Assessing the Costs of Regulation: The Case of Dual Trading
Author(s): Tom Smith and Robert E. Whaley
Source: *The Journal of Law & Economics*, Vol. 37, No. 1 (Apr., 1994), pp. 215-246
Published by: The University of Chicago Press for The Booth School of Business, University of Chicago and The University of Chicago Law School
Stable URL: https://www.jstor.org/stable/725610

ASSESSING THE COSTS OF REGULATION: THE CASE OF DUAL TRADING*

TOM SMITH and ROBERT E. WHALEY
Duke University

THE practice of dual trading in futures exchanges across the United States has existed since the advent of organized futures markets in the mid-1800s. Under a dual trader market-making system, market makers are permitted to execute transactions for customers and on personal account. With two sources of income to cover the costs of operation—brokerage commissions and dealer/speculator profits—dual trader markets have greater numbers of market makers than otherwise comparable markets that permit traders to earn income from only brokerage or dealing/speculation. With more market makers, the level of competition for providing the market-making service is increased, thereby increasing market liquidity and lowering the costs of trading (that is, brokerage commissions and the bid/ask spread).[1]

Dual trading prevails on security and futures exchanges in the United States as well as in other countries. Grossman provides a comprehensive list of dual trader markets worldwide.[2] In futures markets, a substantial proportion of total trading volume is executed by dual traders. The Com-

* We thank Todd Petzel at the Chicago Mercantile Exchange for providing us with information regarding the top-step rule in the S&P 500 index futures pit. Margaret Monroe provided extremely useful and insightful comments on an earlier draft of this paper. Comments and suggestions by Steven Figlewski, F. Douglas Foster, Hans R. Stoll, Joseph R. Sweeney, an anonymous referee, and the seminar participants at Duke University, Indiana University, Southern Methodist University, the University of Oklahoma, the 1990 Northern Finance Association meetings in Banff, Alberta, and the 1992 Western Finance Association meetings in San Francisco are also gratefully acknowledged. Research support was received from the Business Associates' Fund (Smith) and the Isle Maligne Fund (Whaley).

[1] Other arguments in support of dual trader markets have also been suggested. Some argue that dual traders operate more efficiently than do exclusive brokers because their own money is at risk. Others argue that dual traders can provide lower transaction costs because they are better informed about the market and/or have superior trading skills as a result of handling both personal and customer orders.

[2] Sanford J. Grossman, An Economic Analysis of Dual Trading 62–72 (working paper, University of Pennsylvania, Wharton School, Rodney L. White Center for Financial Research, 1989).

[*Journal of Law and Economics*, vol. XXXVII (April 1994)]

modity Futures Trading Commission (CFTC) study of futures exchanges in the United States, for example, reports that over 50 percent of trading volume is executed by individuals who transact for themselves as well as customers.[3]

Dual trading is criticized because when a broker can trade on behalf of customers and on his own account a potential conflict of interest arises. One such conflict is called "front-running." Consider a broker who receives a large customer buy order that will surely move the futures price upward. A dual trader could buy futures on personal account first and then execute his customer's order. When execution of the customer order drives the price upward, the broker liquidates his position at a profit. Front-running is difficult to detect.[4] Consequently, the CFTC, from time to time, contemplates restricting or banning dual trading in futures markets.

Fishman and Longstaff analyze the effects that restricting dual trading may have on trading costs.[5] In markets where dual traders do not front-run, they show that a ban on dual trading will decrease bid/ask spreads. The intuition for this result is that market makers can afford to decrease spreads because they will face fewer information-based (money-losing) trades. To compensate for lost profits on dual trading, brokers raise commissions. In markets where dual traders front-run, Fishman and Longstaff show that banning dual trading may or may not increase spreads.

Restricting or banning dual trading activity, however, provides no assurance that front-running activity or trading costs will be reduced. Even in the absence of dual trading, brokers could conspire with others to profit from the trades of informed customers. While the theoretical argu-

[3] Commodity Futures Trading Commission, Economic Analysis of Dual Trading on Commodity Exchanges (report prepared by the Division of Economic Analysis, November 1989). This proportion is likely understated because, in the CFTC study, dual traders are defined as traders who act for customers and for personal account in the same commodity on the same day. Such a definition fails to include traders who execute transactions for customers in one commodity contract and for themselves in a closely related commodity contract such as a futures option contract, an option contract, or a futures contract on a commodity that is a close substitute.

[4] Detailed time and sales data compiled by the futures exchanges safeguard, as much as possible, such trading abuses. Sanford J. Grossman & Merton H. Miller, Economic Costs and Benefits of the Proposed One-Minute Time Bracketing Regulations, 6 J. Futures Markets 145 (1986), for example, examine the costs and benefits of implementing a one-minute time bracketing regulation and find that "the upper bound to the economic benefits from eliminating abusive dual trading with one-minute bracketing comes to less than 1 cent per contract." Nevertheless, handheld electronic trading card systems that provide instantaneous information about the trading activity of each broker/market maker are being contemplated. If implemented, these systems should alleviate concerns about possible abusive trading practices.

[5] Michael J. Fishman & Francis A. Longstaff, Dual Trading in Futures Markets, 47 J. Fin. 643 (1992).

ments about the effects of restricting dual trading can be persuasive, ultimately, the resolution of the debate will come through empirical examination and evidence. Unfortunately, comparing the operations of dual and nondual trader markets is not as easy as it may seem. There are few instances where dual and nondual traders markets on the same futures operate simultaneously or where a single futures contract market has operated with exclusively dual and nondual trading during different periods of time. The focus of this analysis is on one instance when a futures exchange voluntarily restricted dual trading in an active futures pit—in June 1987, the Chicago Mercantile Exchange (CME) implemented a "top-step rule," which permits only exclusive brokers access to the top step of the S&P 500 index futures pit. An effect of the rule has been to curtail substantially the volume of trading executed by dual traders. The purpose of this study is to assess the economic cost of regulating dual trading by estimating the increase in the effective bid/ask spread resulting from the implementation of the CME's top-step rule in the S&P 500 index futures market.

The outline of the article is as follows. Section I discusses the possible economic implications of restricting dual trading, including increases in the effective bid/ask spread and brokerage commission rates and a decrease in market liquidity. Our focus is primarily on the cost of immediacy in futures markets—the effective bid/ask spread. Differences or changes in the spread are a means of assessing the cost of restricting trading. In Section II, the measurement of the effective bid/ask spread from futures exchange times and sales data is discussed. Section III contains the empirical analysis. The implementation of the CME's top-step rule in June 1987 appears to have increased the effective bid/ask spread in the S&P 500 futures market. Section IV concludes the study with a brief summary.

I. Restricting Dual Trading and Market Maker Spreads

Futures trading is conducted in a pit on an exchange floor by a number of different types of traders. Some traders are *exclusive brokers,* who execute transactions on behalf of customers (hedgers and long-term speculators) whose orders are called in. Other traders are *exclusive dealers* or *scalpers,* who make markets by quoting bid and ask prices for customers who demand immediate execution. Others are *exclusive short-term speculators,* who trade only for their own account and attempt to profit from intraday price movements. Yet another group of traders are *dual traders,* who may act both as broker and dealer/speculator. Some dual traders tend to scalp more than broker. Occasionally, they act as broker when customer demand is high. In contrast, some dual traders earn their income primarily from brokerage commissions and scalp/speculate only when customer demand is low. Seldom do any of these traders, dual or nondual, hold positions overnight.

This study deals with the possible consequences of restricting or banning dual trading. In the absence of trading abuse and fraud, the cost of trading cannot be reduced by forcing dual traders to operate exclusively as brokers or dealer/speculators. Presumably, both sources of income are important to dual traders, otherwise they would voluntarily choose to engage exclusively in one activity or the other. Likely outcomes of restricting or banning dual trading, therefore, are that dual traders who predominantly scalp will increase bid/ask spreads to compensate for lost brokerage commissions and dual traders who predominantly broker will increase brokerage commission rates to compensate for lost dealer/speculator profits. In addition, the supply of both services is likely to be reduced, thereby reducing market liquidity.

