Journal of Financial Economics 1 (1974) 353-364. © North-Holland Publishing Company

DETERMINANTS OF BID-ASKED SPREADS IN THE OVER-THE-COUNTER MARKET*

George J. BENSTON

Graduate School of Management and Center for Research in Government Policy and Business, University of Rochester, Rochester, N.Y. 14627, U.S.A.

Robert L. HAGERMAN

School of Management, State University of New York, Buffalo, N.Y. 14214, U.S.A.

Received July 1974, revised version received August 1974

Security market regulators, among others, are concerned to know whether or not dealers are natural monopolists. Based on a randomly drawn sample of 314 over-the-counter stocks, the results of this study suggest that while there are economies of scale, they are not on the dealer level. In addition, both systematic and unsystematic risk were tested for association with the transaction costs in this market. The evidence suggests unsystematic risk is related to spread.

1. Introduction

The mark-up charged by dealers to consumers in the securities market, as in any other market, is a function of the operational efficiency of the dealers and the nature of the product. Because the security markets are regulated, the specific determinants of this mark-up need to be estimated to answer public policy questions as well as to satisfy intellectual and managerial interest in the dealers' production functions. The importance of these determinants is illustrated by the recent debate over whether or not specialists are natural monopolists, a question central to the furor over the relationship between the New York Stock Exchange (NYSE) and the third market. These questions make it essential that the nature of transactions costs in these markets be understood. The purpose of this paper is to analyze the determinants of spread in the over-the-counter market (OTC), to determine if the dealership function is a natural monopoly and to test other hypotheses.[1]

---

*We wish to acknowledge the helpful suggestions made by Michael Canes, Michael Jensen and an anonymous referee.

[1] Several other studies analyze the determinants of the spread between bid and asked per share prices. Demsetz (1968) developed a theory of transactions costs in the securities markets (on which we rely, in large measure) and provides some empirical verification of the theory by analyzing the specialists' spread on NYSE stocks. Tinic and West (1972) used Demsetz's analysis to study the spreads on OTC stocks. These studies made important contributions to the theory and measurement of transactions costs but, as the authors pointed out, the data used are not sufficient to allow more than tentative support for Demsetz's theory. In addition, the treatment of risk in both studies is inadequate (Demsetz does not discuss risk; Tinic and West use a poor measure).

The analysis is based on standard demand theory. The product offered by security dealers [as Demsetz (1968) points out] is an immediate exchange of titles to securities instead of a delayed exchange. Dealers provide this immediate exchange by matching buy and sell orders and by holding an inventory of securities which is used to fill unmatched orders. The price charged for this product is the spread, the difference between the buying (bid) and selling (asked) price per share. The spread is a function of the market demand curve (the amount of immediacy demanded by investors), the competitiveness of the market, and the dealers' cost curves. In this study we take investors' demand for immediacy as given, and analyze per share spreads as a function of dealers' costs and market structure. This analysis allows testing of hypotheses about whether natural monopoly characterizes the share-trading market, whether the market is competitive and the prevalence and effects of insiders.

2. Determinants of the bid–asked price per share

An important factor affecting the dealers' costs is the amount of inventory required to provide the immediate transfer of shares they offer to investors. The amount of inventory a dealer must carry of a particular stock is a function of the volume of that stock's transactions. As volume increases so does the number of limit orders, which facilitate immediate exchange. These limit orders are a substitute for inventory; the greater the number of transactions, the lower the amount of inventory that must be held per transaction. Even without considering limit orders, standard inventory theory suggests that the inventory a dealer must hold to effect trading immediacy is less than a proportionate function of the number of transactions he expects to make. Thus the per unit cost of immediacy, i.e., the spread, should decline as the transactions rate for the security increases. The elasticity of the spread with respect to the number of transactions provides a measure of economies of scale from dealing in a particular stock, cet. par.

