of information asymmetry on the behavior of market makers have been put forward by Bagehot (1971), Copeland and Galai (1983), Glosten and Milgrom (1985), and Easley and O'Hara (1987), among others. These studies have suggested that market makers recoup losses suffered in trades with informed traders through gains in trades with liquidity traders by increasing the size of the spread as the extent of information asymmetry between market makers (dealers/specialists) and informed traders becomes greater. In an attempt to examine the empirical validity of the theory, previous studies have employed various measures of information asymmetry. Benston and Hagerman (1974) use a security's unsystematic risk as an empirical proxy for the degree of market makers' exposure to informed traders. They hypothesize that more frequent occurrence of firm specific events leads to greater unsystematic risk and, consequently, greater opportunity for informed traders to trade against market makers. Based upon this view, Benston and Hagerman (1974) predict a positive correlation between spreads and unsystematic risk. Stoll (1978) suggests that market makers' losses to informed traders will be greater for stocks with greater trading volume. Chiang and Venkatesh (1988) employ insider ownerships and institutional holdings as proxies for the degree of information asymmetry faced by market makers.

Although previous studies provide useful insights into the effect of information asymmetry on spreads, the measures of the extent of information asymmetry among market participants used in these studies are indirect and thus not completely satisfactory. In addition, these studies have not explored the endogenous nature of information production, that is, the firm's information environment (e.g., the supply of and demand for information) has been an exogenous variable in these studies. The main focus of our study is to reexamine the information motive for bid-ask spreads when the production of information is endogenous.

Recently there has been considerable interest in the theory of information production. Numerous studies have examined the factors determining the number of analysts who follow a firm. Bhushan (Bhushan, 1989a; Bhushan, 1989b), Moyer et al. (1989), and O'Brien and Bhushan (1990) have suggested that financial analysts' information acquisition is closely related to the information environment and agency costs. Viewing the number of analysts following a firm as a proxy for the equilibrium total expenditure on analyst services for that firm, Bhushan and Moyer et al. show how the demand for and supply of analyst services are influenced by various firm characteristics (e.g., size, ownership structure, growth potential, and risk), representing the firm's information environment and potential agency costs. [1]

Pearson (1991) develops a multi-asset noisy rational expectations model describing the determinants of the cross-sectional distribution of the production of information about security returns. Following Bhushan (1989b), Pearson uses the number of analysts following a firm as a proxy for the level of information production and performs empirical testing on the cross-sectional distribution of information production. Brennan and Hughes (1991) develop a model relating information production and stock splits and conclude that a firm's analyst following is negatively related to the level of its stock price after controlling for firm size. These studies suggest that one of the main forces driving the cross-sectional distribution of analyst attention is the extent of information asymmetry about firms' future prospects since the value of private information increases with the degree of information asymmetry.

Based on these considerations, we conjecture that market makers deduce the extent of the adverse selection problem associated with a stock (and set up the bid-ask spread accordingly) by observing how many financial analysts are following that stock. Market makers do this based on the belief that more stock analysts would follow the stock with a greater extent of information asymmetry since the value of private information increases with informational asymmetry. Similarly, financial analysts deduce the profit potential of a stock from the size of the market makers' spread (based on the expectation that spreads would be greater for a stock with greater information asymmetry). This structural view of the process determining the bid-ask spread and analyst following is empirically tested using a simultaneous equations regression analysis. The empirical results are generally consistent with this view. Therefore, our study supports the notion that the decisions of two major players in the financial markets (i.e., market makers and financial analysts) are made interactively.

The paper is organized as follows. Section 2 states the hypotheses and describes the data. Section 3 provides the empirical results. The final section provides a summary.

[1] Although Bhushan (Bhushan, 1989a; Bhushan, 1989b) and Moyer et al. (1989) arrive at similar predictions regarding the relationship between analyst following and firm characteristics, they use somewhat different arguments. Moyer et al. stress the role of analysts as an efficient monitoring device for controlling agency-related costs of debt and equity, whereas Bhushan focuses on the factors influencing the demand for and supply of information driven by profit seeking agents.

2. Statement of hypotheses and data description

2.1. Production of information, information asymmetry, and the bid-ask spread

Financial analysts from major brokerage houses, independent research services, and banks produce and analyze information on a large number of firms. The demand for financial analysts' information collection (and, ultimately, the equilibrium supply of such activity in a competitive market for information) is expected to be positively related to the potential economic value of the information being generated. Once an analyst identifies a useful piece of information, he is most likely to first use it for his company or transmit it to his favored clients, then put it in a newsletter to subscribers, and finally disclose it publicly.

If there were little economic benefit to his company and/or his clients expected from the information on a firm, the analyst would have very little incentive to follow that firm. To the extent that the value of private information increases with the degree of information asymmetry about the firm's future prospects, we expect that a larger number of financial analysts would follow a firm with a greater degree of information asymmetry. Considering that more stock analysts will be enticed to follow a stock with a greater degree of information asymmetry (thus greater profit potential), we hypothesize that market makers deduce the extent of adverse selection risk associated with a stock by observing how many financial analysts are following that stock. Since market makers set up greater spreads for stocks with greater adverse selection risk, we expect that the spread increases with a larger number of analysts following the firm. Hence we have:

Hypothesis 1a: the number of analysts following a firm positively affects the bid-ask spread.

Contrarily, it may be argued that the degree of information asymmetry about an asset's intrinsic value will be less if more information is available about the asset from independent sources. At one extreme, if no independent financial analyst follows the firm, the information about the firm available to investors will be mostly internally-generated and self-reported (e.g., annual reports). For this case, insiders may possess valuable private information and thus be able to exploit market makers who do not possess that information. On the other hand, if the firm is followed by many financial analysts and thereby subject to considerable scrutiny, the likelihood of such exploitation will be less because much of the information regarding the firm will be available to investors and market makers through analyst services. These considerations suggest that the adverse selection risk faced by market makers may be negatively correlated with the number of analysts following the firm which, in turn, implies that the bid-ask spread may be lower for firms followed by a larger number of financial analysts. Hence we now have a competing hypothesis to Hypothesis 1a:

Hypothesis 1b: the number of analysts following a firm negatively affects the bid-ask spread.

Hence, depending on whether market makers perceive larger analyst following as an indication of a greater extent of informational asymmetry or, conversely, the extent of adverse selection risk faced by market makers actually decreases through analyst services, we would expect either a positive or negative effect of analyst following on spreads. Since theories predict an ambiguous effect, the ultimate verdict on which theory is a more reasonable representation of the real world can only be made by empirical confirmation.

In order to envision the effect of spreads on the production of information, consider that financial analysts deduce the extent of information asymmetry (and profit potential) through the size of spreads established by market makers. Analysts do this based on the belief that market makers would establish greater spreads for stocks with greater information asymmetry. Then we would expect that the number of analysts following a firm increases with the bid-ask spread. In this case, the difference in spreads causes the difference in the number of analysts following firms, rather than vice versa. Hence we have:

Hypothesis 2: the bid-ask spread positively affects the number of analysts following a firm.