In this study, we focus on the effects of restricting dual trading on the effective bid/ask spread. In June 1987, the CME voluntarily implemented the top-step rule in the S&P 500 futures pit. Under this rule, only exclusive brokers are permitted access to the top step of the pit, where a high proportion of total transactions are executed. If the effect of this rule has been to reduce competition and hence market liquidity, spreads should be higher in the period after the implementation of the rule than before. In contrast, if the restricted dual trader market after the implementation of the rule is no less competitive, the effective spread should be no different.

Before examining the differences between the effective spreads during the two time periods, however, it is necessary to have a clear understanding of the economic determinants of the effective bid/ask in the futures markets. This section is divided into three parts. In the first, the economic determinants of market maker spreads are discussed, and a summary of empirical results of investigations in stock markets is provided. In the second, the focus turns to modeling the bid/ask spread in futures markets. Finally, in the third subsection, we discuss how the economic benefits/costs of dual trader markets may be assessed in part by examining the effects of the implementation of the top-step rule.

A. *Determinants of the Bid/Ask Spread*

Models of market maker spreads have appeared in the finance and economics literature for many years.[6] In general, these studies model the

[6] The first model to appear in the literature is that of Harold Demsetz, The Cost of Transacting, 82 Q. J. Econ. 33 (1968). Subsequent stock market studies include Seha Tinic,

market maker's bid/ask spread, $SPRD_t$, as a function of the security transaction rate (the number of contracts traded per day, for example), VOL_t, the number of competing dealers, ND_t, and risk, $RISK_t$, that is,

$$SPRD_t = f(VOL_t, ND_t, RISK_t). \tag{1}$$

The market maker's spread should vary inversely with volume of trading. The higher the rate of inventory turnover, the less the market maker needs to charge per transaction. The spread should also vary inversely with the number of competing market makers since competition presumably causes each market maker to operate more efficiently. Finally, since the value of the portfolio of securities that the market maker holds at a particular point of time may change unexpectedly, the market maker needs to be compensated for risk.[7]

Empirical work examining various specifications of (1) is largely focused on stock markets and concludes that the determinants posited are well-supported by the data. Benston and Hagerman, for example, find that spread varies significantly and inversely with number of shareholders and the number of competing dealers and significantly and directly with unsystematic risk for a sample of 314 over-the-counter (OTC) stocks during the five-year period 1963–67.[8] The number of shareholders serves as a proxy variable for volume of shares traded since share volume for the sample's stocks was unavailable. The use of the unsystematic risk variable as a risk variable is motivated by the fact that the market maker is expected to earn a rate of return commensurate to his inventory's systematic risk level. Only to the extent that the market maker is not well-diversified (that is, to the extent that the market maker sustains unsystematic risk) should he be rewarded through the bid/ask spread. Price per share is also included as an independent variable since the regression is cross-sectional. Holding other factors constant, a twenty-dollar stock

The Economics of Liquidity Services, 86 Q. J. Econ. 79 (1972); George Benston & Robert Hagerman, Determinants of Bid/Asked Spreads in the Over-the-Counter Market, 1 J. Fin. Econ. 353 (1974); Hans R. Stoll, Dealer Inventory Behavior: An Empirical Investigation of NASDAQ Stocks, 11 J. Fin. Quant. Anal. 359 (1976); Hans R. Stoll, The Supply of Dealer Services in Securities Markets, 33 J. Fin. 1133 (1978); and Hans R. Stoll, The Supply of Dealer Services: An Empirical Analysis of NASDAQ Stocks, 33 J. Fin. 1153 (1978). Sarahelen R. Thompson & Mark L. Waller, Determinants of Liquidity Costs in Commodity Futures Markets, 7 Rev. Futures Markets 110 (1988), examine liquidity costs in the corn and oats futures pits at the CBT.

[7] The model presented here focuses only on the order processing and inventory holding costs incurred by the market maker. While adverse information costs could also be modeled within this framework, we chose to mitigate the effects of adverse information by the sample selection procedure described in Section III.

[8] Benston & Hagerman, *supra* note 6. To be more precise about model specification, Benston and Hagerman use the logarithm of variables indicated.

should have a greater spread than a ten-dollar stock since the market maker's total carrying cost is greater. Other investigators such as Stoll control for this effect by using the relative spread (the bid/ask spread divided by the average of the bid and ask prices) as the dependent variable.[9]

B. *Modeling the Bid/Ask Spread in Futures Markets*

The theoretical models of the market maker's spread have generally been tested using stock market data, although the models apply equally well for bid/ask spreads in other markets such as commodity futures markets. Like stock markets, spreads in futures markets should decrease with the volume of trading (economies of scale) and with the number of competing traders (competition). Risk is also an important determinant of the effective spread because the market maker faces the prospect of a significant market movement while he has an open position in the futures.

Risk may not enter the bid/ask spread model in the futures market regression as strongly as it does in the regression model for the stock market. In the stock market, the market maker cannot effectively hedge the value of his inventory. A specialist on the New York Stock Exchange (NYSE), for example, is prohibited from hedging the risk of his stock position using stock options. In contrast, a futures market maker can more effectively manage his risk exposure by hedging with different contract months in the same commodity, futures options, options, or futures on other commodities. In addition, futures market makers can minimize their risk exposure by quickly reversing positions acquired. Silber estimates that the average length of time that a scalper holds a position is about two minutes.[10] In fact, Silber reports that the scalper's returns are negative for positions held longer than three minutes.

Finally, studies of stock market spreads adjust for the price per share of the stock. This is done to account for the differences in spreads across stocks resulting from different capital investments. Higher-priced stocks have higher carrying costs, so the spread must be measured relative to the amount of capital tied up (price per share). In futures markets, however, no capital investment is required. Not even futures margin should affect the size of the spread since market makers (scalpers) only day trade and have no need to post margin.

[9] Stoll, The Supply of Dealer Services: An Empirical Analysis, *supra* note 6.

[10] William Silber, Marketmaker Behavior in an Auction Market: An Analysis of Scalpers in Futures Markets, 39 J. Fin. 950 (1984). This behavior is documented for the New York Futures Exchange's NYSE Composite index futures contract during the period December 1, 1982–January 14, 1983.

C. Assessing the Cost of Restricting Dual Trading

The economic models outlined can, in principle, be used to distinguish between dual traders and exclusive dealers as market makers in the futures pit. The problem of differentiating between the two types of market makers is that all active commodity futures markets currently permit dual trading. Bid/ask spreads may well be lower when dual traders are present in the market than when they are not or vice versa, but examining spreads in a market where both dual traders and exclusive dealers are permitted will not make this apparent. In a competitive market, bid/ask spreads offered by dual traders and exclusive dealers should be the same. If exclusive dealers offered spreads in excess of those of dual traders, they would attract little business from customers and would either lower the price of their services or be forced from the market. However, if dual traders abused their position by trading ahead of customer orders, the effective costs of executing a transaction through dual traders would be higher than through exclusive dealers, so subsequent customer transactions would be placed through exclusive dealers. In a competitive market, the only situation in which dual traders and exclusive dealers can coexist is where both charge the same fee for providing market liquidity (that is, the effective bid/ask spread).[11]

To make a direct comparison of the effectiveness of dual trader and nondual trader markets, the commodity underlying the futures contract in the two markets has to be identical—either (*a*) dual and nondual traders markets on the same commodity operating simultaneously, or (*b*) a single market in which there have been periods of exclusively dual and exclusively nondual trading. In general, few such situations arise. While simultaneous dual/nondual markets exist, the market activity in the nondual trader market is so low that effective comparisons are not possible.[12] Approach (*a*) is therefore infeasible. With respect to approach (*b*), the first instance in which a futures market voluntarily restricted dual trading was on June 22, 1987, when the CME implemented a top-step rule in the S&P 500 index futures pit.[13] Under this rule, only exclusive brokers are allowed access to the top step of the pit, where a high proportion of trades are executed. An effect of this rule has been to reduce substantially

[11] This argument does not necessarily apply to brokerage services since brokerage customers are interested in both low commissions and quick execution.

[12] For example, foreign currency futures contracts trade in dual trader markets at the CME in Chicago and in nondual trader markets at the Philadelphia Board of Trade (PBOT) in Philadelphia.

[13] Recently, the CME prohibited dual trading in all contracts that trade an average of at least 10,000 contracts a day.

the proportion of trades executed by dual traders. In Section III, we investigate empirically the effects that the top-step rule has had on the trading volume and the effective bid/ask spread in the S&P 500 futures market. Prior to conducting the investigation, however, it is necessary to develop a methodology for estimating the effective spread.