Inventory carrying costs per unit are a positive function of the riskiness of holding the inventory, if dealers are risk averse and are unable to eliminate the risk by portfolio diversification. (Since the concept of risk is not discussed extensively in previous studies,[2] an elaboration is provided in the following section.) Unlike most commodities, however, the cost of maintaining an inventory of securities does not include losses in value due to deterioration (although pilferage can be a problem). The cost of capital is also not a relevant

---

[2] Demsetz doesn't discuss risk. Tinic and West's (1972) basic discussion is the following: 'Our initial notion was to hypothesize a positive relationship between spreads and price volatility on the grounds that the greater the variability in price, the greater the risk associated with performance of the dealership function. On further reflection, however, we concluded that we should not try to predict the sign of this coefficient since it might be possible for the influence of price volatility to be negligible if a dealer could diversify his operations sufficiently.' Tinic's study (1970) of spreads on the NYSE reported in Tinic (1972) uses the standard deviation of the price of a security, presumably as a measure of risk (although no explicit rationale for inclusion of the statistic is given in the brief review of his analysis).

cost of carrying the inventory, since the returns from holding the securities normally reflect the opportunity cost of the capital invested. Thus, the inventory carrying costs are primarily due to the risks incurred in holding the inventory.

Dealers also incur costs in matching buy and sell orders. If economies of scale characterize these transactions, per share spreads should be a function of the volume of trades in a specific stock. Transactions costs also may be related to the dollar amount traded. While transactions are stated in terms of the number of shares traded, market participants trade basically in dollar denominated claims. Were all factors other than price per share equal, traders would use limit orders to equalize the spread per dollar regardless of the price per share traded.[3] Consequently, spreads would be proportional to the per share price. This strict proportionality might be eliminated by disportionate broker costs since, if it is costly per dollar traded to enter the market for low-priced securities, the arbitrage mechanism could not equalize the spread per dollar. Thus spread should be positively, though not necessarily proportionally, related to the price of a stock.

Trading with insiders increases the dealers' costs and hence affects the per share spreads as Bagehot (1971) has pointed out. Insiders (by definition) have information which dealers do not. If they cannot identify the traders who are insiders, dealers must increase spreads on those shares which they believe are traded by insiders.

Finally, the extent of competition, measured by the number of dealers who compete in making a market for a stock, should be reflected by the spread. A large number of dealers should keep the spread down to the competitive level. It is also possible that smaller spreads are associated with a larger number of dealers because the presence of other dealers allows any one dealer to offset a temporary inventory imbalance with interdealer trading. The two factors suggest that spread should be negatively related to the number of dealers making a market in the stock.

It should be noted that the number of dealers and the number of shareholders are likely to be correlated with each other since larger companies have more stockholders and more dealers who are interested in making a market in the stock. To the extent that these variables are correlated with company size, their coefficients may measure the relation between the size of the firm and the spread changed by dealers. This proxy relationship should be remembered when the coefficients are interpreted.[4]

In summary, standard economic theory applied to the market for immediate transfer of titles to shares, indicates that

$$\mathrm{SP} = f(\mathrm{NT}, \mathrm{PS}, \mathrm{HR}, \mathrm{IR}, \mathrm{ND}), \qquad (1)$$

---

[3] There is some belief that lower priced shares, as such, have greater variation in price than do higher priced shares. However, Heins and Allison (1966) show that this belief is groundless. Also, as is discussed below, it is irrelevant as a determinant of spreads.

[4] There is also some reason to believe that residual variance, which is our measure of holding risk, is negatively related to company size.

where

SP = spread per share, the price of an immediate transfer of title;

NT = number of transactions in a stock;

PS = price per share;

HR = holding risk due to holding a stock in inventory whose price might change (up or down);

IR = insider losses due to trading with insiders in a stock which, if purchased, is likely to go down in price or, if sold, is likely to go up in price more than expected;

ND = number of competing dealers making a market in a stock.

The relationship between SP and NT provides an estimate of economies of scale that results from savings in inventory and transactions costs, cet. par. The number of transactions in a particular stock by a given dealer and the spread charged by him would be most appropriate for this estimate. Though market spreads are analyzed, appropriate inclusion of the number of dealers in the analysis allows making an estimate of the elasticity of spreads with respect to the total number of transactions, given the number of dealers. (Some additional evidence is brought in below to delineate market from individual dealer economies of scale.) The relationship of SP and HR, cet. par., also provides a measurement of the extent to which dealers can diversify risk and are risk averse. The relationship of SP and IR, cet. par., provides a measure of the extent and cost to dealers of trading with insiders. The relationship of SP and ND provides an indication of the effect of degrees of competition on the price of immediate stock title transfers. PS serves as a 'homogeneity' variable with respect to the transactions costs of transferring titles.

3. Specification of the variables and sources of data