Our next hypothesis concerns the impact of differential opinion on spreads. It is postulated that the degree of information asymmetry among market participants is higher, the greater the disagreement among financial analysts regarding a firm's future prospects, *ceteris paribus*. If there is less discrepancy among assessments made by different analysts regarding a firm's future prospects, the operating environment of the firm is probably more predictable. For this case, the likelihood of possessing *valuable* inside information will be minimal. On the other hand, if different analysts make vastly different forecasts regarding a firm's future performance, there may be greater uncertainty in the firm's operating environment and thus a greater profit potential from insider information.

Information asymmetry, as typically defined in theory, refers to differences in estimates of firm value (or share price). Disagreements among analysts, however, are mostly about current and/or next period's earnings. To the extent that firm value is strongly correlated with expected future earnings, [2] however, the dispersion of analysts' opinion is expected to capture the degree of information asymmetry regarding firm value. Based upon these considerations, this study postulates that the adverse selection risk faced by market makers is positively correlated with the divergence of opinion among financial analysts, which in turn leads to our next hypothesis:

[2] See Brown and Rozeff (1978), Elton et al. (Elton et al., 1981; Elton et al., 1984), Cragg and Malkiel (1982), Fried and Givoly (1982), and Brown et al. (1987) for empirical evidence.

Hypothesis 3: the bid-ask spread is higher for firms with a greater divergence of analysts' opinion regarding their future prospects.

In addition to the main hypotheses described above, we are interested in a number of subsidiary issues concerning the determinants of spreads and analyst following based upon other (than informational) considerations. In the remainder of this section, we discuss these variables.

Price volatility

Studies have suggested that the volatility of share price positively affects bid-ask spreads [see Tinic and West (1972), Benston and Hagerman (1974), Stoll (1978), Copeland and Galai (1983), and McInish and Wood (1992)].

Bhushan (1989a,b) and Moyer et al. (1989) suggest that private information is more valuable for firms with higher price variability. As the variability of share price increases, so does the probability of getting large deviations between the expected price based upon both private and public information and the expected price conditional just on public information. The expected trading profit is an increasing function of this probability and it follows that the expected trading profit based on private information will be higher for a firm with higher price variability. Thus the aggregate demand for analyst services is likely to be an increasing function of a firm's price volatility. Assuming that the cost of information acquisition is not significantly higher for firms with higher price volatility, we would expect that analyst following is positively associated with a firm's price variability.

Trading volume

Previous studies have documented a strong inverse relationship between trading volume and the bid-ask spread [see McInish and Wood (1992) for a review]. High volume typically reflects a lack of consensus, which in turn increases the demand for informedness. Therefore we expect an increase in volume to cause an increase in the number of analysts who follow a stock.

Number of shareholders

To the extent that analysts' forecasts are not a public good (at least before they are disclosed to the public), we would expect the number of analysts who follow the stock to be positively correlated with the number of shareholders.

Share price

Stoll (1978) suggests that the minimum allowable spread of \$1/8 can cause low priced stocks to have artificially high spreads. On the other hand, Brennan and Hughes (1991) develop a model relating information production and stock splits and conclude that a firm's analyst following is negatively related to the level of its stock price after controlling for firm size.

Multiple listing

Brennan and Hughes (1991) argue that firms use stock splits to attract analysts' attention. We suggest that the exchange listing also serves the similar function, and that after controlling for stock price, more analysts will follow multiply listed stocks.

Firm size

Larger firms, on average, release more information than smaller firms. They are also more intensively followed by security analysts and thus are subject to closer scrutiny by the investment community than smaller firms. This suggests that the stock prices of larger firms are relatively more informative and, as a result, the extent of information asymmetry is likely to be lower for larger firms. Hence, we would expect an inverse relation between spreads and firm size.

On the other hand, Bhushan (1989a,b), Moyer et al. (1989), and Brennan and Hughes (1991) have argued that the aggregate demand for analyst services is likely to be an increasing function of firm size. An investor is likely to find private information about a larger firm more valuable than the same information about a smaller firm. This is because the profits that the investor can generate by trading in a larger firm on the basis of this information are likely to be higher than those in a smaller firm. Firm size is also likely to influence the aggregate supply of analyst services. Analysts have incentives to focus on larger firms because they are more widely held and are of interest to more investors. An increase in transactions lowers the cost of providing analyst services and thus increases aggregate supply.

2.2. Description of data and empirical model specification

Data required for the calculation of bid-ask spreads are obtained from tapes provided by the Institute for the Study of Security Markets (ISSM). These data comprise every quotation (market bid and ask) for the New York Stock Exchange (NYSE) and the American Stock Exchange (AMEX) entered into the Consolidated Quotation Systems by NYSE/AMEX specialists for the 1984-1988 period. Each quotation is time stamped to the nearest second at the time it is provided to quotation vendors by the Securities Industry Automation Corporation. For each stock, the percentage bid-ask spread, (Ask-Bid)/{(Ask + Bid)/2}, is calculated for every quotation. The daily time-weighted percentage spread (TWPS) is then estimated by calculating the weighted average of percentage spreads of all quotations for the day, where the weight for each quotation is the number of seconds the quotation is outstanding divided by the number of seconds for the trading day, i.e.,

$$\mathrm{TWPS} = \Sigma(D_k/T)\left[(\mathrm{Ask}_k - \mathrm{Bid}_k)/\{(\mathrm{Ask}_k + \mathrm{Bid}_k)/2\}\right],$$

where D_k is the duration of quotation k in seconds, T is the number of seconds for the trading day, Ask_k and Bid_k are ask and bid prices, respectively, of quotation k, and Σ denotes the summation over $k = 1$ to N, where N is the total number of quotations for the day. Then the bid-ask spread of stock i in year t ($\text{SPREAD}_{i,t}$, $t = 1984, 1985, \ldots, 1988$) is measured by the mean value of TWPSs for each stock.

The number of analysts reporting a one-year ahead earnings forecast for firm i in July of year t ($\text{NAF}_{i,t}$) is obtained from the *I/B/E/S* Inc. database. *I/B/E/S* Inc. compiles earnings expectations from 400 leading brokerage firms on 10,000 publicly traded companies. Divergence of analyst opinion for firm i in year t ($\text{DISPERSION}_{i,t}$) is measured by the coefficient of variation of earnings forecasts made by different analysts in July of year t. Hence, firms followed by fewer than two analysts are excluded. [3]

Stock price variability of firm i in year t ($\text{VOLATILITY}_{i,t}$) is measured by the coefficient of variation of the mid-point of all quoted bid-ask spreads on the ISSM tape during each year. Following McInish and Wood (McInish and Wood, 1991; McInish and Wood, 1992), trading volume is measured by both the average number of transactions ($\text{NOFT}_{i,t}$) and the average transaction size ($\text{SOFT}_{i,t}$) obtained from the ISSM tape. Average share price of stock i in year t ($\text{PRICE}_{i,t}$) is estimated by the weighted average of the mid-points of bid-ask spread quotations for the year, where the weight for each quotation is the number of seconds the quotation is outstanding divided by the number of seconds for each year. The number of shareholders of firm i in year t ($\text{NSH}_{i,t}$) is obtained from the annual industrial COMPUSTAT tapes. $\text{MULTIPLE}_{i,t}$ is defined as zero if firm (stock) i is traded only on the primary exchange and 1.0 if it is also traded on other exchanges during year t. Firm size is measured by the market value of equity ($\text{MVE}_{i,t}$) which is the product of the average daily stock price and the number of shares outstanding at the end of year t.

