Dual Trading and Futures Market Liquidity: An Analysis of Three Chicago Board of Trade Contract Markets

Michael J. Walsh
Stephen J. Dinehart

INTRODUCTION

This article examines the relationship between dual trading activity and market liquidity in Chicago Board of Trade (CBOT) wheat, soybean, and major market index (MMI) futures contract markets. Dual trading allows exchange members to trade for their own account and to execute customer orders during the same day. Supporters of the practice claim dual trading enhances market liquidity and lowers the cost of using futures markets. It is argued also that the presence of dual traders provides more competition for brokerage customers, and thus reduces brokerage commissions. Critics of dual trading claim the practice makes it easier for brokers to illicitly profit by trading ahead of their customers' orders, and presents an inherent conflict of interest.

The liquidity argument is examined by testing for a statistical relationship between dual trader activity and market liquidity as measured by an estimate of the bid–ask spread. The model assumes that own-account traders, including dual traders when they trade for their own account, are liquidity suppliers. The sensitivity of market liquidity to variation in the number and composition of own-account traders is estimated. The hypothesis that market liquidity increases when dual traders comprise a larger share of all liquidity suppliers is tested.

Previous research that directly examines the effects of dual trading includes Stanley (1981), the Commodity Futures Trading Commission (CFTC) (1989), Smith and Whaley (1990), and Grossman (1990). Stanley's simulations indicate that, depending on the assumed cost of an exchange seat and the volume each trader can

The views expressed in this paper are those of the authors only and do not necessarily reflect those of the Chicago Board of Trade or any of its members.

The authors thank Frank Rose, Francis Longstaff, Hang Chang, Betsy Kuhn, Robert McCarthy, and two anonymous referees for helpful comments. Any errors are the authors'.

Michael J. Walsh is an Advisory Economist at the Chicago Board of Trade.

Stephen J. Dinehart is Senior Economist at the Chicago Board of Trade.

The Journal of Futures Markets, Vol. 11, No. 5, 519–537 (1991)

CCC 0270-7314/91/050519-19$04.00

execute per day, a ban on dual trading would result in some combination of increased total brokerage fees charged to customers and slower execution of customer orders. Delayed execution is generally undesirable because it can imply inferior execution prices for customers.

The CFTC study examines dual traders' contribution to liquidity by estimating the relative tendency of dual and nondual traders to take the opposite side of customer orders, and by regressing a measure of the bid–ask spread on the contract volume executed by dual and nondual traders. The CFTC results for the entire industry show that the percent of own-account volume of dual and nondual traders that is opposite of customers' trades is not significantly different, but this result varies among exchanges.

The econometric analysis in the CFTC study regresses a measure of bid–ask spreads—the log of the average absolute value of price changes during a half-hour divided by the price level—on volume executed by dual and nondual traders (using total volume and own-account volume in different estimates) and a measure of price variation. They generally find no significant difference between the effects of dual and nondual volume on their liquidity measure.

The presence of several fundamental problems in the econometric section of the CFTC study suggests the results are not reliable. The most important problem involves the dependent variable used in the regression analyses, i.e., the average absolute value of price changes divided by the price level. In almost all observations in the CBOT T-bond and Chicago Mercantile Exchange (CME) S&P 500 futures markets, the numerator of this variable is equal to one tick, i.e., the smallest possible price change. For these observations, the dependent variable is, in effect, a constant divided by the price level. The price level adjustment is included to adjust for difference across contracts, but independent variation in prices induces variation in the dependent variable that is unrelated to variation in liquidity. Because these two markets should be given the most weight in the CFTC regressions (they are the highest volume contracts included), the results of their regressions appear to be invalid.

Other important problems in the CFTC study (which are discussed further below) include: combining different markets into a single analysis (thereby forcing the relationship between trading activity and liquidity to be the same in all markets), a strong correlation between own-account volume of dual and non-dual traders (which makes it difficult to separate the effects of each volume type), and the bias resulting because bid–ask spreads and own-account activity of dual and nondual traders are simultaneously determined.

Smith and Whaley examine the behavior of several measures of the bid–ask spread before and after the CME restricted dual trading in the S&P 500 futures market in June 1987. They find that spreads increase by a minimum of 5% in the nearby contract, and by 19% and 40% in first and second deferred contracts after the restrictions were adopted. They conclude that bid–ask spreads rose because the restrictions on dual trading reduced competition among liquidity traders. The approach used by Smith and Whaley attributes all changes in bid–ask spreads to the dual trading restrictions when, in fact, other excluded factors may have caused the changes. Specifically, S&P 500 price volatility was increasing even before the October 1987 stock market crash, and market conditions changed fundamentally after the crash; these factors or others may have been the true causes of increased bid–ask spreads observed after June 1987.

Grossman compares the cost of customer order execution by dual traders with that of nondual traders in the CBOT wheat and corn futures markets during June and July 1989. His measure of execution cost compares prices actually received by customers to market prices occurring at the time the customer submitted the order, and to prices occurring after the order was executed. He concludes that the cost of orders executed by dual traders is, on average, roughly 12% less than the cost of orders executed by pure brokers. Grossman offers no explanation for the hypothesis that dual traders might be superior in executing customer orders than nondual traders. Like Smith and Whaley, Grossman also excludes other variables that affect execution cost, such as price volatility, time of day, and transactions volume. Including these variables might alter the results.

In related research, Thompson and Waller (1988) examine the contribution of scalpers (referred to here as locals, liquidity suppliers, or market makers) to liquidity in the CBOT corn and oats contract markets. They find that market liquidity increases as scalpers' share of all executed volume increases. Their findings suggest that if scalpers' share of all volume is larger when dual trading is allowed (the total number of scalpers is presumably higher when dual trading is allowed), prohibition of the practice would lead to reduced liquidity.

The analysis in this study attempts to improve on previous research by examining a broader cross-section of markets, including more factors that may explain market liquidity, and by addressing some of the econometric problems encountered in previous studies. Unlike most previous studies, the data allow the examination of how intra-day variation in dual trading and other factors affect market liquidity.

Why might the liquidity effect of own-account trading by dual traders differ from that of exclusive own-account traders? One plausible explanation mirrors Stanley's arguments. Examining dual trading from the perspective of brokerage, Stanley argues that dual trading permits lower brokerage costs due to the supplemental income provided from speculative trading. How the multi-product nature of a dual trading "firm" can lower its marginal cost of own-account trading is discussed below. This structural advantage is complemented by dual traders' exposure to information from their brokerage activity, which may reduce the asymmetric information effect that contributes to bid–ask spreads.