II. Estimating the Effective Bid/Ask Spread

To be precise in our assessment about the effect of the implementation of the top-step rule on the cost of immediacy, an accurate estimator of the effective bid/ask spread is necessary. The effective or realized spread differs from the quoted spread. The quoted spread is the difference between the market maker's bid and ask quotes, while the effective spread is the difference between the price at which the market maker buys (sells) the futures and the price at which he subsequently sells (buys) it.[14] Our focus is on the effective spread since it is the cost incurred in aggregate by the customers using the futures market.

Measuring directly the effective spread in futures markets is not possible because the trading records of the market makers are not publicly available. Instead, the spread must be inferred from the transaction prices. Futures exchanges do not record a complete history of the transaction prices and trading volumes—they record only "time and sales" data. Time and sales data include only the time and the price of each futures contract transaction, where the price in the transaction is different from the price recorded previously. Bid (ask) prices also appear on the time and sales data file if the bid (ask) price exceeds (is below) the previously recorded transaction price. No trading volume figures are recorded.

Two types of estimators have been used to measure the effective bid/ask spread in the past research using time and sales data. One estimator is based on the first-order, serial covariance of price changes. Roll demonstrates that if true price changes are serially dependent, the effective bid/ask spread can be estimated as

$$s_A = 2\sqrt{-\text{cov}(\Delta P_t^0, \Delta P_{t-1}^0)}, \tag{2}$$

[14] To illustrate, assume that the quoted spread for the S&P 500 futures is .05 (250.00 bid/250.05 ask). Assume further that in the course of market making, a scalper buys fifty contracts at 250.00. Current customer market orders to buy might absorb only part of the scalper's position, say, forty contracts at 250.05. Since the remaining ten contracts pose significant risk, the scalper may attempt to liquidate the position more quickly by temporarily lowering his ask price to, say, 250.00, making his offer the most attractive in the pit. Further customer market orders will automatically flow to the scalper, and his remaining position will be liquidated (through "scratch sales"). While the quoted spread in this example is .05, the effective spread is .04.

where P_t^0 is the observed futures price at time t.[15] In this model, if the observed price of the futures is at the bid (ask), the next price change transaction is equally likely to be zero or plus (minus) the amount of the spread. While the use of (2) seems relatively straightforward, estimates of $\mathrm{cov}(\Delta P_t^0, \Delta P_{t-1}^0)$ using futures time and sales data are frequently positive. This arises because the negative serial dependence induced by futures market bid/ask spreads is small relative to the short-term positive serial dependence in true price changes resulting from new information. With positive serial covariance, the estimate of s_A is meaningless.[16]

The other estimator used to infer the effective spread is the mean absolute value of price change,

$$s_B = \overline{|\Delta P^0|} = \frac{1}{T}\sum_{t=1}^{T} |\Delta P_t^0|, \tag{3}$$

where T is the length of the futures price change series. Thompson and Waller and the CFTC study endorse this approach.[17] If the expected true futures price change and the variance of true price change are both zero, this estimator would capture the effective spread. Unfortunately, while it is reasonable to assume that the expected price change of the futures from transaction to transaction is zero, assuming that the variance of futures price changes is zero is unrealistic. The mean absolute price change, $\overline{|\Delta P^0|}$, therefore consists of two components—the bid/ask spread and the variance of true price changes. The estimator $\overline{|\Delta P^0|}$ is therefore upward-biased, with the magnitude of the bias depending on the variance of true price changes.

Smith and Whaley offer a new approach. They[18] develop a method-of-moments estimator of the effective bid/ask spread that explicitly recognizes that the mean absolute price change contains the price change variance attributable to the bid/ask spread as well as the variance of true price changes. To isolate the two effects, they assume that true price

[15] Richard Roll, A Simple Implicit Measure of the Bid/Ask Spread in an Efficient Market, 39 J. Fin. 1127 (1984).

[16] The problem of imaginary numbers is not the only problem with using the Roll estimator on time and sales data. Roll's model assumes that there is a 50 percent chance that the next observed transaction price will be at its current level. Time and sales data generally do not include consecutive transactions at the same price. If the current transaction price is at a bid (ask) level, in all likelihood the next recorded transaction will be at an ask (bid) level.

[17] Thompson & Waller, *supra* note 6; and Commodity Futures Trading Commission, *supra* note 3.

[18] Tom Smith & Robert E. Whaley, Estimating the Effective Bid/Ask Spread Using Time and Sales Data, J. Futures Markets (June 1994, in press).

changes are distributed normally with zero mean and variance, σ^2, and that bid and ask prices occur with equal probability. Under these assumptions, the expected value of the absolute price change is

$$E(|\Delta P^0|) = \sqrt{\frac{2}{\pi}}\sigma e^{-s_C^2/2\sigma^2} - s_C + 2s_C N\left(\frac{s_C}{\sigma}\right), \qquad (4)$$

where s_C is the bid/ask spread, σ^2 is the variance of true price changes, and $N(d)$ is the cumulative unit normal distribution function with upper integral limit d. Note that even if the bid/ask spread is zero, the expected absolute price change is positive. Substituting $s_C = 0$ into (4) shows that the mean of the distribution of $|\Delta P^0|$ in the absence of a bid/ask spread is

$$E(|\Delta P^0|) = \sigma\sqrt{\frac{2}{\pi}}. \qquad (5)$$

Equation (4), by itself, cannot be used to estimate s_C since both s_C and σ^2 are unknown. In order to develop an estimator of the effective spread, a second equation is needed and can be obtained using the expected value of the squared price change, that is,

$$E(|\Delta P^0|^2) = \sigma^2 + s_C^2. \qquad (6)$$

To estimate the effective spread (and the variance of true price changes), replace $E(|\Delta P^0|)$ and $E(|\Delta P^0|^2)$ with the mean absolute price change and the mean squared price change from the observed futures price change distribution and then perform a nonlinear minimization on the relations (4) and (6), constraining the volatility estimate to be nonnegative.

III. Top-Step Rule Investigation

On June 22, 1987, the CME implemented a top-step rule in the S&P 500 index futures market. Under this rule, only exclusive brokers are allowed access to the top step of the pit, where a high proportion of total transactions are consummated. As a result, most transactions are now executed by brokers, and the amount of trading volume executed by dual traders has been substantially curtailed. Where 50 percent of trading volume was executed by dual traders in February 1987, dual traders accounted for only 11 percent of volume in September 1987.[19]

In this section, the economic effect of the top-step rule is investigated. If there is no difference between markets that permit dual trading and

[19] These figures are taken from excerpts of a 1990 letter from the CME to the CFTC summarizing the effects of the introduction of the top-step rule. We thank Todd Petzel at the CME for providing these excerpts.

those that do not, the rule should have had no effect on the effective bid/ask spread. However, if the restriction reduces competition, the effective spread may have increased. This section is divided into six parts. In the first, the data are described. The sources of the price and volume information and the construction of the variables used in our analysis are provided. In the second subsection, changes in the pattern of trading volume across contract months before and after the introduction of the top-step rule are examined. The third subsection investigates the empirical properties of the mean absolute price change and method-of-moments estimators (s_B and s_C) that were discussed in Section II. The fourth subsection describes the simple and multiple regression models that are used to assess changes in the effective bid/ask spread resulting from the introduction of the top-step rule. The fifth subsection contains the regression results for the S&P 500 futures market. In the sixth subsection, the effective spreads and trading volumes of the S&P 500 and Major Market Index (MMI) futures contract markets are compared.

A. *Data*

To investigate the effect of the top-step rule, we examine the time and sales data for the S&P 500 index futures contracts for the period from contract inception on April 21, 1982, through one week prior to the October 1987 stock market crash, October 9, 1987. These data were made available by the Chicago Mercantile Exchange. In addition, we examine the time and sales data for the (MMI) futures contracts during the period January 2–October 9, 1987. These data were made available by the Chicago Board of Trade (CBT). Since the CBT placed no restriction on dual trading in the MMI futures pit, the MMI futures contract market can serve as a control group.