Firms are included in the final sample only if complete data are available from three data sources (ISSM tapes, *I/B/E/S* Database, and COMPUSTAT tape) for the entire study period. Because the majority of firms have fiscal years ending in December, only such firms are included in the final sample. The final dataset comprises 2,345 pooled time-series and cross-sectional observations (469 firms for each year in the five-year study period). Table 1 presents descriptive statistics of the variables. For each variable, its mean, standard deviation, median, and selected percentile values during the study period are provided. The table shows that the mean and standard deviation of spreads (in percent) are 1.275 and 1.092, respectively. On average, our sample of firms are followed by 17.6 analysts.

We employ the following structural model as an empirical representation of the

[3] Regression results with a sample of firms followed by at least five analysts are qualitatively similar to those presented here.

Table 1
Summary statistics

	Mean	Standard deviation	Median	Percentile			
				5	25	75	95
Bid-ask spread (%) ($SPREAD_{i,t}$) [a]	1.275	1.092	1.057	0.524	0.816	1.373	2.569
Number of analysts following the firm ($NAF_{i,t}$)	17.6	8.5	17	6	11	23	34
Dispersion of analysts' forecasts ($DISPERSION_{i,t}$)	0.153	0.354	0.058	0.016	0.033	0.120	0.525
Stock price volatility ($VOLATILITY_{i,t}$)	0.007	0.004	0.006	0.004	0.005	0.008	0.011
Size of transactions ($SOFT_{i,t}$)	55.0	37.0	48.7	10.2	27.4	74.8	121.1
Number of transactions ($NOFT_{i,t}$)	94.4	112.6	58	10	28	117	306
Number of shareholders ($NSH_{i,t}$)	69.3	212.0	21.2	2.4	7.8	51.3	231.7
Share price ($) ($PRICE_{i,t}$)	36.4	26.4	32.0	9.2	22.6	44.4	73.8
Market value of equity ($MVE_{i,t}$)	2,523	4,572	1,103	111	459	2,684	9,807

[a] $SPREAD_{i,t}$ = the average bid-ask spread of firm i in year t.
$NAF_{i,t}$ = the number of financial analysts following firm i in year t.
$DISPERSION_{i,t}$ = the coefficient of variation of earnings forecasts for firm i in year t.
$VOLATILITY_{i,t}$ = the coefficient of variation of mid-point of all quoted spreads for stock i in year t.
$SOFT_{i,t}$ = the average transaction size of stock i in year t (in round lots).
$NOFT_{i,t}$ = the average number of transactions of stock i in year t.
$NSH_{i,t}$ = the number of shareholders of firm i in year t (in thousand).
$PRICE_{i,t}$ = the average price of stock i in year t.
$MVE_{i,t}$ = the market value of equity of firm i in year t (in millions of dollars).

process determining the bid-ask spread and the number of analysts following a firm:

$$\begin{aligned}\text{SPREAD}_{i,t} = {} & \alpha_0 + \alpha_1 \text{NAF}_{i,t} + \alpha_2 \text{DISPERSION}_{i,t} + \alpha_3 \text{VOLATILITY}_{i,t} \\ & + \alpha_4 \text{SOFT}_{i,t} + \alpha_5 \text{NOFT}_{i,t} + \alpha_6 \text{PRICE}_{i,t} + \alpha_7 \text{MVE}_{i,t} + \epsilon_{i,t}\end{aligned} \quad (1)$$

and

$$\begin{aligned}\text{NAF}_{i,t} = {} & \beta_0 + \beta_1 \text{SPREAD}_{i,t} + \beta_2 \text{SOFT}_{i,t} + \beta_3 \text{NOFT}_{i,t} + \beta_4 \text{NSH}_{i,t} \\ & + \beta_5 \text{PRICE}_{i,t} + \beta_6 \text{MULTIPLE}_{i,t} + \beta_7 \text{MVE}_{i,t} + \epsilon_{i,t}\end{aligned} \quad (2)$$

($i = 1, 2, \ldots, 469$; $t = 1984, 1985, \ldots, 1988$); where $\text{SPREAD}_{i,t}$ = the average bid-ask spread of stock i in year t, $\text{NAF}_{i,t}$ = the number of financial analysts following firm i in year t, $\text{DISPERSION}_{i,t}$ = the coefficient of variation of earnings forecasts made by different analysts for firm i in year t, $\text{VOLATILITY}_{i,t}$ = the coefficient of variation of the mid-point of all quoted bid-ask spreads for stock i in year t, $\text{NOFT}_{i,t}$ = the average number of transactions of stock i in year t, $\text{SOFT}_{i,t}$ = the average transaction size of stock i in year t, $\text{PRICE}_{i,t}$ = the average price of stock i in year t, $\text{NSH}_{i,t}$ = the number of shareholders of firm i in year t, $\text{MULTIPLE}_{i,t}$ = the multiple listing dummy variable (zero if a stock is traded only on the primary exchange and 1.0 if it is also traded on other exchanges), $\text{MVE}_{i,t}$ = the market value of equity of firm i in year t, and $\epsilon_{i,t}$ = the disturbance term.

Notice that our empirical model explicitly recognizes the endogeneity of the number of analysts following a firm as a determinant of the bid-ask spread. Similarly, the bid-ask spread is included in the analyst following equation. The expected sign of α_1 is ambiguous since it depends on whether hypothesis 1a or hypothesis 1b is empirically more relevant. This study will give an empirical answer to a question for which the theory offers only an ambiguous prediction. According to hypothesis 2, the expected sign of β_1 is positive. The expected sign of α_2 is positive according to hypothesis 3. The expected signs of other coefficients are $\alpha_3 > 0$, $\alpha_4 < 0$, $\alpha_5 < 0$, $\alpha_6 < 0$, and $\alpha_7 < 0$ in the bid-ask spread equation and $\beta_2 > 0$, $\beta_3 > 0$, $\beta_4 > 0$, $\beta_5 < 0$, $\beta_6 > 0$, and $\beta_7 > 0$ in the analyst following equation.