MEASUREMENT OF MARKET LIQUIDITY

One measure of futures market liquidity is the bid–ask spread. This is the difference between the "ask" price sought by market makers when selling contracts and the (lower) price they bid when trying to purchase contracts. The bid–ask spread represents the return when an own-account trader buys at the bid and immediately sells at the (higher) offer or "ask" price. Stoll (1989) summarizes the current literature on bid–ask spreads by describing the three components thought to explain the size of the spread: (i) order processing costs; (ii) inventory holding costs due to the risk of unfavorable price moves while holding the position; and (iii) adverse selection or asymmetric information, which occurs if the customer placing the trade has superior information that implies that it is advantageous for the customer to take a position against which the market maker places himself. The possibility that order processing costs, inventory holding costs, and the effect of asymmetric information could be smaller for dual trading market makers than for nondual trading market makers is discussed below. These differences could explain the presence of a relationship between market-making activity of dual traders and bid–ask spreads.

Highly liquid markets are often characterized by a one-tick spread, in which case the difference between bid and ask prices is as small as possible given contract specifications (0.25¢ per bushel, or $12.50 per contract in CBOT grain markets; 0.05 index points, or $12.50 per contract for the Major Market Index contract).

Bid and ask prices are generally not recorded; so, a proxy, the average absolute value of transaction price changes during a half-hour time bracket, is used instead. (The measure does not include zeroes, i.e., transactions executed at the last transaction price.) This proxy is an accurate estimate of bid–asks spreads to the extent that transaction prices move from bid to ask and from ask to bid.[1]

Price changes are smaller in more liquid markets where, by definition, orders can be executed with less price movement. Conversely, less liquid markets are characterized by larger changes in price for a given order size. Order execution in a less liquid market is more costly to the customer because the price displacement needed to have the order filled is larger relative to a more liquid market. It is assumed that the effect of new information on price changes is random and does not distort the measure of market liquidity in any systematic way.

METHODOLOGICAL ISSUES, ECONOMETRIC PROBLEMS

Tests are conducted to measure the relationship, if any, between intraday variation in dual trading activity and corresponding changes in market liquidity. These tests, as well as those used in related research cited above, are used to indirectly address the real question, which asks if market liquidity would be reduced by a prohibition of dual trading.

It should be noted that these tests do not examine evidence regarding other benefits and costs sometimes attributed to dual trading. The purported benefits include superior prices received by customers of dual traders versus pure brokers, enhancement of market depth, order execution speed, order execution capacity, and lower brokerage costs resulting from extra competition for brokerage business. The costs include the possibility of customers receiving inferior prices due to the broker trading ahead of the customer order and easier cheating through use of one's personal account. Effects on market liquidity may be the most important of all the benefits attributed to dual trading because increased liquidity benefits all market participants.

There are several econometric problems that must be addressed when assessing the effects of trading practices on market liquidity. The most prominent problems encountered in this analysis, as well as those of the CFTC and Thompson and Waller (1988), are the following:

1. Simultaneous determination of liquidity and trading activity of dual and nondual traders.
2. Strong multicollinearity; own-account trading activity of dual and nondual traders is strongly correlated.

[1]In their analysis of bid–ask spreads in the CME S&P 500 futures market, Smith and Whaley conclude that the average absolute value of price changes is a biased measure of bid–ask spreads. However, they also find that this measure is strongly correlated with their corrected measure, and that the two measures generate nearly identical regression results. Thompson and Waller use four different measures of liquidity and closely examine two: the average absolute value of price changes, and a transformed (two times the square root of the negative of) first-order serial covariance of price changes. They conclude that the average absolute value of price changes is preferable to the serial covariance measure. The latter is, in many cases, positive (thus, preventing computation of the transformed variable), and regressions on it often yield coefficients with signs that are counterintuitive and opposite those found when other measures of liquidity are used as the dependent variable.

3. Nonnormally distributed or truncated dependent variables; in many markets, the average absolute value of price changes is clustered near one tick.
4. Use of an aggregated dependent variable; Maddalla (1983) notes that OLS, on a dependent variable comprised of truncated (e.g., at one tick) individual observations, can be inappropriate.

These problems are explained further in the discussion of the model. The approach used here tries to avoid the simultaneity and multicollinearity problems. The non-normality and truncation problems (which arise in only one market segment examined) and the aggregation problem are not addressed because doing so requires the use of complicated estimators that would tax the limited sample size and add little to the understanding of the relationships of interest.

The data used in some other studies are pooled samples of observations on several commodities (each having more than one contract month trading) followed over time. When combining observations on more than one commodity or more than one contract market of a single commodity, one implicitly assumes that regression coefficients are the same in all the contracts included in the sample. This assumption, while not acknowledged in previous research, may be unrealistic. It forces the estimated relationship between trading activity and liquidity to be the same across commodities and among different contracts on the same commodity. This study's model is estimated by commodity and contract month groups, and dummy variables are used to allow intercepts to vary within contract month groups.

A SIMPLE MODEL OF MARKET LIQUIDITY

It is assumed that the level of market liquidity during a given time frame depends on the number of active liquidity suppliers, that is, the number of locals and dual traders trading for their own account. It is assumed also that the number of liquidity-supplying traders willing to trade during a given time period depends on the level of liquidity. Some own-account traders are averse to trading when price changes are large, while others are more inclined to trade in those periods. Simultaneous determination of market liquidity and the number of liquidity suppliers means that single equation estimates of the effect of the number of own-account traders on liquidity would be biased and inconsistent. Two-stage least squares is used to correct for this problem.

As noted above, the number of dual and nondual traders that actively trade for their own account in each time period is strongly correlated.[2] This makes it difficult to separate the different effects on market liquidity that the two types of traders might have. To deal with this problem, the number of dual and nondual own-account traders active in each time period are combined, and then an examination is made to determine whether market liquidity is affected by not only the total number of own account traders but also its composition. The composition

[2]The correlations between *DT* and *NT* for the various samples are as follows:

	Full Sample	Nearby Contracts	Deferred Contracts
Wheat	0.81	0.65 (2 contracts)	0.04 (2 contracts)
Soybeans	0.96	0.94 (2 contracts)	0.77 (4 contracts)
MMI	0.79	0.70 (2 contracts)	0.29 (1 contract)

Similarly, strong correlations are found between dual own-account volume and non-dual own-account volume.

variable is the number of dual traders actively trading for their own account divided by the total number of own-account traders in each time period. A finding that the latter variable affects market liquidity would be evidence in support of the hypothesis that own-account trading of dual traders has a different effect on market liquidity than does own-account trading of nondual traders.

