FUTURES MARKET TRANSACTION COSTS

PETER R. LOCKE
P. C. VENKATESH

INTRODUCTION

Financial market transaction costs are of general interest to academic researchers, as well as being of particular interest with regard to the optimal design and regulation of market microstructure. For example, regulators might be interested in knowing the relationship between microstructure regulation, such as dual trading restrictions, and transaction costs [Chang and Locke (1994), and Smith and Whaley (1994b)]. Likewise, the analysis of transaction costs features prominently in the debate over the relative efficiency of screen based trading [Fremault Vila and Sandmann (1995)]. Holding all else constant, if two market structures are similar in other respects, then the market with lower transaction costs is more efficient. In seminal research leading to asset market transaction cost analysis, Demsetz (1968) points out that the appropriate measure of asset market transaction costs for regulatory analyses is the aggregate dollar flow from customers (as a group) to marketmakers. Information about the counterparty identity (i.e., marketmaker versus customer) to each trade is essential to this computation, but is typically not publicly available. Lacking this key information, the quoted bid–ask spread or the

The comments of Corinne Bronfman, and the comments and assistance of Steve Mann, are greatly appreciated. Views stated herein are those of the authors and do not necessarily reflect the views of the Commodity Futures Trading Commission, the Office of the Comptroller of the Currency, or their respective staffs. Direct comments to Peter Locke, C.F.T.C., Three Lafayette Centre, Washington, D.C., 20581, phone: 202-418-5287.

- *Peter R. Locke is a Financial Economist at the Commodity Futures Trading Commission.*
- *P. C. Venkatesh is a Financial Economist at the Office of the Comptroller of the Currency.*

The Journal of Futures Markets, Vol. 17, No. 2, 229–245 (1997)

CCC 0270-7314/97/020229-17

implied effective bid–ask spread computed from the available transaction price series is often used as a proxy for transaction costs.[1,2]

As discussed below, there exist a number of factors which may confound the mapping from the implied effective bid–ask spread to actual transaction costs. The usefulness of bid–ask spread estimators for regulatory analysis requiring transaction costs is clearly an important matter for empirical evaluation. This article provides such an analysis. Futures floor trader data is employed, which includes the requisite counterparty information, to compute futures transaction costs, i.e., the dollar flows from customers to marketmakers. This direct measure of transaction costs is compared to two common effective spread estimators. Based on these results, the estimators do not appear to be applicable to futures markets. The appropriateness of applying these estimators to other markets such as equities is not covered by this article.

The intuitive appeal of the bid–ask spread as a social transaction cost is readily seen in the typical microstructure model in which customers trade only with marketmakers (i.e., in markets where there are no intercustomer trades), with the bid–ask spread, on average, centered around the intrinsic asset value.[3] In such a situation, customers are said to incur a transaction cost equal to the difference between the true price and the bid or ask price on every trade [Demsetz (1968)]. In a round trip, that is, a purchase and sale, the customer pays the bid–ask spread to the marketmaker. Stoll (1989) observes that even when all trades are between customers and marketmakers, the quoted bid–ask spread overstates transaction costs if: (i) a subset of customers is better informed than the marketmakers; or (ii) marketmakers adjust the bid and ask prices to manage inventory level. In place of the quoted spread, he advocates the use of

[1]Smith and Whaley (1994b, p. 222) lay out this reasoning in defense of the need to develop the method of moments estimator: "Measuring directly the effective spread (i.e., actual transaction cost) in futures markets is not possible because the trading records of the marketmakers are not publicly available." Thus, it should be clear that the value under discussion is the *social* transactions costs, i.e., the flow of cash to the marketmakers, rather than the private transactions cost defined on a trade-by-trade basis.

[2]Roll (1984) demonstrated how the effective bid–ask spread might be estimated from the time-series of transaction prices. Neiderhofer and Osborne (1962) had earlier suggested that serial correlation in stock prices is induced by bid–ask bounce. Glosten and Harris (1988) extend Roll to incorporate marketmakers who trade against (noisily) informed traders, although their model requires quantities as well as prices. Stoll (1989) combines the time-series of prices and bid–ask quotations to infer the components of the bid–ask spread. Smith and Whaley (1994a) adjust Roll's measure to accommodate the time series of futures prices.