3. Empirical results

3.1. A brief look at the correlation structure

Since some of our independent variables are expected to be highly correlated (and thereby potentially cause a multicollinearity problem), we first examine the correlation structure among the variables. [4] Table 2 shows the correlation matrix for our variables. Most variables are highly correlated with firm size (e.g., the correlation between NAF and MVE is 0.634 and both measures of trading volume, i.e., SOFT and NOFT, also show strong correlations with firm size). [5] Share price (PRICE) is highly correlated with both the average size of transactions (SOFT) and the spread (SPREAD). [6]

Note that correlation estimates presented in Table 2 may suggest misleading notions on the relationship between certain variables. The inverse correlation between SPREAD and NAF may be spurious since both variables are strongly correlated with size. That is, since NAF is highly positively correlated with firm size and larger firms generally have lower bid-ask spreads than smaller firms, the negative correlation between SPREAD and NAF may be due merely to the size of the firm. To examine this more we calculate the correlation matrix (see panel B of Table 2) after variables are made orthogonal to firm size. [7] The results in panel B are consistent with the view that the inverse correlation between SPREAD and NAF is spurious. After controlling for the effect of firm size on both variables, SPREAD and NAF are actually positively associated. Similarly, since NAF is highly positively correlated with firm size and larger firms generally have lower price volatility than smaller firms, the negative correlation between NAF and VOLATILITY (in panel A) may be also due to firm size. After controlling for the effect of firm size on both variables, the results reveal that NAF and VOLATILITY are actually positively related, which is again consistent with our expectations.

3.2. *A diagnostic investigation of multicollinearity*

Since the correlation matrix suggests that selected variables (e.g., PRICE and SOFT, NAF and MVE) are highly correlated with each other, we follow the

[4] In our preliminary data analysis, we perform a Box and Cox (1964) analysis in order to identify the best (in an empirical sense) functional form from data rather than imposing one on an *ad hoc* basis. Since linear and log-linear models are frequently employed in the empirical literature, we also perform the likelihood ratio test to determine whether the 'goodness of fit' of either model is significantly different from that of the Box-Cox model. Results of the likelihood ratio test suggest that the goodness of fit of the log-linear model is not different from that of the Box-Cox model at the one percent significance level in most years. Given this result, and considering that the log-linear model is simple and gives easily interpretable coefficient estimates (i.e., elasticities), we use log-transformed variables in the empirical investigation.

[5] For notational simplicity, we hereafter omit the firm (i) and time (t) subscript from the variable name unless it is essential.

[6] This high correlation between PRICE and SOFT is largely attributable to the fact that SOFT is defined as a function of PRICE (i.e., transaction size = share price × the number of shares traded).

[7] That is, we regress each variable against firm size and calculate the residuals of the regression model. Panel B presents the correlation matrix based on these residuals.

Table 2
Pearson correlation coefficients

Panel A: Correlation coefficients using log of variables

	SPREAD	NAF	DISP.	VOLAT.	SOFT	NOFT	NSH	PRICE	MULTI.	MVE
SPREAD [a]	1.000	−0.378	0.369	0.730	−0.809	−0.398	−0.304	−0.907	−0.186	−0.664
NAF		1.000	−0.110	−0.185	0.478	0.584	0.571	0.317	0.430	0.634
DISPERSION			1.000	0.438	−0.253	0.030	−0.105	−0.405	0.084	−0.253
VOLATILITY				1.000	−0.516	−0.019	−0.259	−0.668	0.029	−0.404
SOFT					1.000	0.412	0.204	0.827	0.294	0.693
NOFT						1.000	0.626	0.227	0.642	0.684
NSH							1.000	0.136	0.420	0.572
PRICE								1.000	0.122	0.603
MULTIPLE									1.000	0.462
MVE										1.000

Panel B: Correlation coefficients after variables are made orthogonal to firm size [b]

	SPREAD	NAF	DISP.	VOLAT.	SOFT	NOFT	NSH	PRICE	MULTI.	MVE
SPREAD	1.000	0.074	0.278	0.675	−0.648	0.103	0.122	−0.849	0.161	0.000
NAF		1.000	0.067	0.100	0.069	0.266	0.328	−0.107	0.178	0.000
DISPERSION			1.000	0.380	−0.111	0.289	0.050	−0.327	0.208	0.000
VOLATILITY				1.000	−0.359	0.385	−0.038	−0.582	0.235	0.000
SOFT					1.000	−0.118	−0.324	0.711	−0.035	0.000
NOFT						1.000	0.392	−0.320	0.447	0.000
NSH							1.000	−0.320	0.190	0.000
PRICE								1.000	−0.197	0.000
MULTIPLE									1.000	0.000
MVE										1.000

[a] See footnote a in Table 1 for the definition of the variables.

[b] We first regress each variable against firm size and calculate the residuals of the regression model. Panel B presents the correlation matrix based on these residuals.

diagnostic procedure of Belsley et al. (1980) and examine the extent of multicollinearity among the variables. First, we search for the presence of linear dependencies and attempt to isolate which explanatory variables are interrelated. We then assess any adverse effects of linear dependencies on the precision of our estimates. Panels A and B in Table 3 present the eigenvalue of **X'X** matrix, the condition index, and the proportion of the variance of the estimate accounted for by each principal component when regression models (1) and (2) are applied to our data. The condition indices are the square roots of the ratio of the largest eigenvalue to each individual eigenvalue. The largest condition index is the condition number of the scaled **X** matrix. As a rule of thumb, Belsley et al. (1980) suggest that when the condition number is greater than about 30, then a linear dependency among the columns of **X** exists that may seriously affect the standard errors of the estimated coefficients. Belsley et al. also suggest that the collinearity problem exists when a component associated with a high condition index contributes strongly (as a rule of thumb, 0.50 or greater) to the variance of two or more variables.

Table 3
Collinearity diagnostics

Panel A: SPREAD equation [a]			Variance proportion							
No.	Eigen-value	Condition index [b]	Intercept	NAF	DISP.	VOLAT.	SOFT	NOFT	PRICE	MVE
1	7.74431	1.0000	0.0000	0.0003	0.0016	0.0000	0.0002	0.0004	0.0001	0.0003
2	0.12647	7.8255	0.0000	0.0063	0.6213	0.0000	0.0005	0.0330	0.0002	0.0061
3	0.05508	11.8578	0.0114	0.0001	0.1444	0.0067	0.0005	0.0643	0.0041	0.1107
4	0.03723	14.4231	0.0047	0.0419	0.1113	0.0006	0.0703	0.1486	0.0523	0.0426
5	0.01824	20.6056	0.0021	0.7768	0.0037	0.0010	0.0003	0.3867	0.0075	0.0084
6	0.01196	25.4512	0.0153	0.0946	0.0491	0.0130	0.1733	0.2365	0.0133	0.7565
7	0.00516	38.7575	0.0163	0.0780	0.0436	0.0004	0.7508	0.0564	0.8015	0.0287
8	0.00105	85.7250	0.9502	0.0020	0.0251	0.9783	0.0041	0.0720	0.1210	0.0467