Expressions (1) through (3) summarize the model. Expression (1) states that liquidity (L) is a function of the number of active liquidity suppliers: dual traders (DT) plus nondual traders (NT) trading for their own account; dual traders as a proportion of all own-account traders ($DT/(DT + NT)$); other factors included in the vector X; and an error term. The number of active liquidity suppliers affects bid–ask spreads not only by providing actual competition through their trading, but also because their presence in the pit (indicated by doing at least one trade during the half-hour) provides additional potential competition among buyers and sellers.

Expressions (2) and (3) state that the number and composition of dual own-account plus nondual own-account traders active in each time period depends on the level of market liquidity and other factors included in the vector Z and an error term. Function "g" is assumed to differ from function "h."

$$L = f((DT + NT), DT/(DT + NT), X) + e_1 \tag{1}$$

$$DT + NT = g(L, Z) + e_2 \tag{2}$$

$$DT/(DT + NT) = h(L, Z) + e_3 \tag{3}$$

where

L = liquidity during a given time period,
DT = number of dual traders who trade for their own account during a given time period,
NT = number of nondual traders who trade for their own account during a given time period,
X, Z = vectors comprised on other relevant variables, and
e_i = error terms.

OLS regression of L on $DT + NT$, $DT/(DT + NT)$, and X is inappropriate because $DT + NT$ and $DT/(DT + NT)$ are correlated with e_1. This problem forces the use of a systems estimation technique, such as two-stage least squares, to generate instruments for $DT + NT$ and $DT/(DT + NT)$ that are uncorrelated with e_i.

In the first stage regressions, all the predetermined variables in the system of equations (those contained in vectors X and Z) are used to generate fitted values of $DT + NT$ and $DT/(DT + NT)$. The fitted values are included on the right-hand side of the second stage regression. An appropriate number of predetermined variables are omitted in the second stage to conform to the order condition for identification of that equation.

It is expected that market liquidity increases (average bid–ask spreads are smaller) as the total number of own-account traders ($DT + NT$) increases. The coefficient on $DT/(DT + NT)$ is examined to test the hypothesis that the composition of the pool of own-account traders is important, i.e., that dual own-account traders have a larger effect on market liquidity than do nondual own-account traders.

Included in the "X" vector from expression (1): price volatility in the previous half-hour period; the average number of contracts traded per transaction during the half-hour; customers' share of all volume; a scale variable that increases with time until contract expiration (values are 0 and 1 for each subsample having the two nearby wheat, soybean and MMI contracts, and for the two deferred wheat contracts subsample, and 0 through 3 for the four deferred soybean contracts subsample); and a dummy variable indicating the contract is a new crop year contract (July 1989 expiration for wheat, November 1989 for soybeans).

Bid–ask spreads are expected to be larger when prices are more volatile because marketmakers must compensate for the larger inventory risk they face when holding positions. Larger transactions are expected to increase the average bid–ask spread because larger trades require greater price displacement to be executed. This may reflect a perception of adverse information possessed by the party who is counter to the market maker, or it may reflect liquidity supply and demand conditions. It is assumed that market makers perceive there to be more adverse information when volume executed for customers comprises a larger share of all volume; therefore, bid–ask spreads should be positively associated with this variable. (It is assumed that increased customer volume also implies an increase in demand for market liquidity.) Bid–ask spreads are expected to be larger the longer the time until a contract expires because those contract markets are less active, thus it is riskier to hold positions in those markets. It is expected that bid–ask spreads in the new crop year contract market are smaller than those in other deferred contracts because there is more active trading.

The vector "Z" included in expressions (2) and (3) contains variables thought to explain variation in the number and composition of dual plus nondual own-account traders active in each period. The number of active own-account traders is assumed to be positively related to the level customer (brokerage) volume executed in the previous time period as traders respond to profit opportunities that such volume presents. An increase in brokerage volume in one time period signals to traders in other pits that a market is active. To the extent that traders can profitably trade in more than one pit (and their membership status allows trading in other markets), some may move from less active pits to the markets that are more active. Other sources of additional own-account traders include: those who are in the pit, but inactive; traders active in one expiration, and see that the other expirations are active; and spread traders who see opportunity in activity.

The number of traders willing to do own-account trades is expected to be nonlinearly related to price volatility. To allow for this, price variance in the previous half-hour period is included, along with its square as explanatory variables. A nonlinear relationship between volatility and the number of active traders is allowed because it is assumed that slightly higher volatility attracts some traders and discourages others; but, this effect changes as volatility increases further. Because current-period volatility (price variance) is endogenous, i.e., determined simultaneously with *DT* plus *NT*, the price variance from the previous time bracket is used. Finally, the number of own-account traders is expected to be lower the longer the time until contract expiration except when the deferred contract is a new crop year contract. A dummy variable is used to indicate the latter. Because it is expected that the number of own-account traders is greater in the first two and final half-hour time brackets of each day—those offering more profit opportunities—dummy variables are included to denote those brackets. Two-stage least squares requires inclusion of the other predetermined variables in the first stage regressions (as a simplification, it is assumed that average trade size and customers' share of all volume are exogenous).

The first stage regressions [expressions (4) and (5)] use all the predetermined variables to generate predicted levels of $DT + NT$ and $DT/(DT + NT)$.

$$DT_t + NT_t = F[\text{price variance}_{t-1}, (\text{price variance}_{t-1})^2, \text{customer volume}_{t-1}, \text{time until contract expiration, new crop year contract dummy, open and close bracket dummies, average trade size}_t, \text{customer's share of all volume}_t] \quad (4)$$

$$DT_t/(DT_t + NT_t) = G[\text{price variance}_{t-1}, (\text{price variance}_{t-1})^2, \text{customer volume}_{t-1}, \text{time until contract expiration, new crop year contract dummy, open and close bracket dummies, average trade size}_t, \text{customer's share of all volume}_t] \quad (5)$$

In the second stage, regression [expression (6)] the estimated bid–ask spread is the dependent variable and the independent variables are fitted values of $DT + NT$, and $DT/(DT + NT)$, as well as other relevant variables.

$$\text{Bid–ask spread}_t = H[DT + NT, DT/(DT + NT), \text{price variance}_{t-1}, \text{average trade size}_t, \text{customers' share of all volume}_t, \text{time until contract expiration, new crop year dummy}] \quad (6)$$

Rather than assume a particular functional form, linear specifications are used in both stages as a first approximation of what may be a more complicated relationship.