By way of history, the CME introduced trading of the S&P 500 futures contract on April 21, 1982. These futures contracts followed a quarterly expiration cycle (March, June, September, December) and generally had times to maturity as long as one year. For a short period of time, the CME experimented with contract maturities extending out more than one year. On June 17, 1983, six quarterly contract maturities became available. On June 18, 1984, this number dropped to five, and on September 24, 1984, the number fell to four.

The CBT introduced trading of the MMI futures contract in July 1984. The original MMI futures contract was relatively small in denomination (100 times the index value), had minimum price increments of .10 index points, and followed a monthly expiration cycle. In July 1985, the CBT introduced a larger MMI futures contract (250 times the index level), the

"Maxi MMI," with minimum price increments of .05 index points and a quarterly expiration cycle. By mid-August 1985, this larger contract had assumed the greatest trading volume. Subsequently, the smaller contract was delisted and a monthly expiration cycle was introduced for the larger contract. Our analysis includes only the larger MMI contract during a period in which the contract was traded on a monthly expiration cycle—January 2–October 9, 1987.

From the time and sales data of each futures exchange, two samples are created. The first sample, Sample 1, treats all records, including records reporting bid and ask quotes, as if they are actual transactions.[20] The second sample, Sample 2, excludes bid and ask quotes as well as prices of transactions adjacent to quotes. The prices in adjacent transactions are eliminated because the price movement from the transaction record prior to a reported bid (ask) record to the transaction record after the bid (ask) record likely represents a price movement due to new information.

To clarify the differences between the two samples, an illustration is provided. The left column of Table 1 contains a short sequence of prices that appeared in the CME's time and sales data for the September 1982 futures contract on April 23, 1982. In the center column are the price changes of Sample 1 computed from the time and sales data. Note that all prices are treated as if they are actual transactions. If n records are reported for the futures contract on a particular day, $n - 1$ price changes are computed. Note also that zero price changes appear where a transaction price appears after a bid/ask price at the same level. In the right column are the price changes of Sample 2. The number of price changes for a given day depends on the number of recorded bid/ask prices. A greater number of bid/ask quotes as a proportion of total price records for the day results in fewer computed price changes. Table 1, for example, shows only six computed price changes in a sequence of twenty recorded prices. On certain days, the number of bid/ask quotes for a distant maturity contract is so large that no Sample 2 price changes appear.

Table 1 also serves to illustrate that only a single price is reported on each record in the time and sales data file. Simultaneous bid/ask spread quotes do not appear in the time and sales data as they do in many stock market transaction files. Bid/ask spread quotes would have provided us with an alternative, direct measure of trading costs. The absence of direct

[20] Also on the file are records for transactions that are canceled or corrected. The canceled records are eliminated because the transaction (price) was erroneously reported. The corrected (price) records are eliminated because the transactions may be out of chronological order.

TABLE 1

AN EXAMPLE OF TIME AND SALES DATA AND THEIR RELATION TO SAMPLE 1 AND SAMPLE 2 PRICE CHANGES

Time and Sales Data: Price*	Sample 1 Price Change	Sample 2 Price Change
119.90	. . .	. . .
120.00	.10	.10
119.95	−.05	−.05
120.00	.05	.05
120.05B	.05	. . .
120.10B	.05	. . .
120.15A	.05	. . .
120.15	.00	. . .
120.10A	−.05	. . .
120.05A	−.05	. . .
120.05	.00	. . .
120.00A	−.05	. . .
119.95	−.05	. . .
119.90	−.05	−.05
119.85	−.05	−.05
119.90B	.05	. . .
119.95	.05	. . .
120.00B	.05	. . .
120.00	.00	. . .
119.95	−.05	−.05
Mean value	.0026	−.0083
Mean absolute value	.0447	.0583

NOTE.—The data are for the September 1982 S&P 500 index futures contract on April 23, 1982.
* A denotes an ask quote; B denotes a bid quote.

spread quotes, however, is a primary motivation for using the bid/ask spread estimators described in Section II.

The effects of censoring the time and sales data to generate Sample 2 are perhaps best understood by comparing the frequency of price changes in the two samples. Table 2 contains a summary of the frequency of price changes by contract maturity and tick size for the time and sales data for the S&P 500 futures contracts during the overall sample period April 21, 1982–October 9, 1987. The tick size for the S&P 500 futures contract is .05, and price changes are categorized by actual and absolute price value. Note that the nearby futures contract has by far the greatest number of recorded price changes. The Sample 1 results indicate that there were more than 2.35 million price changes reported in the sample. Of these, the largest majority are one-tick price moves, however price moves of two ticks and larger are not uncommon. For more distant contracts, two-

TABLE 2

Summary of Actual and Absolute Price Change Distributions by Contract Maturity and Number of Ticks

Contract Month	Number of Price Changes							
	Total	Less than −2 ticks	−2 ticks	−1 tick	0 ticks	1 tick	2 ticks	More than 2 ticks
Sample 1:								
Actual price changes:								
1	2,350,030	1,404	28,427	1,129,993	28,604	1,131,158	29,108	1,336
2	910,771	3,072	38,562	365,415	93,863	368,526	38,404	2,929
3	127,460	1,366	39,425	13,234	17,827	14,803	39,706	1,099
4	19,231	247	7,203	741	2,328	961	7,496	255
5	544	9	218	14	55	18	218	12
6	302	23	140	10	26	9	90	4
Absolute price changes:								
1	2,350,030				28,604	2,261,151	57,535	2,740
2	910,771				93,863	733,941	76,966	6,001
3	127,460				17,827	28,037	79,131	2,465
4	19,231				2,328	1,702	14,699	502
5	544				55	32	436	21
6	302				26	19	230	27

Sample 2:								
Actual price changes:								
1	2,279,864	1,061	25,410	1,103,620	19,379	1,103,273	26,118	1,003
2	394,797	1,044	10,778	184,024	3,273	183,521	11,066	1,091
3	2,472	42	448	698	91	729	440	24
4	90	0	19	14	9	24	24	0
5	1	0	0	0	0	0	1	0
6	0	0	0	0	0	0	0	0
Absolute price changes:								
1	2,279,864				19,379	2,206,893	51,528	2,064
2	394,797				3,273	367,545	21,844	2,135
3	2,472				91	1,427	888	66
4	90				9	38	43	0
5	1				0	0	1	0
6	0				0	0	0	0

NOTE.—The data presented are for the S&P 500 index futures contract for all days in the sample period April 21, 1982–October 9, 1987. The tick size for the S&P 500 futures contract is .05.

tick price changes are more common than one-tick price changes. The Sample 1 results show this for the third through sixth contract maturities.

The Sample 2 results, when contrasted with the Sample 1 results, show that the data censoring reduces the number of price changes in the sample, with the relative size of reduction increasing by contract maturity. For example, the nearby contract had a total of 2.35 million price changes in Sample 1 and 2.28 million transactions in Sample 2. This represents about a 3 percent decrease. The second, third, and fourth contract maturities experienced relative decreases of 56.6, 98.1, and 99.5 percent. Moreover, the censoring virtually eliminates all price changes for the fifth and sixth contract maturities. All of these observations reflect the fact that trading volume decreases with the time to contract maturity. The nearby contract has the most active and continuous market, so few bid/ask quotes appear in the time and sales data. To the contrary, distant maturities have few trades, so recorded time and sales data consist largely of bid/ask quotes, reflecting market movements without accompanying trades. Since our interest is in estimating the effective bid/ask spread from the price movements attributable to the bid/ask spread and not to new information, most of the interpretations in this section will focus on Sample 2 results.

Another interesting aspect of Table 2 is that the price change distributions are almost perfectly symmetric for both Sample 1 and Sample 2. This aspect is particularly reassuring considering that the method-of-moments estimator presented in Section II and applied later in this section assumes that true price changes are normally distributed.

Besides the time and sales data for the S&P 500 index and MMI futures contracts, other data sources are used. Since trading volume has been shown to be an important determinant of the bid/ask spread in stock markets, daily trading volume information for each S&P 500 and MMI futures contract during the sample period was obtained from Tick Data, Inc. In the regressions that follow, the trading volume variable is expressed in millions of futures contracts.