[3]Other social transaction costs may include brokerage commissions, transaction fees, etc. Unlike the bid–ask spread, these components typically do not depend directly on the trading mechanism/environment or on the security's return properties (e.g., volatility), although they may depend on the price level. To a first approximation, they may be considered less important when considering the impact of regulatory policy changes, etc.

the effective or realized bid–ask spread, defined as the average price at which the marketmaker sells minus the average price at which the marketmaker buys, over some interval. Based on this concept, estimators of the effective bid–ask spread, which utilize only the time-series of transaction prices, have been developed.

Violation of the modeling assumptions could cause the spread estimators to be either biased or noisy estimators of the actual transaction cost. Two specific "real-world" factors which are typically not incorporated into the underlying models should be noted.[4] The first factor is that inter-customer trades can and do take place in most financial markets. Using Demsetz' methodology, customer-to-customer trading entails zero transaction costs to customers in the aggregate.[5] For such trading there is clearly no flow of funds from customers to marketmakers. The second factor arises from the minimum price change, or tick. The equilibrium transaction cost per contract need not be an exact integer multiple of the minimum tick. This will be increasingly problematic the smaller the transaction cost is in relation to the minimum tick. For example, in Eurodollar or T-bond futures, the futures price may be observed fluctuating between only two ticks for an extended period of time. Yet, marketmakers and customers during this period may be observed buying and selling at both prices, so that the effective spread (the customers' per-contract transaction cost) is less than the minimum tick.[6] Without a formal model of marketmaking behavior under these conditions, the effect on the time series of prices and estimators based thereon is indeterminate.

A direct measure of transaction costs circumvents these inferential problems, since the costs are calculated by simply measuring aggregate customer trading profits over an interval. Early efforts in the use of trader data for this purpose include Beebower and Priest (1980), and Baesel et al. (1983), who calculated transaction costs for stock and options cus-

[4]Another factor which will confound the inference of transactions costs from the time series of prices is when the demand for liquidity is nonstationary. If there are jumps in the demand for liquidity, these may correspond to larger swings in price than would be indicated by a bid–ask spread. Manaster and Mann (1995) refer to this as the "timing" component of marketmaker income.

[5]This result does not depend on the location of the transaction price relative to the "intrinsic value." Customer-to-customer trades may involve a transfer from one customer to another customer with lower demand for immediacy, but this is simply a transfer and not a social cost. Even in equity markets such as the New York Stock Exchange (NYSE), it has long been known that specialist participation rates are far less than 100%. For some recent evidence, see Madhavan and Smidt (1993).

[6]Such behavior is consistent with strategies which locals have apparently been forced into employing. Since incoming customer orders are probabilistic, the optimal strategy for a competitive marketmaker constrained by price discreteness may involve a mixed strategy, varying bids and offers across a variety of prices above and below the expectation of the intrinsic value. Note that buying above (selling below) the intrinsic value may be optimal when the marketmaker has a short (long) inventory. In addition, such behavior could be related to marketmakers' attempts to disguise their trading strategy or position from other marketmakers.

tomers, respectively. Chang and Locke (1994) used the direct futures transaction costs measure in a regulatory cost analysis. In the extreme case, if all trades are between customers, there are no execution-related transaction costs. On the other hand, if all trades are between customers and marketmakers, then customer costs would equal marketmaker revenues. Most real world microstructures lie somewhere between these extremes. However, since counterparty data has not been widely available, estimators have been used that are based on prices alone, with the second scenario above (only customer-to-marketmaking trading) as an underlying assumption.

Here, futures market transaction costs are compared to the values obtained from two price-series based estimators: Roll's (1984) estimator and Smith and Whaley's (1994a) estimator. These estimators have been applied to futures markets in general analysis [Laux and Senchak (1992)] or to estimate differences in transactions costs across time or markets [Smith and Whaley (1994b); Fremault Vila and Sandmann (1995)]. Thus, it is important and timely to evaluate the applicability of these estimators to futures markets. This study finds that these estimators bear no relationship to futures market transaction costs. This finding shows that the incorporation of actual market stylized facts into microstructure modeling appears to be a fruitful and perhaps critical subject for future microstructure research. See Chordia and Subrahmanyam (1995) and Chakravarty and Holden (1994) for recent efforts in this direction, applied toward the NASDAQ quote issue.