DATA SUMMARY AND ANALYSIS

The model is estimated using half-hour data on activity in the CBOT wheat, soybean, and MMI futures contract markets for three weeks in the fourth quarter of 1988. The estimated bid–ask spread is based on price data from the CBOT time and sales register, and the other variables are based on CBOT clearing corporation data. These data are chosen as the low, high, and average volume weeks of the quarter.

The same data are gathered for corn, T-bonds, and ten-year note markets. Because instances where the estimated bid–ask spread exceeds one tick are infrequent (fewer than 10% of cases) in these markets, the model is not useful for explaining the few instances of reduced liquidity. Those results are not reported here. Research on futures market microstructure cannot assess effects on bid–ask spreads in markets where one-tick spreads are dominant due to the lack of variation in the dependent variable. Therefore, while it is possible that dual trading helps keep bid–ask spreads small in one-tick markets, the only observable marginal effects of dual trading in such markets are better customer fills, enhanced market depth, immediacy, or from information effects.

For purposes of this study, a trader is designated a dual trader if, within the same trading day, he executes trades for both his own account and for a customer's account. This segmentation of the trading population into dual and nondual traders is done using proprietary records of the Board of Trade Clearing Corporation (BOTCC) and the computerized trade reconstruction system of the CBOT. The definition of dual trader presence during a particular half-hour bracket has some limitations. First, it is assumed that if a trader executes both customer and own-account trades at any point in the day, that trader could potentially do either at any time during that day. Second, it is assumed that a trader is present during the entire half-hour bracket if that trader executes a trade at any time during the bracket. Third, it is assumed that mere dual-trader presence (as indicated by doing at least one trade during a bracket), not necessarily dual trading activity, is sufficient to affect the bid–ask spread. The last point means it is assumed that potential dual-trading activity is sufficient to affect the bid–ask spread.

Table I shows summary statistics for the key variables used in the model. The data are reported for two subsamples—the nearby and first deferred month, and all other deferred month activity. The full sample is divided because the average contract volume per half-hour in the nearby and first deferred contracts (December 1988 and March 1989 for wheat, November 1988 and January 1989 for soybeans, and October and November 1988 for MMI) is much higher than in deferred expiration contracts. Because the volume levels, number of traders, and estimated bid–ask spreads differ significantly in the two subsamples, it is assumed that the relationship examined differs across the subsamples, so the subsamples are kept separate.

Note that the minimum average absolute value of price change is equal to one tick (0.25¢), in some cases, in the wheat and soybean markets. As expected, this estimate of the bid–ask spread is, on average, larger in the deferred contracts. Figure 1 shows the distribution of the average of absolute value of price changes for nearby and deferred samples. Table I also shows that dual traders comprise around 40% of all own-account traders in the wheat market, around 20% of all own-account traders in the soybean market, and less than 10% of all own-account traders in the MMI market.

Table II shows the results from the second stage OLS regression of expression (6) for the two nearby contracts and for deferred contracts.[3] The equations used to generate the fitted values of $DT + NT$ and $DT/(DT + NT)$ for each of the samples are shown in the Appendix. The coefficient standard errors reported in Table II are the original standard errors reported by the regression package multiplied by an adjustment factor required for the second stage of the 2SLS procedure.[4] In all cases, this adjustment reduces the statistical significance of the reported coefficients.

The coefficient on the $DT + NT$ variable is negative in all the nearby samples but is statistically significant (at the 90% level of confidence) only for nearby soybeans. Thus, for nearby soybeans, the evidence suggests that estimated average bid–ask spreads are smaller when more liquidity "suppliers" (own-account traders) are active.

The coefficient on $DT/(DT + NT)$, the ratio of dual traders who do own-account trades to all own-account traders, is negative in all samples. That coefficient is statistically significant at the 95% level of confidence for the deferred wheat and nearby MMI samples and it is significant at the 99% level of confidence in the nearby expirations of soybeans.

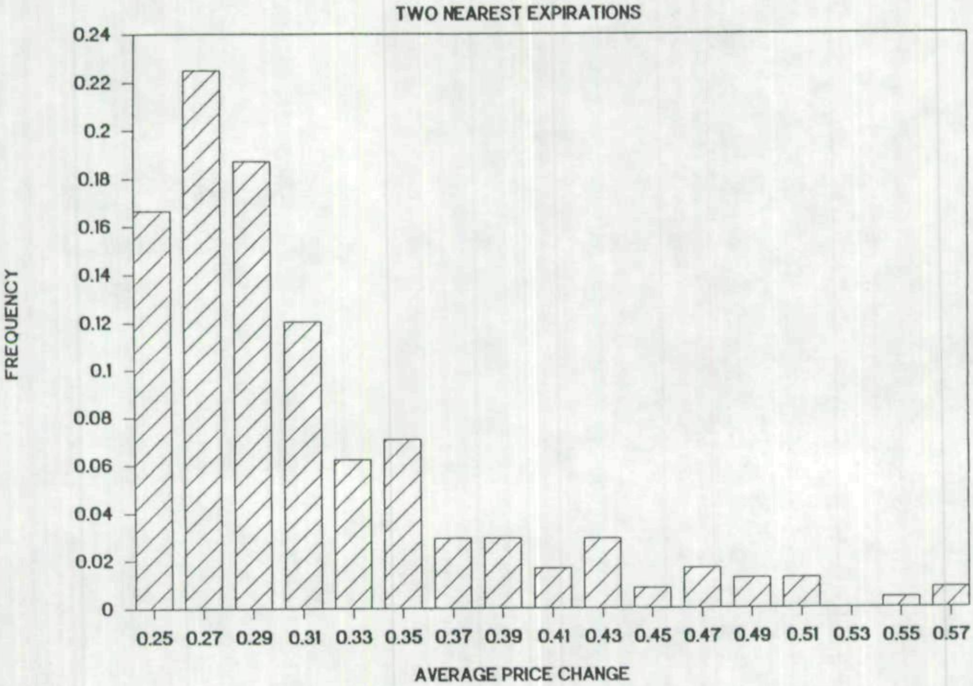
This evidence suggests that, for some samples of wheat, soybeans, and MMI contracts, estimated bid–ask spreads are smaller when dual traders trading for their own account comprise a larger share of all own-account traders. In these cases the marginal contribution of each additional own-account trader to lower bid–ask spreads is larger when that own-account trader is a dual trader. The total marginal

[3]Because many (16%) of the observations in the two nearby wheat contracts are equal to the truncation value (one tick), the Tobit estimator is applied also to the second stage for that sample. Tobit estimation is appropriate when the dependent variable takes values equal to and above (or equal to and below) a limit value. The results are not qualitatively different from those reported here.