The bid/ask spread has also been shown to be sensitive to risk, so a risk variable is estimated and used in the empirical analysis. Since the focus will be on explaining the variation in the effective daily bid/ask spread, risk must be measured on a daily basis. Two simple estimators of risk are the daily trading range[21] or the square of the daily trading range; however, these estimators incorporate little information and are therefore extremely noisy. Another possibility is the standard deviation of the price changes or returns of the futures across the transactions

[21] The daily trading range is the difference between the high and the low of the day.

recorded on the time and sales data; however, the volatility estimate obtained from transaction prices during the day is upward biased due to the influence of the bid/ask spread. Yet another possibility is the standard deviation estimate obtained as a by-product in the method-of-moments procedure described in Section II. Unfortunately, this standard deviation depends on the average time between prices, and the average time changes from contract to contract and from day to day. In addition, this estimate is spuriously correlated with the estimate of the implied spread by virtue of the simultaneous parameter estimation procedure.[22]

The risk variable that we use in our analysis is the implied volatility from S&P 500 futures option prices. The implied volatility estimate is computed using a nonlinear regression of observed call and put option transaction prices for the S&P 500 futures option contract on model prices. The options are constrained to have a maturity between 15 and 110 days. The futures price preceding the option transaction is used as an estimate of the true futures price. Where the futures option price is below the intrinsic value of the option, the transaction is eliminated from the sample. The model used to price the American-style futures options is the quadratic approximation of Barone-Adesi and Whaley.[23] The riskless rate of interest is the yield on the Treasury bill maturing just after the option expiration or the thirty-day Treasury bill, whichever has the longest maturity. The Treasury-bill discount rates from which the yields are computed are obtained from the *Wall Street Journal*. A single estimate of volatility is computed each day during the sample period. The S&P 500 futures options did not begin trading until March 1, 1983, so a shorter sample is used in the regression analysis that follows. The risk variable is an annualized rate-of-return standard deviation expressed in percentage form.

B. Trading Volume Patterns

Trading volume is used as an explanatory variable in the regression tests that follow. Before examining the results of those tests, however, it is interesting to note that there was a redistribution of trading volume by contract maturity after the implementation of the top-step rule. Table 3 contains a summary of the trading volume of the S&P 500 index futures for the first four contract maturities.[24] The first period from April 21,

[22] For a discussion of properties of the volatility estimated using the method of moments procedure, see Smith & Whaley, *supra* note 18.

[23] Giovanni Barone-Adesi & Robert E. Whaley, Efficient Analytic Approximation of American Option Values, 42 J. Fin. 301 (1987).

[24] The fifth and sixth contract maturities are dropped since they are not traded in all of the subperiods examined in the table.

TABLE 3

TOTAL TRADING VOLUME OF S&P 500 INDEX FUTURES CONTRACTS BY CONTRACT MATURITY

	NUMBER OF CONTRACTS TRADED (% of Total)			
CONTRACT MATURITY	April 21, 1982–June 19, 1987	June 24, 1985–October 9, 1985	June 23, 1986–October 9, 1986	June 22, 1987–October 9, 1987
Nearby	58,659,872	3,542,105	5,154,978	5,210,671
	(86.418)	(86.237)	(86.436)	(88.008)
2	9,132,846	561,275	804,010	699,427
	(13.455)	(13.665)	(13.481)	(11.813)
3	80,004	3,883	4,830	10,293
	(.118)	(.095)	(.081)	(.174)
Distant	6,483	156	93	262
	(.010)	(.004)	(.002)	(.004)
Total	67,879,205	4,107,419	5,963,911	5,920,653
	(100.000)	(100.000)	(100.000)	(100.000)

NOTE.—The data are provided for the overall pre-top-step-rule sample period April 21, 1982–June 19, 1987, and during the three subperiods (*a*) June 24–October 9, 1985, (*b*) June 23–October 9, 1986, and (*c*) June 22–October 9, 1987.

1982–June 19, 1987, is the period from when the S&P 500 futures began trading until when the top-step rule was first implemented. Note that the relative trading volumes of the four contract maturities are 86.41, 13.46, .12, and .01 percent, respectively.

The second and third periods represent control group trading patterns. In general, the pattern of trading volume shifts by time of year as the nearby contract matures and an additional contract is added. Since the seventy-eight-day, post-top-step-rule period represents a specific interval in 1987 (June 22–October 9) among contract cycles, it is important that we evaluate trading volume patterns against the same interval during other years. The control groups that we form are for the period June 24–October 9, 1985, where the relative trading volumes of the four contract maturities are 86.24, 13.66, .10, and .00 percent, respectively, and for the period June 23–October 9, 1986, where the relative trading volumes are 86.44, 13.48, .08, and .00 percent, respectively. Note that these trading patterns are not different from those of the overall pre-top-step-rule sample period.

Interestingly, in the seventy-eight-day period after the introduction of the top-step rule, the trading volume pattern is significantly different from either of the control groups.[25] In particular, the proportion of total trading

[25] χ^2 tests of the null hypotheses that 1985 and 1986 control group periods have the same percentages of trading volume by contract maturity as the 1987 treatment group period are rejected at the .0001 probability level.

volume accounted for by the nearby contract, 88.01 percent, is considerably higher than the 86.24 percent reported for 1985 and the 86.44 percent reported for 1986. A possible explanation for this result is that dual traders tend to make markets in more distant contract maturities. When the top-step rule restricted their trading activity, they either (*a*) became exclusive brokers in the nearby contract month where trading volume is highest or (*b*) left the S&P 500 futures pit because their ability to earn two sources of income, from brokerage and from scalping/speculation, had been impaired.

The trading volume tests show that the implementation of the top-step rule appears to have influenced trading activity in the S&P 500 futures pit. The important question to answer, however, is not whether *trading patterns* have changed but rather whether *trading costs* have changed.

C. *Empirical Properties of the S&P 500 Spread Estimates*

Before beginning the analysis of the changes in the effective bid/ask spread resulting from the introduction of the top-step rule, we examine the empirical properties of the mean absolute price change and method-of-moments estimators of the effective bid/ask spread. Each estimator is computed for each S&P 500 futures contract each day during the sample period[26] using the price changes in Sample 1 and Sample 2. Table 4 contains the mean and the standard deviation of the daily spread estimates across days from the S&P 500 futures contract inception on April 21, 1982, through one week prior to the October 1987 crash. The correlations between the daily estimates for the two spread estimation approaches are also reported.

Several interesting results emerge from Table 4. First, comparing the descriptive statistics for the s_B and s_C estimates shows that the mean of the estimates of s_C is consistently less than s_B. This result is anticipated since the mean absolute price change s_B contains not only the effective spread but also the variance of true price changes. Second, both the s_B and s_C estimates for the nearby S&P 500 futures contract are close to the minimum price movement (tick size). Considering that the nearby S&P 500 futures contract is widely perceived to be a single-tick market, finding evidence to the contrary would have been surprising. Third, the correlation between the estimates is extremely high, indicating that the proportion of the mean absolute price change attributable to the bid/ask spread is fairly stable from day to day, particularly when price changes are censored in the manner of Sample 2. Finally, the spreads tend to increase

[26] We implicitly assume that the variance of true changes and the effective spread are constant throughout the trading day.

TABLE 4

MEAN, STANDARD DEVIATION, AND CORRELATION OF ESTIMATES OF THE S&P 500 INDEX FUTURES CONTRACT EFFECTIVE SPREAD

Contract Maturity	No. of Observations	Spread Estimate Mean (Standard Deviation)		Correlation between s_B and s_C
		s_B*	s_C†	
Sample 1:				
Pooled	6,103	.0669 (.0198)	.0636 (.0198)	.809
Nearby	1,385	.0505 (.0035)	.0503 (.0028)	.478
2	1,385	.0474 (.0075)	.0459 (.0067)	.773
3	1,385	.0757 (.0134)	.0715 (.0167)	.753
Distant	1,378	.0850 (.0123)	.0796 (.0167)	.304
Sample 2:				
Pooled	3,205	.0556 (.0137)	.0546 (.0130)	.843
Nearby	1,385	.0509 (.0043)	.0508 (.0036)	.680
2	1,382	.0552 (.0104)	.0543 (.0096)	.728
3	430	.0715 (.0257)	.0681 (.0262)	.850
Distant	8	.0703 (.0188)	.0666 (.0252)	.928

NOTE.—Data are provided for all days in the sample period April 21, 1982–October 9, 1987.