Panel B: NAF equation			Variance proportion							
No.	Eigen-value	Condition index	Intercept	SPREAD	SOFT	NOFT	NSH	PRICE	MULTI.	MVE
1	7.57583	1.0000	0.0001	0.0000	0.0002	0.0003	0.0013	0.0001	0.0022	0.0003
2	0.22826	5.7610	0.0019	0.0004	0.0021	0.0008	0.0225	0.0016	0.4312	0.0061
3	0.11667	8.0580	0.0006	0.0001	0.0029	0.0008	0.5176	0.0009	0.2506	0.1107
4	0.03690	14.3288	0.0685	0.0017	0.0195	0.0050	0.0238	0.0026	0.0170	0.0426
5	0.02465	17.5306	0.0004	0.006	0.0109	0.5518	0.2977	0.0193	0.2314	0.0084
6	0.01076	26.5392	0.1214	0.0001	0.2011	0.2595	0.0601	0.0161	0.0280	0.7565
7	0.00588	35.9075	0.0758	0.0107	0.7610	0.0875	0.0217	0.2775	0.0089	0.0287
8	0.00105	84.9419	0.7312	0.9863	0.0023	0.0943	0.0553	0.6819	0.0307	0.0467

Panel C: SPREAD equation			Variance proportion							
No.	Eigen-value	Condition index	Intercept	NAF	DISP.	VOLAT.	SOFT	NOFT	LOWP	MVE
6	0.01734	19.8007	0.0017	0.9814	0.0003	0.0019	0.0576	0.0917	0.0303	0.0355
7	0.01113	24.7179	0.0125	0.0101	0.0625	0.0145	0.7212	0.1432	0.0519	0.5541
8	0.00101	81.9419	0.9711	0.0026	0.0342	0.9737	0.0111	0.0992	0.1359	0.0720

Panel D: NAF equation			Variance proportion							
No.	Eigen-value	Condition Index	Intercept	SPREAD	SOFT	NOFT	NSH	LOWP	MULTI.	MVE
6	0.01948	18.4738	0.0018	0.0005	0.0583	0.8831	0.1104	0.0916	0.1315	0.0900
7	0.00940	26.5923	0.1187	0.0019	0.6263	0.0657	0.0997	0.0891	0.0238	0.5092
8	0.00210	56.2274	0.8219	0.9883	0.3006	0.0005	0.0247	0.1405	0.0328	0.0145

[a] See footnote a in Table 1 for the definition of the variables.

[b] The condition indices are the square roots of the ratio of the largest eigenvalue to each individual eigenvalue. The largest condition index is the condition number of the scaled **X** matrix. As a rule of thumb, Belsley, Kuh, and Welsch (1980) suggest that when the condition number is greater than about 30, then a linear dependency among the columns of **X** exists that may seriously affect the standard errors of the estimated coefficients. Belsley et al. also suggest that the collinearity problem exists when a component associated with a high condition index contributes strongly (as a rule of thumb, 0.50 or greater) to the variance of two or more variables.

Inspection of panels A and B of Table 3 reveals that there are strong linear dependencies between SOFT and PRICE and between the intercept and VOLATILITY in the spread equation. In the analyst following equation, there is a linear dependency among the intercept, SPREAD, and PRICE. It is not surprising to observe a strong linear dependency between PRICE and SPREAD and between PRICE and SOFT because the measurement of both SPREAD and SOFT involves the share price. Because these results suggest that the common source of multicollinearity is the share price and since the primary purpose of including the share price in regression models is to examine whether the bid-ask spread and analyst following are greater for low-priced stocks, we replicate the preceding analysis after we replace PRICE with a dummy variable (LOWP) representing low priced stocks. [8]

Panels C and D in Table 3 present the last three rows of the variance decomposition matrix when LOWP (instead of PRICE) is used in the regression. The results show that the multicollinearity problem associated with share price does not exist when the dummy variable is used. Linear dependencies between the intercept and VOLATILITY in the spread equation and between the intercept and SPREAD in the analyst following equation, however, remain intact. These linear dependencies adversely affect the precision of α_0 and α_3 estimates in the spread equation and the precision of β_0 and β_1 estimates in the analyst following equation. Since the inspection of Table 1 reveals that the variance of VOLATILITY in the spread equation and the variance of SPREAD in the analyst following equation are substantially less than that of other variables in each equation, the imprecision of estimated coefficients may be largely attributable to the relative lack of variation in these variables. Lastly, although not enough to represent a strong dependency, the condition numbers of 24.7 and 26.6 nevertheless suggest that a moderate linear dependency exists between SOFT and MVE.

Although the above analysis reveals that there exists a linear dependency among selected variables, the main focus of this study (i.e., the association of NAF and DISPERSION with spreads) is free from such dependency. Furthermore, the regression results reported in Table 4 show that all the variables we suggest as the determinants of SPREAD and NAF display statistical significance, despite any inflated variances due to linear dependency. Since an application of the ordinary least squares (OLS) method to a structural model is subject to simultaneous equation bias and thus may yield biased and inconsistent parameter estimates, we estimate our structural equations (1) and (2) using the three-stage least squares

[8] The dummy variable is defined as 1.0 if the share price is less than $20.00 and zero otherwise. The use of $20.00 as a breaking point is somewhat arbitrary, but this ensures a reasonable number of sample observations in each category. The results are qualitatively similar to those presented here when the regression analysis is replicated using other dollar values.

(3SLS) method. In order to observe the extent of simultaneous equation bias when the two equations are estimated separately, we also report the results of the OLS regression.

Table 4

Determinants of spreads and the number of analysts following the firm. $SPREAD_{i,t} = \alpha_0 + \alpha_1 NAF_{i,t} + \alpha_2 DISPERSION_{i,t} + \alpha_3 VOLATILITY_{i,t} + \alpha_4 SOFT_{i,t} + \alpha_5 NOFT_{i,t} + \alpha_6 LOWP_{i,t} + \alpha_7 MVE_{i,t} + \epsilon_{i,t}$, $NAF_{i,t} = \beta_0 + \beta_1 SPREAD_{i,t} + \beta_2 SOFT_{i,t} + \beta_3 NOFT_{i,t} + \beta_4 NSH_{i,t} + \beta_5 LOWP_{i,t} + \beta_6 MULTIPLE_{i,t} + \beta_7 MVE_{i,t} + \epsilon_{i,t}$

	OLS Estimates		3SLS Estimates	
	SPREAD [a,b]	NAF	SPREAD	NAF
Intercept	0.310	1.615	0.194	1.665
	(3.31 **)	(16.30 **)	(1.92)	(11.09 **)
Number of analysts	0.081		0.229	
Following the firm (NAF)	(7.66 **)		(6.88 **)	
Bid-ask spread (SPREAD)		0.237		0.241
		(7.76 **)		(4.79 **)
Dispersion of analysts'	0.022		0.019	
forecasts (DISPERSION)	(4.91 **)		(4.53 **)	
Stock price volatility	0.685		0.693	
(VOLATILITY)	(35.42 **)		(34.83 **)	
Size of transactions	−0.291	0.278	−0.308	0.275
(SOFT)	(−33.79 **)	(14.25 **)	(−31.73 **)	(11.25 **)
Number of transactions	−0.101	0.062	−0.121	0.047
(NOFT)	(−16.26 **)	(4.99 **)	(−15.44 **)	(3.79 **)
Number of shareholders		0.131		0.127
(NSH)		(16.78 **)		(16.01 **)
Low price dummy (LOWP)	0.164	0.182	0.122	0.169
	(6.80 **)	(4.17 **)	(4.53 **)	(3.53 **)
Multiple listing dummy		0.031		0.101
(MULTIPLE)		(1.31)		(4.57 **)
Market value of equity	−0.015	0.097	−0.037	0.099
(MVE)	(−2.42 *)	(8.85 **)	(−4.78 **)	(8.85 **)
F-value	1,593.72 **	370.29 **		
Adjusted R^2	0.826	0.524		
Root MSE	0.206	0.377		
System weighted R^2			0.682	
System weighted MSE			1.236	
Number of observations	2,345	2,345	2,345	

[a] See footnote a in Table 1 for the definition of the variables.