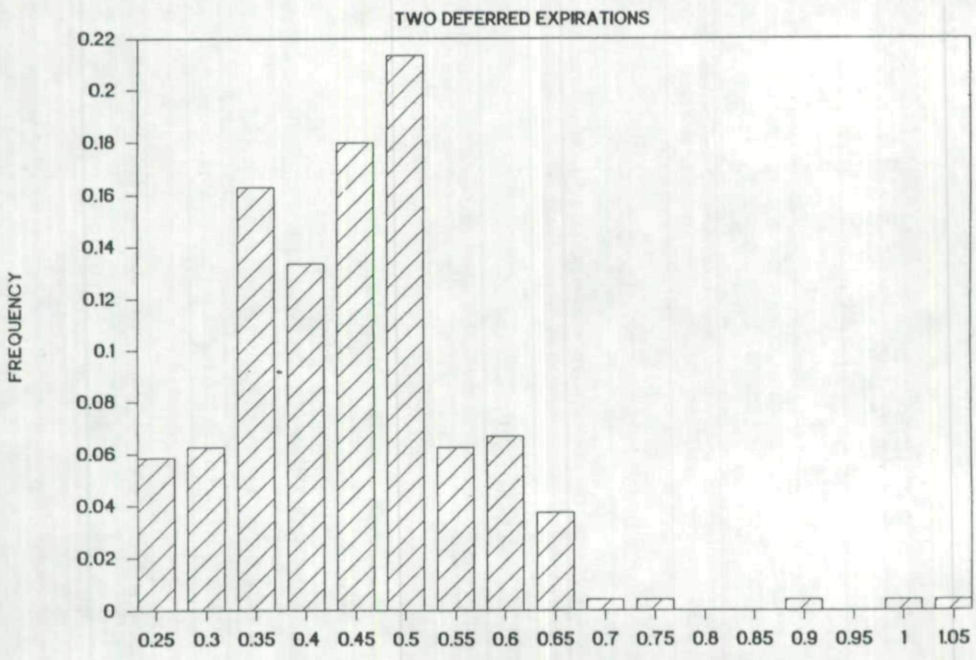
[4]Maddala (1977) shows that standard errors for second-stage regression coefficients are downward biased. He derives a correction factor to correct for this bias. It is the ratio of the regression standard error computed by substituting original (not fitted) endogenous variable values in the second-stage regression equation divided by the regression standard error of the actual second-stage equation.

Table I
DATA SUMMARY

	Wheat		Soybeans		Major Market Index (MMI)	
	First 2 Expirations	2 Deferred Contracts	First 2 Expirations	4 Deferred Contracts	First 2 Expirations	1 Deferred Contract
Average absolute value of price changes						
Average	0.303 ¢/bu.	0.436 ¢/bu.	0.523 ¢/bu.	0.612 ¢/bu.	0.112 pts/contract	0.151 pts/contract
Std. dev.	0.065	0.142	0.098	0.139	0.037	0.047
Maximum	0.569	1.75	1.26	1.5	0.3	0.4
Minimum	0.25	0.25	0.313	0.25	0.07	0.066
No. of dual own-account traders (DT)						
Average	9.12	2.23	15.6	2.35	2.4	0.11
Std. dev.	4.29	1.72	11.8	2.45	2	0.32
Maximum	21	8	41	14	9	1
Minimum	1	0	0	0	0	0
No. of non-dual own-account traders (NT)						
Average	12.07	3.17	49.4	9.87	21.8	4.9
Std. dev.	5.61	1.49	36.5	6	12.7	2.9
Maximum	29	7	126	36	56	14
Minimum	3	0	0	2	0	0
$DT/(DT + NT)$						
Average	0.431	0.39	0.23	0.169	0.084	0.015
Std. dev.	0.109	0.23	0.07	0.116	0.06	0.04
Maximum	0.765	1	0.66	0.5	0.33	0.25
Minimum	0.125	0	0	0	0	0
Volume						
Average	1685	187	4271	554	454	94
Std. dev.	1606	231	4427	740	360	117
Maximum	8695	1628	21,415	6067	1855	613
Minimum	89	3	17	17	1	0
Contracts included	Dec 88 Mar 89	May 89 Jul 89	Nov 88 Jan 89	Mar 89, May 89 Jul 89, Nov 89	Oct 88 Nov 88	Dec 88



(A)



(B)

Figure 1
(A and B) Wheat distribution of half-hour averages, in cents per bushel (*continued*).