* The estimator s_B is the mean absolute value of the price changes, $s_B = \overline{|\Delta P^0|} = 1/T \sum_{t=1}^{T} |\Delta P_t^0|$, where T is the length of the time series.

† The estimator s_C is determined simultaneously with σ^2 in the solution to the system of equations, $\overline{|\Delta P^0|} = \sqrt{2/\pi}\, \sigma e^{-s_C^2/2\sigma^2} - s_C + 2s_C N(s_C/\sigma)$ and $\overline{|\Delta P^0|^2} = \sigma^2 + s_C^2$, where $N(d)$ is the cumulative unit normal distribution function with upper integral limit d.

with contract maturity, reflecting the lower market liquidity in the distant contracts.

In the regression tests that follow, we rely primarily on the results using the method of moments estimate s_C as the measure of the effective bid/ask spread to draw our inferences regarding the effect of the implementation of the top-step rule. To safeguard against incorrect inferences, however, we also report the regression results using the mean absolute price change estimate s_B.

D. *Methodology*

The methodology employed in the analysis of the effect of the top-step rule involves two regression models. In the first, the estimated effective spread, SPRD_t, is regressed on a dummy variable, D_t, that is assigned a value zero in the period preceding June 22, 1987—the date on which the top-step rule was first used—and one in all subsequent days, that is,

$$\text{SPRD}_t = \alpha_0 + \alpha_1 D_t + \epsilon_t. \tag{7}$$

The coefficient α_1 is the key to our analysis. If α_1 is not different from zero, the null hypothesis that there is no difference between the effective bid/ask spreads in the dual trader/nondual trader markets cannot be refuted. If α_1 is significantly less than zero, the null hypothesis is refuted in favor of the alternative hypothesis that dual trader markets are more costly than nondual trader markets. If α_1 is significantly greater than zero, the null hypothesis is refuted in favor of the alternative hypothesis that dual trader markets are less costly.[27]

In the second testing procedure, the multiple regression model

$$\text{SPRD}_t = \alpha_0 + \alpha_1 D_t + \alpha_2 \text{RISK}_t + \alpha_3 \text{VOL}_t + \epsilon_t \tag{8}$$

is performed. This regression specification is based on the arguments presented in Section I; that is, the effective bid/ask spread should be an inverse function of the trading volume and a direct function of risk. No information is available on the number of market makers in the pit. The interpretation of the multiple regression results is the same as that of the simple regression results—the focus is on α_1. The additional variables are included to ensure that the dummy variable is not acting as a proxy for trading volume and/or risk rather than the dual/nondual trader periods.

In the regression results that follow, the sample period begins on March 1, 1983, because S&P 500 futures options did not trade until that date, so no implied volatility estimates (recall that the implied volatility from the S&P 500 futures option prices is used as a proxy for risk) could be estimated. In addition, the sample ends on October 9, 1987—one week prior to the October 1987 market crash. This reason for this is that the market crash may have affected (increased) the effective bid/ask spread in the S&P 500 futures pit due to factors other than those in equation (8). Many market makers suffered large, unexpected losses in the days surrounding the crash. In some cases, these losses may have precipitated increases in the degree of risk aversion of market makers and hence increases in the rate of compensation (spread) for bearing the risk of the inventory position that they assume in order to accommodate customers. The crash may even have precipitated the departure of market makers from the futures pit, causing a decrease in competition and an increase in spread.

[27] This test is tantamount to a t-test of the difference between the mean spreads in the pre- and post-top-step-rule periods.

E. S&P 500 Regression Results

Tables 5 and 6 contain the regression results for the S&P 500 futures contracts for Sample 1 and Sample 2, respectively. The pooled results indicate that the hypothesis that there is no difference in the effective spread between markets that permit dual trading and those that restrict dual trading is rejected. For example, the Sample 2 (the sample which eliminates the effects of bid and ask quotes from the data) results reported in Table 6 indicate that during the period immediately after the top-step rule was implemented, the mean implied spread is .0116 higher on average than it was during the period in which dual trading was permitted. For the same sample, the mean absolute price change is .0138 higher. Both of these differences are significant from a statistical standpoint and are very meaningful in an economic sense. The parameter estimates indicate that the effective bid/ask spread has increased by more than 20 percent as a result of the top-step rule, independent of which measure of the spread is used.

On a contract-by-contract basis, the Sample 2 results are very similar. The nearby contract, for example, has an increased implied spread of .0035, and this increase is significantly greater than zero, as is indicated by the t-ratio of 8.05. The increase in spread for the second nearby contract is .0148, and the increase in spread for the third nearby contract is .0397. Both of these increases are also significant in a statistical sense. The contract-by-contract results using the mean absolute price as the dependent variable show that the spread increases after the top-step rule was implemented are generally larger and more significant than those of the implied spread. Presumably, this result arises because the mean absolute price change is an inflated measure of the effective bid/ask spread.

The simple linear regression results for Sample 1 (reported in Table 5) are very similar to those of Sample 2. Both the implied spread and the mean absolute price change are significantly higher in the post-top-step-rule period in the pooled regressions as well as on contract-by-contract bases. Before attaching too much importance to these results, however, we must be reminded that the reported simple linear regression results may be driven by greater risk and/or lower trading volume during the post-top-step-rule period. To properly evaluate any change in the effective spread, we must account for the spread's determinants (see Section I).

Tables 5 and 6 also contain the Sample 1 and Sample 2 results of the multiple regressions where risk and trading volume are included as explanatory variables. In general, the coefficients of the risk and trading volume variables appear with the expected signs. The spread increases with risk and decreases with trading volume. The risk variable enters more strongly than does trading volume. The t-ratios for the risk coefficient are large positive values. The t-ratios for the trading volume variable are smaller in absolute magnitude, and occasionally the sign of the coefficient is positive.

Finally, it is important to note that the results reported in Tables 5 and 6 are based on a linear regression model. This facilitates the economic interpretations below, where aggregate increased trading costs are estimated. It is worth noting, however, that the model was also estimated in log-linear form, with no change in the statistical inferences.

For purposes of economic interpretation, we focus on the multiple regression results where the implied spread is used as the dependent variable and Sample 2 is used and attempt to measure aggregate costs. Overall, the pooled results indicate that the implied spread for the S&P 500 futures contract has increased significantly. The estimated increase is .0087, and its t-ratio is 9.23. To evaluate the economic importance of this result, consider that in the seventy-eight-day period following the implementation of the top-step rule, total trading volume for the three nearby futures was 5,920,391 contracts. Estimated increased trading costs are therefore $.0087 \times 5{,}920{,}391 \times 500$ or \$25,753,701. Alternatively, since the pooled regression equally weights each contract maturity on each day of the sample period, it may be more appropriate to use the contract-specific parameter estimates, .0027, .0104, and .0272, for the three nearby contracts, respectively, and the contract-specific trading volumes in the seventy-eight-day period following the implementation of the top-step rule, 5,210,671, 699,427, and 10,293, respectively. Under this approach, the estimated increase in trading costs for the seventy-eight-day period are

$$(.0027 \times 5{,}210{,}671 + .0104 \times 699{,}427 + .0272 \times 10{,}293) \times \$500 = \$10{,}811{,}411.$$

These aggregate cost estimates are intended to be illustrative only. Implicitly, we have assumed that all trading volume is customer related. This is not the case. Some of the trading volume is between traders, hence our estimate of aggregate trading costs is overstated. Customer volume is almost certainly greater than half of the total, so halving the cost estimates should provide conservative estimates of aggregate costs, and these estimates remain economically meaningful. In addition, we have also implicitly assumed that trading volume (and pattern) is as high in the post-top-step-rule period as it would have been without the implementation of the rule. To examine the implication of this assumption,

TABLE 5

REGRESSION ANALYSIS OF THE S&P 500 INDEX FUTURES CONTRACT EFFECTIVE DAILY SPREAD

$$\text{SPRD}_t = \alpha_0 + \alpha_1 D_t + \alpha_2 \text{RISK}_t + \alpha_3 \text{VOL}_t + \epsilon_t$$