[b] The expected signs of regression coefficients (and their theoretical origins) are as follows: $\alpha_1 = ?$ (hypotheses 1a and 1b), $\alpha_2 > 0$ (hypothesis 3), $\alpha_3 > 0$ (Tinic and West, 1972; Stoll, 1978), α_4, $\alpha_5 < 0$ (McInish and Wood, 1992), $\alpha_6 > 0$ (Stoll, 1978), $\alpha_7 < 0$ (this study), $\beta_1 > 0$ (hypothesis 2), β_2, $\beta_3 > 0$ (this study), $\beta_4 > 0$ (this study), $\beta_5 > 0$ (Brennan and Hughes, 1991), $\beta_6 > 0$ (this study), and $\beta_7 > 0$ (Bhushan, 1989a; Bhushan, 1989b; Moyer et al., 1989; and Brennan and Hughes, 1991).

** Significant at the 1% level.

* Significant at the 5% level.

3.3. Interpretation of empirical results

The regression results show that hypothesized variables jointly account for more than 80 percent of the variation in bid-ask spreads and 50 percent of the variation in the number of analysts following the firm. As a system of equations, the variables jointly explain 68.2 percent of the variation in spreads and analyst following. (All statistical tests are at the one percent level unless otherwise indicated.) The comparison of estimated coefficients of the 3SLS regression with those of the OLS regression reveals that the OLS method, which ignores the simultaneity of the process determining the spread and analyst following, gives biased estimates. For example, the OLS estimate of the coefficient for NAF in the spread equation is 0.081 whereas the 3SLS regression gives an estimate of 0.229 for the same variable. Since the 3SLS estimates are presumably more accurate than OLS estimates, the interpretation of empirical results will be based upon the former.

The results show that the number of analysts following a firm has a significant positive influence on the bid-ask spread, supporting our hypothesis 1a. The estimated coefficient suggests that a ten percent increase in analyst following results in a 2.29 percent increase in the bid-ask spread. This result provides empirical support for the view that since the value of private information increases with the degree of information asymmetry about the firm's future prospects, market makers set up higher spreads for stocks followed by more stock analysts. Contrarily, the notion that the extent of adverse selection risk faced by market makers may decrease through analyst services is not supported by our empirical results.

Empirical results reveal a strong positive effect of the divergence of analyst opinion on spreads, supporting our hypothesis 3. This result is consistent with the view that the adverse selection problem faced by market makers is greater for firms with greater divergence of opinion among financial analysts, and thus market makers will tend to maintain greater bid-ask spreads for these firms. This result is also compatible with the empirical finding by Amihud and Mendelson (1989) and Amihud et al., 1990 that excess returns and spreads are positively correlated, given the qualification that securities for which a considerable divergence of analyst opinion exists will earn excess returns [see Barry and Brown (1985) for this point].

A strong positive correlation exists between spreads and risk (i.e., VOLATILITY), which is consistent with the results of earlier studies by Tinic and West (1972), Benston and Hagerman (1974), Stoll (1978), Copeland and Galai (1983), and McInish and Wood (1992). Regression results also show that greater trading volume (i.e., SOFT and NOFT) results in lower spreads. This result is consistent with previous empirical evidence [see Demsetz (1968), Tinic and West (1972), Benston and Hagerman (1974), and McInish and Wood (1992)] that the bid-ask spread is inversely correlated with trading activity. The coefficient of the dummy variable for low priced stocks exhibits a positive sign, showing the effect of the minimum allowable spread of \$1/8. Lastly, larger firms exhibit low spreads, suggesting that the extent of information asymmetry is less for larger firms.

The results show that the spread has a significant positive effect on analyst following which is supportive of our hypothesis 2. According to the estimated coefficient, a ten percent increase in spreads will result in a 2.41 percent increase in analyst following. This result supports the view that financial analysts deduce

Table 5
Determinants of spread and the number of analysts following the firm (Regression results with firm-size orthogonalized variables). $\text{SPREAD}_{i,t} = \alpha_0 + \alpha_1\text{NAF}_{i,t} + \alpha_2\text{DISPERSION}_{i,t} + \alpha_3\text{VOLATILITY}_{i,t} + \alpha_4\text{SOFT}_{i,t} + \alpha_5\text{NOFT}_{i,t} + \alpha_6\text{LOWP}_{i,t} + \epsilon_{i,t}$ $\text{NAF}_{i,t} = \beta_0 + \beta_1\text{SPREAD}_{i,t} + \beta_2\text{SOFT}_{i,t} + \beta_3\text{NOFT}_{i,t} + \beta_4\text{NSH}_{i,t} + \beta_5\text{LOWP}_{i,t} + \beta_6\text{MULTIPLE}_{i,t} + \epsilon_{i,t}$

	OLS estimates		3SLS estimates	
	SPREAD [a,b]	NAF	SPREAD	NAF
Intercept	−0.008	−0.040	−0.005	−0.088
	(−1.70)	(−2.28 *)	(−1.16)	(−5.17 **)
Number of analysts	0.083		0.243	
Following the firm (NAF)	(7.83 **)		(7.35 **)	
Bid-ask spread (SPREAD)		0.237		0.230
		(7.78 **)		(4.59 **)
Dispersion of analysts'	0.023		0.020	
Forecasts (DISPERSION)	(5.04 **)		(4.60 **)	
Stock price volatility	0.693		0.699	
(VOLATILITY)	(36.10 **)		(35.26 **)	
Size of transactions	−0.294	0.276	−0.312	0.272
(SOFT)	(−34.55 **)	(14.25 **)	(−32.53 **)	(11.18 **)
Number of transactions	−0.101	0.060	−0.123	0.045
(NOFT)	(−16.24 **)	(4.90 **)	(−15.60 **)	(3.74 **)
Number of shareholders		0.131		0.125
(NSH)		(16.76 **)		(15.86 **)
Low price dummy (LOWP)	0.133	0.166	0.095	0.187
	(6.10 **)	(4.12 **)	(3.97 **)	(4.22 **)
Multiple listing dummy		0.041		0.106
(MULTIPLE)		(2.03 *)		(5.46 **)
F-value	864.61 **	101.20 **		
Adjusted R^2	0.689	0.204		
Root MSE	0.206	0.377		
System weighted R^2			0.533	
System weighted MSE			1.261	
Number of observations	2,345	2,345	2,345	

[a] See footnote a in Table 1 for the definition of the variables.
[b] The expected signs of regression coefficients (and their theoretical origins) are as follows: $\alpha_1 = ?$ (hypotheses 1a and 1b), $\alpha_2 > 0$ (hypothesis 3), $\alpha_3 > 0$ (Tinic and West, 1972; Stoll, 1978), α_4, $\alpha_5 < 0$ (McInish and Wood, 1992), $\alpha_6 > 0$ (Stoll, 1978), $\beta_1 > 0$ (hypothesis 2), β_2, $\beta_3 > 0$ (this study), $\beta_4 > 0$ (this study), $\beta_5 > 0$ (Brennan and Hughes, 1991), and $\beta_6 > 0$ (this study).
** Significant at the 1% level.
* Significant at the 5% level.

the extent of information asymmetry (and profit potential) through the size of spreads established by market makers.