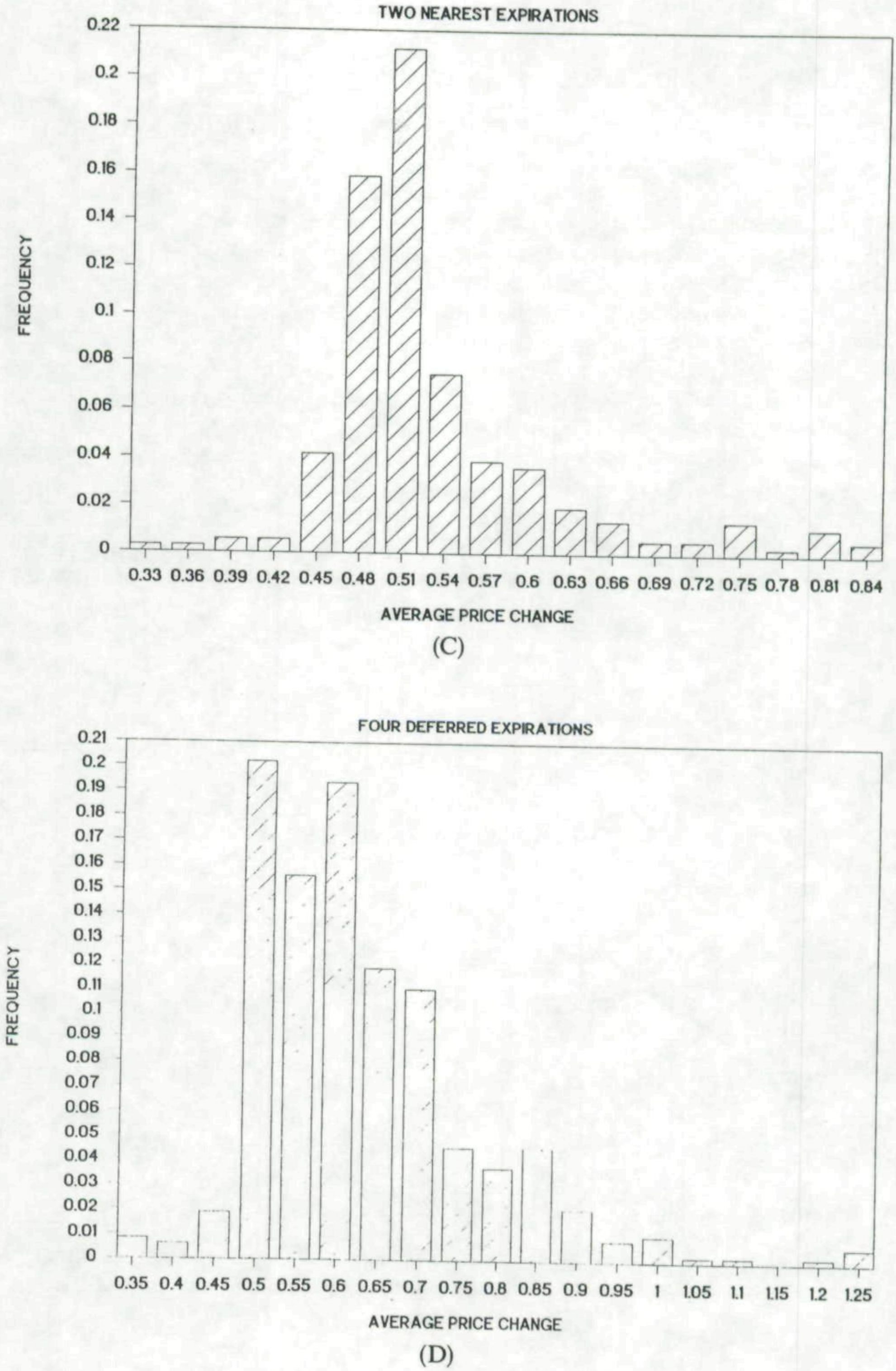
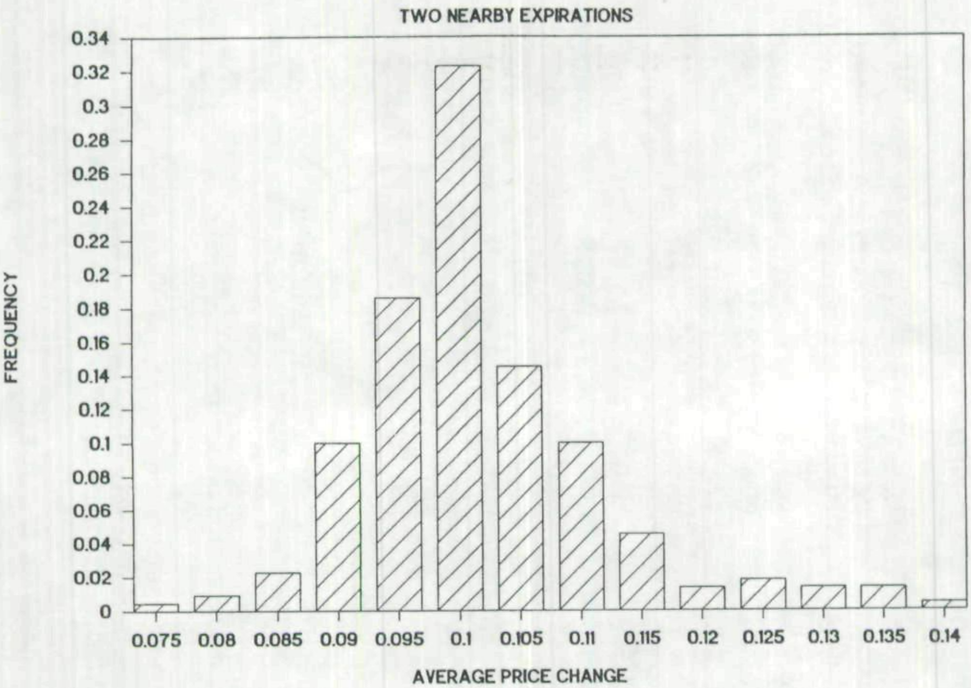
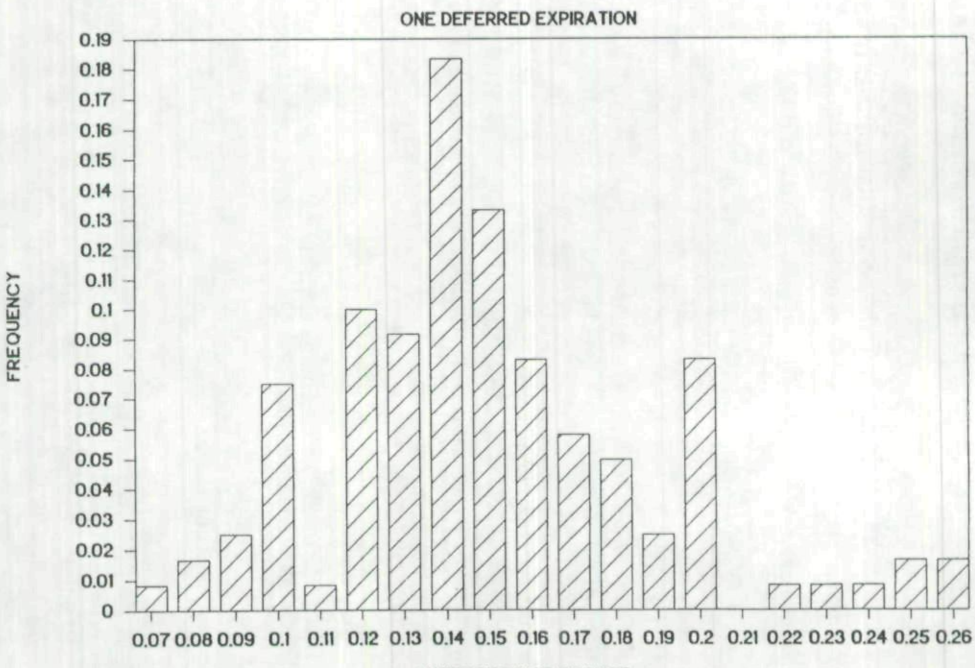


Figure 1 (continued)
(C and D) Soybeans distribution of half-hour averages, in cents per bushel (*continued*).



(E)



(F)

Figure 1 (continued)
Major Market Index distribution of half-hour averages, in index points per contract.