CONTRACT MATURITY	NO. OF OBSERVATIONS	PARAMETER ESTIMATES							
		$\hat{\alpha}_0$	$t(\hat{\alpha}_0)$*	$\hat{\alpha}_1$	$t(\hat{\alpha}_1)$*	$\hat{\alpha}_2$	$t(\hat{\alpha}_2)$*	$\hat{\alpha}_3$	$t(\hat{\alpha}_3)$*
Sample 1: $\text{SPRD}_t \equiv \hat{s}_B$ = mean absolute price change on day t:									
Pooled	5,232	.0667	237.30	.0055	4.81				
Nearby	1,166	.0505	446.99	.0032	7.35				
2	1,166	.0465	210.14	.0135	15.72				
3	1,166	.0743	191.04	.0127	8.46				
Distant	1,166	.0842	228.80	.0042	2.93				
Pooled	5,232	.0505	40.83	.0033	3.18	.0014	17.10	−.3595	−37.79
Nearby	1,166	.0438	82.65	.0019	4.46	.0004	12.54	−.0013	−.28
2	1,166	.0258	29.18	.0093	13.04	.0013	23.22	.0656	6.83
3	1,166	.0362	23.48	.0055	4.40	.0025	24.48	−9.3587	−4.14
Distant	1,166	.0630	37.15	−.0004	−.30	.0014	12.84	−35.4470	−3.89

Sample 1: $SPRD_t \equiv \hat{s}_C$ = implied spread on day t:									
Pooled	5,232	.0633	224.29	.0064	5.52				
Nearby	1,166	.0501	562.47	.0032	9.41				
2	1,166	.0452	220.01	.0082	10.26				
3	1,166	.0702	144.71	.0152	8.13				
Distant	1,166	.0787	159.98	.0081	4.24				
Pooled	5,232	.0512	39.11	.0048	4.36	.0010	12.37	−.2937	−29.19
Nearby	1,166	.0453	106.99	.0023	6.72	.0003	11.38	−.0002	−.06
2	1,166	.0291	34.02	.0049	7.18	.0010	18.31	.0970	10.45
3	1,166	.0304	14.63	.0075	4.48	.0026	19.58	−6.5563	−2.15
Distant	1,166	.0671	27.75	.0056	2.91	.0008	4.91	−8.0607	−.62

NOTE.—Results are provided for the period March 1, 1983–October 9, 1987. The effective daily spread of the contract is regressed on a dummy variable whose value is zero in the period preceding the implementation of the top-step trading rule on June 22, 1987, and one in the period June 22, 1987–October 9, 1987, on the implied volatility from S&P 500 futures option prices, and on the trading volume of the futures contract. The implied volatility is computed each day during the sample period using a nonlinear regression of observed call and put option transaction prices for the S&P 500 futures option contract on model prices. The options were constrained to have a maturity between 15 and 110 days. The futures price preceding the option transaction was used as an estimate of the true futures price. Where the futures option price was below the intrinsic value of the option, the option transaction was eliminated from the estimation. The model used to price the American-style futures options is the quadratic approximation of Giovanni Barone-Adesi & Robert E. Whaley, Efficient Analytic Approximation of American Option Values, 42 J. Fin. 301 (1987).

* The t-ratio corresponds to a test of the null hypothesis that the coefficient α equals zero.

TABLE 6

REGRESSION ANALYSIS OF THE S&P 500 INDEX FUTURES CONTRACT EFFECTIVE DAILY SPREAD

$$SPRD_t = \alpha_0 + \alpha_1 D_t + \alpha_2 RISK_t + \alpha_3 VOL_t + \epsilon_t$$

		PARAMETER ESTIMATES							
CONTRACT MATURITY	NO. OF OBSERVATIONS	$\hat{\alpha}_0$	$t(\hat{\alpha}_0)$*	$\hat{\alpha}_1$	$t(\hat{\alpha}_1)$*	$\hat{\alpha}_2$	$t(\hat{\alpha}_2)$*	$\hat{\alpha}_3$	$t(\hat{\alpha}_3)$*
Sample 2: $SPRD_t \equiv \hat{s}_B$ = mean absolute price change on day t:†									
Pooled	2,707	.0549	203.06	.0138	12.95				
Nearby	1,166	.0507	364.26	.0034	6.34				
2	1,165	.0544	182.02	.0196	16.97				
3	370	.0691	52.85	.0424	6.96				
Pooled	2,707	.0366	40.00	.0105	10.62	.0015	19.18	−.1595	−20.31
Nearby	1,166	.0439	67.46	.0022	4.30	.0006	13.22	−.0342	−6.24
2	1,165	.0290	23.31	.0144	14.37	.0017	21.39	−.0935	−6.92
3	370	.0230	4.12	.0280	4.75	.0031	8.06	2.1646	.28

Sample 2: $SPRD_t \equiv \hat{s}_C$ = implied spread on day t:†									
Pooled	2,707	.0540	213.02	.0116	11.55				
Nearby	1,166	.0505	445.11	.0035	8.05				
2	1,165	.0537	188.66	.0148	13.45				
3	370	.0658	50.62	.0397	6.54				
Pooled	2,707	.0386	34.03	.0087	9.23	.0012	16.91	−.1356	−17.98
Nearby	1,166	.0454	86.44	.0027	6.47	.0005	13.57	−.0380	−8.60
2	1,165	.0323	26.12	.0104	10.45	.0014	18.12	−.0790	−5.88
3	370	.0262	4.59	.0272	4.53	.0027	6.80	2.1545	.27

NOTE.—Results are provided for the period March 1, 1983–October 9, 1987. The effective daily spread of the contract is regressed on a dummy variable whose value is zero in the period preceding the implementation of the top-step trading rule on June 22, 1987, and one in the period June 22, 1987–October 9, 1987, on the implied volatility from S&P 500 futures option prices, and on the trading volume of the futures contract. The implied volatility is computed each day during the sample period using a nonlinear regression of observed call and put option transaction prices for the S&P 500 futures option contract on model prices. The options were constrained to have a maturity between 15 and 110 days. The futures price preceding the option transaction was used as an estimate of the true futures price. Where the futures option price was below the intrinsic value of the option, the option transaction was eliminated from the estimation. The model used to price the American-style futures options is the quadratic approximation of Giovanni Barone-Adesi & Robert E. Whaley, Efficient Analytic Approximation of American Option Values, 42 J. Fin. 301 (1987).

* The t-ratio corresponds to a test of the null hypothesis that the coefficient α equals zero.

† No meaningful results are obtained for the distant contract since there are no days during the period June 22–October 9, 1987, in which the distant contract had Sample 2 price changes.

we recompute aggregate trading costs using the trading volume figures reported in Table 3 for the period June 23–October 22, 1986, 5,154,978, 804,010, and 4,830, and find that the aggregate cost is $11,205,760, slightly higher than when the 1987 trading volume figures are used.

Regardless of the cost estimate used, the implementation of the top-step rule appears to have increased significantly the effective bid/ask spread in the S&P 500 index futures pit. Throughout the S&P 500 test results, increased spreads in the period after the introduction of the top-step rule are observed.

F. MMI Control Group Tests

The multiple regression results of the last subsection document increased trading costs in the S&P 500 futures market after the introduction of the top-step rule, independent of sample construction and effective trading cost estimator. The robustness of the results, however, may be influenced by (*a*) the lack of a control group, (*b*) the length of the pre-top-step-rule estimation period, and/or (*c*) regression model misspecification. To guard against these possibilities, we perform control group tests using the MMI futures contract during the period January 2–October 9, 1987. The MMI futures is a close substitute for the S&P 500 futures. Although the MMI is more narrowly based with only twenty stocks, all of the stocks have large market capitalizations and comprise a significant proportion of the value-weighted S&P 500 index. The correlation between their weekly price movements during 1989, for example, was only slightly less than .95.[28] The only significant differences between the contracts are that the MMI contract is about half the size of the S&P 500 futures contract and that the MMI contract has a monthly expiration cycle while the S&P 500 contract expires quarterly. With respect to this latter consideration, our tests of relative spreads and trading volumes in this section compare the first three nearby (monthly) MMI contracts with the nearby (quarterly) S&P 500 contracts.

The regression analyses conducted in this section attempt to circumvent possible model misspecification by making the implicit assumption that the economic factors affecting bid/ask spread and trading volume are the same for the MMI futures contract as the S&P 500 futures contract. The tests are straightforward. First, we regress the ratio of the spread of the nearby S&P 500 futures to the average spread of the three nearby MMI futures on a dummy variable that is zero during the pre-top-

[28] See Hans R. Stoll & Robert E. Whaley, Futures and Options: Theory and Applications 106 (1993).