The results also show that greater trading volume triggers more analyst following. This result supports the idea that since high trading volume reflects a lack of consensus, the demand for informedness rises with trading volume. As predicted, firms with larger investor bases are followed by more analysts, supporting the notion that analysts' forecasts are not a public good. Consistent with the finding of Brennan and Hughes (1991), empirical results show that analyst following is negatively associated with share price and positively associated with firm size. The results also show that more analysts follow multiply listed stocks.

In order to examine whether the results in Table 4 are driven by spurious correlations among variables operating through their correlations with firm size, we also replicate Table 4 using firm-size neutralized variables (i.e., after they are made orthogonal to firm size). The results (reported in Table 5) with firm-size neutralized variables are qualitatively similar to those with original variables. Hence we conclude that the results in Table 4 are not driven by spurious correlations.

Since our data includes both time-series and cross-sectional observations, we also employ the Fuller and Battese (1974) error component model in order to allow for a more general error structure. The Fuller-Battese error component model assumes that the disturbance term is composed of three independent components–one associated with time, another associated with the cross-sectional units, and the third varying in both dimensions. Specifically, we assume that the disturbance term in regression models (1) and (2) is defined as:

$$\epsilon_{i,t} = u_i + v_t + w_{i,t} \quad (i = 1,2,\ldots,469;\ t = 1984,1985,\ldots,1988). \tag{3}$$

If the behavior of the disturbances over the cross-sectional units is different from the behavior of the disturbances of a given cross-sectional unit over time, the regression model with this error structure would give more accurate coefficient estimates. The results of the Fuller-Battese estimation are presented in Table 6. The results show that the variance component for cross-sections is much larger than the variance component for error in the analyst following equation, and the two components are similar in magnitude in the spread equation. In both equations the variance component for time-series is comparatively very small. Despite different assumptions regarding the error structure of the regression model, however, the results in Table 6 are qualitatively similar to those in Table 4. [9] In addition, the unreported results of the regression models (2) and (3) using cross-sectional data for each year of 1984-1988 are also quite similar to the results

[9] Although the Fuller-Battese model permits a more general error structure, it does not capture the structural feature of our model. Hence, the results of the error component model are directly comparable only to those of the OLS.

Table 6
Determinants of spreads and the number of analysts following the firm (Fuller-Battese model estimates). $SPREAD_{i,t} = \alpha_0 + \alpha_1 NAF_{i,t} + \alpha_2 DISPERSION_{i,t} + \alpha_3 VOLATILITY_{i,t} + \alpha_4 SOFT_{i,t} + \alpha_5 NOFT_{i,t} + \alpha_6 LOWP_{i,t} + \alpha_7 MVE_{i,t} + u_i + v_t + w_{i,t}$ $NAF_{i,t} = \beta_0 + \beta_1 SPREAD_{i,t} + \beta_2 SOFT_{i,t} + \beta_3 NOFT_{i,t} + \beta_4 NSH_{i,t} + \beta_5 LOWP_{i,t} + \beta_6 MULTIPLE_{i,t} + \beta_7 MVE_{i,t} + u_i + v_t + w_{i,t}$

	SPREAD [a,b]	NAF
Intercept	0.147	1.971
	(1.44)	(24.25 *)
Number of analysts	0.066	
Following the firm (NAF)	(5.12 **)	
Bid-ask spread (SPREAD)		0.086
		(4.59 **)
Dispersion of analysts'	0.023	
Forecasts (DISPERSION)	(4.66 **)	
Stock price volatility	0.660	
(VOLATILITY)	(32.02 **)	
Size of transactions	−0.264	0.078
(SOFT)	(−28.72 **)	(6.01 **)
Number of transactions	−0.101	0.043
(NOFT)	(−15.41 **)	(3.34 **)
Number of shareholders		0.176
(NSH)		(16.06 **)
Low price dummy (LOWP)	0.185	0.134
	(7.73 **)	(3.12 **)
Multiple listing dummy		0.065
(MULTIPLE)		(2.26 *)
Market value of equity	−0.018	0.034
(MVE)	(−2.93 **)	(4.64 **)
Variance component for cross-sections	0.017	0.115
Variance component for time-series	0.001	0.004
Variance component for error	0.025	0.026
Root MSE	0.158	0.168
Number of cross-sections	469	469
Number of time-series	5	5

[a] See footnote a in Table 1 for the definition of the variables.
[b] The expected signs of regression coefficients are as follows: $\alpha_1 = ?$, $\alpha_2 > 0$, $\alpha_3 > 0$, α_4, $\alpha_5 < 0$, $\alpha_6 > 0$, $\alpha_7 < 0$, $\beta_1 > 0$, β_2, $\beta_3 > 0$, $\beta_4 > 0$, $\beta_5 > 0$, $\beta_6 > 0$, and $\beta_7 > 0$.
** Significant at the 1% level.
* Significant at the 5% level

of pooled cross-sectional and time-series data. This leads us to conclude that our empirical results are reasonably robust.

The robustness of our results is further examined by re-estimating the structural model with an alternative definition of stock price volatility, the annual average of (Daily High-Daily Low)/{(Daily High + Daily Low)/2}. The results (see Table 7, panel A) are, again, qualitatively identical to those in Table 4. Similarly, panel B of Table 7 presents the results when VOLATILITY (instead of SPREAD) is employed as a proxy variable for the expected profit associated with private

Table 7
Determinants of spread and the number of analysts following the firm (Testing model robustness using different measures of variables)