Table II
SECOND-STAGE REGRESSION RESULTS FOR LIQUIDITY MODEL*

Independent Variable	Wheat 2 Nearby Expirations	Wheat 2 Deferred Expirations	Soybeans 2 Nearby Expirations	Soybeans 4 Deferred Expirations	Major Market Index (MMI) 2 Nearby Expirations	Major Market Index (MMI) 1 Deferred Expiration
Total CTI1 traders ($DT + NT$)	−.00055	.0086	−.0004	−.00046	−.00043	.0061
	(.0088)	(.012)	(.00025)[c]	(.0038)	(.00053)	(.0067)
Dual CTI1 traders/total	−.52	−.45	−.98	−.75	−.47	−.77
($DT/(DT + NT)$)	(1.35)	(.22)[b]	(.35)[a]	(.94)	(.26)[b]	(.69)
Price variance ($t - 1$)	.035	.023	.012	.017	.029	−.046
	(.20)[b]	(.02)	(.0031)[a]	(.004)[a]	(.02)[c]	(.06)
Customer's share of all volume	.14	.18	.095	−.028	.0082	.011
	(.15)	(.084)[b]	(.068)[c]	(.067)	(.018)	(.027)
Average trade size	.00077	−.0013	.00071	−.0009	−.00024	.00082
	(.0019)	(.0015)	(.00074)	(.0008)	(.00074)	(.0019)
Back-month variable	.0028	—	−.034	.045	.02	—
	(.17)		(.016)[b]	(.017)[a]	(.007)[a]	
New crop year dummy	—	−0.087	—	−.007	—	—
		(.035)[b]		(.035)		
Constant	.43	.70	.71	.60	.15	.13
	(.80)	(.08)[a]	(.10)[a]	(.13)[a]	(.012)[a]	(.04)[a]
Adjusted R^2	.29	.14	.35	.23	.30	.00
Durbin–Watson	1.56	1.59	1.46	1.49	1.23	1.38
Standard error adjustment factor	1.52	1.31	1.34	1.19	1.06	1.25
Sample size	240	238	238	480	253	123

*Number of dual and non-dual CTI1 traders ($DT + NT$), and ratio of dual to total CTI1 traders ($DT/(DT + NT)$ are fitted values. Reported standard errors are original standard errors multiplied by adjustment factor listed at table bottom. Dependent variable: average absolute value of price changes in time bracket t.

[a]Coefficient is statistically significant at the 99% level of confidence.

[b]Coefficient is statistically significant at the 95% level of confidence.

[c]Coefficient is statistically significant at the 90% level of confidence.

effect of a dual own-account trader on bid–ask spreads is, on average, the coefficient on $DT + NT$ plus $B\{DT/(DT + NT)^2\}$, where B is the coefficient on the $DT/(DT + NT)$ variable.

As expected, estimated bid–ask spreads are generally found to be higher when previous period prices are more volatile, as measured by the variance of prices. This is consistent with the hypothesis that bid–ask spreads increase when there is greater risk to holding positions in inventory. In the agricultural markets, three of the four coefficients on the customer's share of all volume variable are positive; but those coefficients are statistically significant only for the deferred months in wheat and nearby contracts in soybeans. The significant results can be interpreted as evidence in support of the notion that market makers raise bid–ask spreads when a larger share of volume is information trading (customer trading), or they can be interpreted as an upward shift in demand for liquidity.

In no case is the coefficient on average trade size statistically significant. Except for the nearby soybean sample, estimated bid–ask spreads are generally higher the longer the time until contract expiration, all else constant.[5] Estimated bid–ask spreads appear to be lower in new crop year contracts, but the coefficient on this variable is statistically significant only in the deferred wheat sample. The adjusted R^2 values indicate that, in the nearby contract samples (those representing the great majority of all volume), around 30% of the variation in estimated bid–ask spreads is associated with variation in the independent variables used in the model. Explanatory power is lower for the deferred contracts.

EXPLANATION

The results presented above indicate that, for some samples of the wheat, soybean, and MMI futures markets, estimated bid–ask spreads decline when dual traders trading for their own account comprise a larger share of all own-account traders. The results might be explained by the cost structure of dual trading "firms" and a differing extent of the asymmetric information effect.

One reason own-account trading by dual traders may provide a greater contribution to tightening bid–ask spreads is related to order processing costs. Dual traders are permitted to produce two products: brokerage services and speculative or liquidity services. This multiple product capability suggests that the dual trading firm's behavior is directly analogous to "job shops" in the manufacturing sector, whose economic behavior is analyzed by Clemens (1951). In the context of Clemens' analysis, dual trading firms have market access and skills that can be used to produce multiple products. They enter the market with an ability to produce execution services for others and for themselves. Thus, the dual-trading firm faces two separate markets and demand curves for its production capacity. As a profit maximizing entity, each dual-trading firm is expected to first specialize in the market offering him the most inelastic demand curve. The firm continues to specialize in producing this particular service until its marginal revenue equals marginal cost. At that point, if the firm has excess capacity for execution services, it produces services for the market with the more elastic demand curve. This occurs until mar-

[5]The significant negative coefficient on the back-month variable in the nearby soybean sample simply indicates that bid–ask spreads were tighter in the March 1989 contract than in the November 1988 contract. During the time period of this study, the November 1988 contract was approaching expiration, and the March 1989 was becoming the nearby contract.

ginal revenue from production of this service equals marginal cost.[6] More importantly, as Clemens explains, regarding the "secondary" market:

> In theory this market is one with the most elastic demand, and the price established for it would be barely in excess of marginal revenue and marginal cost.... The virtual equivalence of marginal cost, marginal revenue and demand in the marginal market merely recognizes the fact that practically all firms produce some products at little or no profit which are just on the margin of being dropped. (pp. 7–8)

The implication of Clemens' argument is that if dual traders have excess capacity, they will use their excess resources in their secondary market becoming the marginal low-cost supplier to that market. The price in the market for liquidity approximates the marginal cost of liquidity trading.[7] Dual-trading firms that specialize in brokerage are expected to be the low-cost providers of liquidity services as their secondary product; and, given the competitive pressure to equate marginal cost and price, the bid–ask spread (the price of liquidity services) is expected to narrow with an increased supply of liquidity trading by dual traders. This behavior is consistent with results that show average bid–ask spreads to be smaller when dual traders comprise a larger share of all liquidity traders.