BACKGROUND

Price Series and Transaction Costs

This section presents a thumbnail summary of the evolution of the price time series bid–ask spread estimators. The fundamental assumption is that there is a group of competitive floor traders (or a single zero-profit marketmaker) executing trades for their personal accounts and providing liquidity services through this trading. Further, an often overlooked assumption in these models is the restriction that all trades occur between marketmakers and customers.[7]

The sequence of transaction prices from this microstructure, P_t, is:

[7]While it is easy to see that any customer-to-customer trading would complicate the use of the time series of prices for the calculation of transaction costs, note that inter-marketmaker trading must also be assumed away in these models to eliminate the effect that these transaction prices would have on the time series of prices. In other words, both Roll's model and Smith and Whaley's extension have rather restrictive assumptions regarding trading.

$$P_t = V_t + Q_t \frac{\theta}{2} \tag{1}$$

where V_t is generally referred to as the marketmaker's prediction of the assets' intrinsic value, θ represents the round trip (a combination of a buy and a sell) cost of trading incurred by marketmakers (including, perhaps, the *per contract* opportunity time cost), and Q_t takes on the values of either 1 or -1, denoting a trade at the ask or bid, respectively. If changes in V_t are independent of Q_t, and trades at the bid or the offer are always equally likely, then:

$$E[\Delta P_t \cdot \Delta P_{t-1}] = -\frac{1}{4}\theta^2 \tag{2}$$

where Δ is the lag difference operator. Glosten (1987) analyzes the case where changes in V_t are correlated with Q_t, such as when trading is a noisy signal from informed customers to the marketmaker about changes in the intrinsic value. This extension is particularly important in the case of individual equities where there is a greater likelihood of trading based on private information.

The value θ_R is defined by:

$$\theta_R = -2\ \mathrm{sgn(COV)}\ \sqrt{|\mathrm{COV}(\Delta P_t \Delta P_{t-1})|} \tag{3}$$

where sgn(COV) is the sign of the covariance of price changes. This is the convention used by Roll (1984) and the succeeding literature, which "preserves the monotonicity of the estimator." In fact, the notion of a negative effective bid–ask spread (i.e., the result when there is a positive serial covariance of price changes) makes sense over finite intervals, since it coincides with periods when the marketmaker is on the "wrong" side of a moving market (keeping in mind that the assumptions of these models force the marketmaker to participate in all trades). For the futures price series, the value of θ_R is one-half of the value implied by (3) (since futures prices are only recorded after a price change).

To derive the method of moments estimator, Smith and Whaley (1994a) model the price changes, ΔY_t, in the futures time and sales record directly:

$$\Delta Y_t = u_t + \delta_t \theta_{MM} \tag{4}$$

where δ_t is defined in a similar fashion to Roll's Q_t, except that there is only a realization when there is a record on the futures time and sales report (generally, after a price change). Repeated transactions at the bid or offer are eliminated in the Smith and Whaley model, to correspond to the time and sales record supplied by the futures' exchanges. The variable u_t is a normally distributed innovation associated with each price change. Estimates of θ_{MM} and σ_u^2, the variance of u_t, can be obtained numerically from the first two moments of the empirical distribution of the absolute value of price change.[8] Note that when the absolute value of a price change is bounded below by a minimum price change, as is the case with the futures price series, this places a lower bound on the estimate of θ_{MM}.[9]

Futures Market Microstructure

This section describes the futures market microstructure and discusses the application of the above mentioned transaction costs estimators to futures markets. The first major point to note is that there are no designated marketmakers in seasoned futures instruments.[10] Futures exchange members are free to broker for customers and/or trade for their own account, often with the ability to move about the exchange floor among different commodities.[11] Floor traders have no marketmaking obligations, and those who appear to be making a market are known as "scalpers." It is well established that a set of futures traders trade actively for their own account in a manner which resembles marketmaking [Working (1964); Silber (1984); Kuserk and Locke (1993)]. For regulatory purposes, a meaningful distinction must be drawn between those individuals (or firms) who use the futures markets for price basing and hedging, i.e., the exchanges' *customers*, and the individuals (or firms) whose personal trading facilitates those customers' trading, i.e., *marketmakers*.[12] Applying the analysis of Demsetz (1968) to futures markets, transaction costs

[8]Smith and Whaley (1994a) attempt to cure the problem of asymmetric information by eliminating futures prices recorded after a bid or an offer has been entered onto the time and sales record. For high volume futures contracts (such as the S&P) there are rarely bids and offers entered into the time and sales record. Further, for macroeconomic variables such as the S&P 500 index futures, the entire concept of asymmetric information is not well defined. Theoretically, one could imagine a trader privy to the Open Market Committee hearings making a play on interest rate sensitive contracts. Most of the time, however, there is no "private" information which would affect these macroeconomic markets.