	A. 3SLS estimates [a]		B. 3SLS estimates [b]	
	SPREAD [c,d]	NAF	SPREAD	NAF
Intercept	−0.960	1.242	0.181	1.867
	(−12.48 **)	(7.58 **)	(1.78)	(10.31 **)
Number of analysts		0.306	0.228	
Following the firm (NAF)	(7.87 **)		(6.84 **)	
Bid-ask spread (SPREAD)		0.129		
		(2.78 **)		
Dispersion of analysts'	0.003		0.021	
Forecasts (DISPERSION)	(6.84 **)		(4.79 **)	
Stock price volatility	0.532		0.689	0.192
(VOLATILITY) [c]	(29.93 **)		(34.34 **)	(5.12 **)
Size of transactions	−0.341	0.219	−0.308	0.215
(SOFT)	(−33.11 **)	(8.44 **)	(−31.75 **)	(12.70 **)
Number of transactions	−0.208	0.042	−0.121	0.018
(NOFT)	(−20.17 **)	(3.43 **)	(−15.45 **)	(1.37)
Number of shareholders		0.120		0.135
(NSH)		(15.06 **)		(16.15 **)
Low price dummy (LOWP)	0.149	0.241	0.122	0.208
	(5.02 **)	(4.90 **)	(4.54 **)	(4.70 **)
Multiple listing dummy		0.126		0.106
(MULTIPLE)		(5.57 **)		(4.76 **)
Market value of equity	−0.033	0.092	−0.037	0.094
(MVE)	(−3.89 *)	(8.17 **)	(−4.71 **)	(8.53 **)
System weighted R^2	0.648		0.702	
System weighted MSE	1.231		1.087	
Number of observations	2,345		2,345	

[a] The results when stock price volatility is measured by the annual average of (daily high-daily low)/{(daily high + daily low)/2}.
[b] The results when VOLATILITY (instead of SPREAD) is employed as a proxy variable for the expected profit associated with private information.
[c] See footnote a in Table 1 for the definition of the variables.
[d] See footnote b in Table 4 for the expected signs of regression coefficients.
** Significant at the 1% level.
* Significant at the 5% level.

information. Our results confirm the earlier finding of Bhushan (Bhushan, 1989a; Bhushan, 1989b) and Moyer et al. (1989) that the aggregate demand for analyst services is an increasing function of a firm's price volatility.

4. Summary

In this paper, we present empirical results which support the notion that two major players of financial markets (i.e., market makers and financial analysts) make their decisions interactively. More specifically, the results show that market makers establish the bid-ask spread of a stock according to how many financial analysts are following that stock. Similarly, the results show that more financial analysts follow a stock with a greater spread. These results are consistent with the view that one group deduces the extent of informational asymmetry associated with a stock from the behavior of the other group.

Although much of the theoretical literature emphasizes the effect of informational asymmetry on the bid-ask spread, corresponding empirical evidence is comparatively scarce. This study provides empirical evidence on the significance of the information hypothesis of bid-ask spreads by employing new measures of information asymmetry and by treating the production of information endogenous. Our results also reinforce the notion that the demand for (and in equilibrium, the supply of) information about securities is positively related to the level of information asymmetry.

Acknowledgements

The authors thank *I/B/E/S* Inc. for providing earnings expectations data. The authors are grateful to two anonymous referees for their very helpful comments. The authors are solely responsible for the contents of the paper.

References

Amihud, Y. and H. Mendelson, 1989, The effects of beta, bid-ask spread, residual risk, and size on stock returns, Journal of Finance 44, 479-486.

Amihud, Y., H. Mendelson and R. A. Wood, 1990, Liquidity and the 1987 stock market crash, Journal of Portfolio Management 16, 65-69.

Bagehot, W., 1971, The only game in town, Financial Analysts Journal 22, 12-14.

Barry, C.B. and S.J. Brown, 1985, Differential information and security market equilibrium, Journal of Financial and Quantitative Analysis 20, 407-422.

Benston, G.J. and R.L. Hagerman, 1974, Determinants of bid-ask spreads in the over-the-counter market, Journal of Financial Economics 1, 353-364.

Belsley, D.A., E. Kuh, and R.E. Welsch, 1980, Regression Diagnostics (John Wiley and Sons, Inc., New York, NY).

Bhushan, R., 1989a, Collection of information about publicly traded firms, Journal of Accounting and Economics 11, 183-206.

Bhushan, R., 1989b, Firm characteristics and analyst following, Journal of Accounting and Economics 11, 255-274.

Box, G.E.P. and D.R. Cox, 1964, An analysis of transformation, Journal of the Royal Statistical Society, Series B 26, 211-243.

Brennan, M.J. and P.J. Hughes, 1991, Stock prices and the supply of information, Journal of Finance 46, 1665-1692.

Brown, L.D., R.L. Hagerman, P.A. Griffin and M.E. Zmijewski, 1987, Security analyst superiority relative to univariate time-series models in forecasting quarterly earnings, Journal of Accounting and Economics 9, 61-87.

Brown, L.D. and M.S. Rozeff, 1978, The superiority of analyst forecasts as measures of expectations: Evidence from earnings, Journal of Finance 33, 1-16.

Chiang, R. and P.C. Venkatesh, 1988, Insider holdings and perceptions of information asymmetry: A note, Journal of Finance 43, 1041-1048.

Copeland, T. and D. Galai, 1983, Information effects on the bid-ask spread, Journal of Finance 38, 1457-1469.

Cragg, J.G. and B.G. Malkiel, 1982, Expectations and the structure of share prices (The University of Chicago Press, Chicago, IL).

Demsetz, H., 1968, The cost of transacting, Quarterly Journal of Economics 82, 33-53.

Easley, D. and M. O'Hara, 1987, Price, trade size, and information in securities market, Journal of Financial Economics 19, 69-90.

Elton, E., M. Gruber and M. Gultekin, 1981, Earnings expectations and share prices, Management Science 27, 975-987.

Elton, E., M. Gruber and M. Gultekin, 1984, Professional expectations: Accuracy and diagnosis of errors, Journal of Financial and Quantitative Analysis 19, 351-363.

Fried, D. and D. Givoly, 1982, Financial analysts' forecasts of earnings: A better surrogate for market expectations, Journal of Accounting and Economics 4, 85-107.

Fuller, W.A. and G.E. Battese, 1974, Estimation of linear models with crossed-error structure, Journal of Econometrics 2, 67-78.

Glosten, L.R. and P.R. Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, Journal of Financial Economics 14, 71-100.

McInish, T.H. and R.A. Wood, 1991, Hourly returns, volume, trade size, and number of trades, Journal of Financial Research 14, 303-315.

McInish, T.H. and R.A. Wood, 1992, An analysis of intraday patterns in bid/ask spreads for NYSE stocks, Journal of Finance 47, 753-764.

Merton, R.C., 1987, A simple model of capital market equilibrium with incomplete information, Journal of Finance 42, 483-510.

Moyer, R.C., R.E. Chatfield and P.M. Sisneros, 1989, Security analyst monitoring activity: Agency costs and information demand, Journal of Financial and Quantitative Analysis 24, 503-512.

O'Brien, P.C. and R. Bhushan, 1990, Analyst following and institutional ownership, Journal of Accounting Research 28, 55-82.

Pearson, N.D., 1991, Determinants of the production of information (Working Paper, University of Rochester).

Stoll, H.R., 1978, The pricing of security dealer services: An empirical study of NASDAQ stocks, Journal of Finance 33, 1153-1172.

Tinic, S.M. and R.R. West, 1972, Competition and the pricing of dealer service in the over-the- counter stock market, Journal of Financial and Quantitative Analysis 7, 1707-1727.