An additional explanation for the findings is related to the asymmetric information problem associated with trading against informed customers. Dual traders may not be as exposed to such adverse effects because they are well informed about some small portion of all customer trading patterns; that is, they clearly see the pattern of trades they execute for customers. The probability of loss on a dual traders' own-account trades is expected to differ due to differences in information access. Specifically, the dual trader is exposed to the information available to pure locals in the pit, plus the information generated from his brokered order flow. Assuming dual traders' customers possess a different information set, own-account trades executed by dual traders after executing customer trades are less information-deficient than those of nondual traders (of course, federal law and exchange regulations prohibit trading ahead of customer orders). Being exposed to more information reduces the probability of loss, and, hence, lowers the marginal cost of speculative trades for the dual trader. Although all speculative trades add liquidity, dual traders have a slightly smaller asymmetric information problem than nondual liquidity traders; thus, they may be able to pressure the market to provide more competitive prices, i.e., smaller bid–ask spreads.[8]

CONCLUSIONS

The results presented here suggest that bid–ask spreads in some segments of CBOT wheat, soybean and MMI futures markets are lower when dual traders comprise a larger share of all own-account traders. The strength of the conclusions is limited due to the short time frame covered by the data, the use of a proxy for bid–ask spreads, and the presence of several tricky econometric problems.

[6]This hypothetical production pattern is validated by the fact that most dual-trading firms specialize in brokerage or liquidity services. See CFTC (1989).

[7]Note that the minimum bid–ask spread is institutionally constrained by the minimum tick size of a contract.

[8]The price of brokerage services is not examined; however, in theory, similar results should be realized due to parallel arguments. Dual traders should have lower marginal costs for brokerage due to their ability to more effectively address the cost of out-trades.

The findings may be explained by the fact that dual traders are multiple product producers and, thus, have an order processing cost advantage in the production of liquidity services (i.e., speculative trading). Also, they can accept smaller bid–ask spreads than nondual traders since their information deficiency relative to informed customers is smaller. These factors allow dual traders to put downward pressure on bid–ask spreads, the price of liquidity services.

The evidence presented here provides support for the argument that restrictions on dual trading would hurt liquidity in some markets. Restrictions on dual trading activity would alter the structure of dual trading firms by restricting their production to a single product (brokerage or liquidity trading), thus eliminating the basis for their advantage in the production of liquidity services. Such changes would restrict competition in the provision of liquidity services, and could impose higher transaction costs on market users.

Evidence is found that dual trading enhances liquidity in some markets. The hypothesized benefits that dual trading provides to the brokerage business are not examined. The policy process should compare these benefits with the hypothesized (but not quantitatively identified) benefits of restricting dual trading.

APPENDIX

FIRST-STAGE REGRESSIONS USED TO PREDICT $DT + NT$ AND $DT/(DT + NT)$

	Constant	Broker Volume $(t - 1)$	Price Variance $(t - 1)$	(Price Variance $(t - 1))^2$	Back-Month Variable	Open Bracket Dummy	Close Bracket Dummy	Average Trade Size	Customer's Share of All Volume	New Crop Year Dummy	Adjusted R^2
Wheat,											
2 nearby contracts											
$DT + NT$	20.6[a]	0.00038[a]	0.086	−0.28[c]	−13.1[a]	6.1[a]	3.2[a]	0.26[a]	−4.7[b]	—	0.79
$DT/(DT + NT)$	0.47[a]	−0.000003	0.025	−0.0007	—	0.038[b]	0.026	−0.0028[a]	0.13[a]	−0.46[a]	0.10
2 deferred contracts											
$DT + NT$	0.44	0.00078[a]	−0.15	0.022	—	1.76[a]	0.74[b]	0.049[a]	−2.5[a]	1.79[a]	0.43
$DT/(DT + NT)$	0.24	0.000083[a]	0.046	−0.0053	—	0.063[b]	0.0078	0.0021	0.10	0.0096	0.06
Soybeans,											
2 nearby contracts											
$DT + NT$	91.4[a]	0.0037[a]	−10.1[a]	0.69[a]	−14.2[a]	2.8	10.5[b]	−0.57[a]	−62.8[a]	—	0.59
$DT/(DT + NT)$	0.28[a]	0.0000014[b]	−0.013[b]	0.0011[b]	0.15[b]	−0.033[a]	−0.037[a]	−0.00039	−0.086[b]	—	0.10
4 deferred contracts											
$DT + NT$	21.8[a]	0.0017[a]	0.097	0.0019	−3.8[a]	2.1[a]	1.1[c]	0.0042	0.21	1.24[c]	0.57
$DT/(DT + NT)$	0.18[a]	0.0000079[b]	0.0036	−0.00015	−0.0029	−0.01	0.0054	0.00048	−0.036	−0.023	0.03
MMI											
2 nearby contracts											
$DT + NT$	22.5[a]	0.039[a]	−33.8[a]	34.1[b]	−12.5[a]	6.3[a]	−1.2	−0.16[b]	0.42	—	0.55
$DT/(DT + NT)$	0.063[a]	0.000073[a]	−0.11[c]	0.15	−0.006	−0.0073	−0.022[c]	−0.0028[a]	0.046[b]	—	0.08
1 deferred contract											
$DT + NT$	4.7[a]	0.0025	10.4[b]	−15.7	—	1.39[b]	1.96[b]	0.0042	−1.3	—	0.05
$DT/(DT + NT)$	0.011	0.000073	−0.17[b]	0.20	—	0.0043	−0.0025	0.0016	0.012	—	0.01

DT = number of dual traders trading for their own account per time bracket, NT = number of nondual CTI1 (own-account) traders active per time bracket.
[a]Statistically significant at the 99% level of confidence.
[b]Statistically significant at the 95% level of confidence.
[c]Statistically significant at the 90% level of confidence.

Bibliography

Clemens, E.W. (1951): "Price Determination and the Multiple-Product Firm," *Review of Economic Studies,* 19:1–11.

Commodity Futures Trading Commission (1989, November): "Economic Analysis of Dual Trading on Commodity Exchanges," Division of Economic Analysis, Washington, DC.

Grossman, S. J. (1990, July): "An Empirical Analysis of the Quality of Order Execution of Dual and Nondual Trading Brokers," unpublished manuscript.

Maddala, G. S. (1977): *Econometrics,* New York: McGraw-Hill.

Maddala, G. S. (1983): *Limited Dependent and Qualitative Variables in Econometrics,* Cambridge, UK: Cambridge University Press.

Smith, T., and Whaley, R. E. (1990, June): "Assessing the Costs of Regulation: The Case of Dual Trading," unpublished manuscript.

Stanley, K. L. (1981): "Measuring the Operational Costs of Dual Trading: An Analytical Framework," *The Journal of Futures Markets,* 1:329–336.

Stoll, H. R. (1989): "Inferring the Components of the Bid–Ask Spread: Theory and Empirical Tests," *The Journal of Finance,* March, 115–134.

Thompson, S. R., and Waller, M. L. (1988): "Determinants of Liquidity Costs in Commodity Futures Markets," *Review of Futures Markets,* Spring, 111–125.