TABLE 7

RELATIVE DAILY IMPLIED BID/ASK SPREADS AND RELATIVE TRADING VOLUME REGRESSIONS

	No. of Observations	Parameter Estimates			
		$\hat{\alpha}_0$	$t(\hat{\alpha}_0)$*	$\hat{\alpha}_1$	$t(\hat{\alpha}_1)$*
Relative spread regressions: $\frac{SPRD_{S\&P\ 500,t}}{SPRD_{MMI,t}} = \alpha_0 + \alpha_1 D_t + \epsilon_t$					
Sample 1	196	.6744	17.59	.0610	1.00
Sample 2	196	.6701	17.29	.0801	1.30
Relative trading volume regression: $\frac{VOL_{S\&P\ 500,t}}{VOL_{MMI,t}} = \alpha_0 + \alpha_1 D_t + \epsilon_t$					
	196	6.9456	34.14	−1.0695	−3.32

NOTE.—Results are presented of the regression of the ratio of daily implied bid/ask spread (trading volume) of the nearby S&P 500 futures contract to the average daily implied bid/ask spread (total trading volume) of the three nearby MMI futures contracts on a dummy variable whose value is zero in the period preceding the introduction of the top-step trading rule, January 2–June 19, 1987, and one in the period after, June 22–October 9, 1987.

* The t-ratio corresponds to a test of the null hypothesis that the coefficient α equals zero.

step-rule period, January 2–June 22, 1987, and one during the post-top-step rule from June 22–October 9, 1987, that is,

$$\frac{SPRD_{S\&P\ 500,t}}{SPRD_{MMI,t}} = \alpha_0 + \alpha_1 D_t + \epsilon_t. \tag{9}$$

Next, we regress the ratio of the nearby S&P 500 futures contract trading volume to the trading volume of the three nearby MMI futures on the same dummy variable,

$$\frac{VOL_{S\&P\ 500,t}}{VOL_{MMI,t}} = \alpha_0 + \alpha_1 D_t + \epsilon_t. \tag{10}$$

The coefficient α_1 in the two regressions is an indicator of whether or not the bid/ask spread (trading volume) of the S&P 500 contract changed relative to the MMI contract.

Table 7 contains the results of the bid/ask spread regressions. The estimates of the intercept term, $\hat{\alpha}_0$, indicate that prior to the introduction of the top-step rule, the spread in the S&P 500 futures market was about 67 percent of the average size of the spread of the three nearby MMI futures contracts. The relative spread increased after the introduction of the rule. The slope coefficient estimates are positive in both the Sample

1 and Sample 2 regressions. The Sample 2 results indicate that the nearby S&P 500 futures spread increased to a level of about 75 percent of the average MMI futures spread, an increase of about 12 percent. Although this increase is only marginally significant, it is considerably larger than the Table 6 result, where the increase in the bid/ask spread of the nearby S&P 500 futures contract was reported to be .0027, an increase of about 5.4 percent (assuming the pre-top-rule spread is .05).

Table 7 also contains the relative trading volume regression results. Prior to the top-step rule, the trading volume of the nearby S&P 500 futures appears to be about 6.94 times higher than the total trading volume of the three nearby MMI contracts. After the rule, the factor drops significantly to about 5.88. The implementation of the rule apparently caused the trading volume in the nearby S&P 500 futures to drop relative to the total trading volume of the three nearby MMI futures contracts.

IV. Summary and Conclusions

The purpose of this study is to analyze the merits of dual trading. Specifically, we consider the cost of implementing the CME's top-step rule in the S&P 500 index futures pit in June 1987. Under this rule, only exclusive brokers are permitted access to the top step of the S&P 500 index futures pit. An effect of this rule has been to substantially curtail the volume of trading executed by dual traders. The empirical analysis conducted in this study indicates that the effective bid/ask spread increased and trading volume decreased as a result of restricting dual trading. During the period immediately following the implementation of the rule, June 22–October 9, 1987, the estimated increase in trading costs is at least $5 million.

If restricting dual trading is so costly, why then did the CME choose to implement the top-step rule and then subsequently to prohibit dual trading in all of its futures markets that trade an average of at least 10,000 contracts a day? One plausible explanation is that the CME continues to be threatened by more intrusive regulation by the CFTC. By self-imposing restrictions on highly active contracts, the CME is voluntarily incurring modest costs in highly active contract markets in an attempt to appease the CFTC and avoid the much higher expected costs of a dual trading ban.

Bibliography

Barone-Adesi, Giovanni, and Whaley, Robert E. "Efficient Analytic Approximation of American Option Values." *Journal of Finance* 42 (1987): 301–20.

Benston, George, and Hagerman, Robert. "Determinants of Bid/Asked Spreads in the Over-the-Counter Market." *Journal of Financial Economics* 1 (1974): 353–64.

Chicago Mercantile Exchange. Summary of S&P 500 top-step analysis, 1990. Excerpts from a letter to the Commodity Futures Trading Commission.

Commodity Futures Trading Commission. "Economic Analysis of Dual Trading on Commodity Exchanges." Unpublished report prepared by the Division of Economic Analysis with technical support from the Office of Information Resources Management, November 1989.

Commodity Futures Trading Commission. "Proposed Rule concerning Restriction on Dual Trading by Floor Brokers, Trading on Commodity Exchanges." Unpublished staff document prepared by the Division of Trading and Markets, January 3, 1990.

Demsetz, Harold. "The Cost of Transacting." *Quarterly Journal of Economics* 82 (1968): 33–53.

Fishman, Michael J., and Longstaff, Francis A. "Dual Trading in Futures Markets." *Journal of Finance* 47 (1992): 643–71.

Grossman, Sanford J. "An Economic Analysis of Dual Trading." Working paper. Philadelphia: University of Pennsylvania, Wharton School, Rodney L. White Center for Financial Research, 1989.

Grossman, Sanford J., and Miller, Merton H. "Economic Costs and Benefits of the Proposed One-Minute Time Bracketing Regulations." *Journal of Futures Markets* 6 (1986): 141–66.

Grossman, Sanford J., and Miller, Merton H. "Liquidity and Market Structure." *Journal of Finance* 43 (1988): 617–37.

Roll, Richard. "A Simple Implicit Measure of the Bid/Ask Spread in an Efficient Market." *Journal of Finance* 39 (1984): 1127–39.

Silber, William L. "Marketmaker Behavior in an Auction Market: An Analysis of Scalpers in Futures Markets." *Journal of Finance* 39 (1984): 937–53.

Smith, Tom, and Whaley, Robert E. "Estimating the Effective Bid/Ask Spread Using Time and Sales Data." *Journal of Futures Markets* (June 1994, in press).

Stoll, Hans R. "Dealer Inventory Behavior: An Empirical Investigation of NASDAQ Stocks." *Journal of Financial and Quantitative Analysis* 11 (1976): 359–80.

Stoll, Hans R. "The Supply of Dealer Services in Securities Markets." *Journal of Finance* 33 (1978): 1133–51.

Stoll, Hans R. "The Supply of Dealer Services: An Empirical Analysis of NASDAQ Stocks." *Journal of Finance* 33 (1978): 1153–72.

Stoll, Hans R. "Inferring the Components of the Bid/Ask Spread: Theory and Empirical Tests." *Journal of Finance* 44 (1989): 115–34.

Stoll, Hans R., and Whaley, Robert E. "The Dynamics of Stock Index and Stock Index Futures Returns." *Journal of Financial and Quantitative Analysis* 25 (1990): 441–68.

Stoll, Hans R., and Whaley, Robert E. *Futures and Options: Theory and Applications*. Cincinnati, Ohio: South-Western Publishing Company, 1993.

Thompson, Sarahelen R., and Waller, Mark L. "Determinants of Liquidity Costs in Commodity Futures Markets." *Review of Futures Markets* 7 (1988): 110–26.

Tinic, Seha. "The Economics of Liquidity Services." *Quarterly Journal of Economics* 86 (1972): 79–93.

Whaley, Robert E. "On Valuing American Futures Options." *Financial Analysts Journal* 42 (1986): 49–59.

Whaley, Robert E. "Valuation of American Futures Options: Theory and Empirical Tests." *Journal of Finance* 41 (1986): 127–50.