[9]In the extreme case when the variance is an insignificant fraction of the minimum price change, the estimate of the bid–ask spread is nearly identical to the minimum price change.

[10]For new contracts, exchanges may provide incentives for individuals to formally declare themselves as marketmakers with the attendant obligations (of making a continuous two-sided market).

[11]On the CME, in relatively liquid delivery months for some contract markets, the floor trader may choose either brokering or personal trading, but not both, over any day.

[12]As with the stock market, there may be other firms which are marketmaking with their trades. But, also, as is the case with stocks, such traders are not easily identified. In particular, clearing member proprietary trading may take the form of marketmaking in some markets.

are the flows ("leakage") from customers (or end-users) to marketmakers, or floor traders engaged in personal trading.[13] Thus, a legitimate consideration in regulatory evaluation is to examine changes in these transactions costs, such as the recent efforts by Smith and Whaley (1994b) and Chang and Locke (1996) with respect to dual trading.

Corresponding to the absence of official marketmakers, there is no formal, quoted, binding bid–ask spread in futures markets, since offers to trade are only good as long as the words are in the mouth of the trader. However, clerks hired by brokerage firms hand-flash up-to-date quotations from the pit to the phone desks, so that a quoted bid–ask spread does exist in that sense. Although these quotes are never formally recorded, casual empiricism suggests that the spread quoted by these clerks (for most active futures markets) tends to be the minimum price change. These quotes are partly based on the observation by the clerks of marketmaker quotations. However, in the futures pit floor brokers hold decks of customer limit, stop, and market-if-touched orders, all of which they are obligated to conceal until execution. In addition, floor brokers receive instantaneous market orders from phone desks which are generally hand-flashed into the pit through clerks. Bids and offers shouted in the pits, as well as the traders hitting the bids and offers, will be a mixture of marketmakers and floor brokers executing customer orders. Thus, in addition to trades between marketmakers and customers, inter-marketmaker trading and inter-customer trading are not only possible, but probable.

It might seem that, *on average,* the minimum price change (tick) would act as a lower bound on transaction costs. However, competitive forces on the floor may dictate that the average cost customers pay be smaller than the minimum tick, or not equal to an integer multiple of the minimum tick. Marketmakers faced with this constraint may be forced to adopt complex strategies leading to such an outcome.[14] For example, their strategy may lead to a positive revenue on only a fraction of contracts. Consider the following: the clerks are reporting a spread of 436.35 bid, 436.40 offered. A floor trader's bid of 436.35 for 20 S&Ps is hit; so the floor trader asks 436.40 for 20 and is able to sell 10 at this price. No

[13]Several articles have examined the marketmaking behavior of futures floor traders. Kuserk and Locke (1993, 1994), Manaster and Mann (1993), Silber (1984) and Working (1964) investigated the personal trading strategies and profits of futures floor traders. Chang and Locke (1994); Chang, Locke, and Mann (1994); and Fishman and Longstaff (1992) were particularly interested in the behavior and liquidity consequences of dual trading, the practice whereby a floor trader executes trades for their customers and their own account within a given time interval, typically a day.

[14]The point here is that the equilibrium return to holding a position in the futures market for a marketmaker should depend on such factors as price risk and marketmakers' aversion to risk, and not the minimum price change.

more purchasers come forward immediately, so the floor trader lowers the ask to 436.35, while some traders are bidding 436.30. The trader sells 5 more contracts. Finally, to close out the position, the trader sells 5 contracts "at a loss" for 435.30. This sequence leads to 0.0125 in revenue per contract for the trader, although the quoted spread continually flashed by the clerks would be the minimum price change, or 0.05. If these transactions are all against customers, then customers' transaction costs over this interval would be 0.0125 per contract, or 25% of the minimum tick.

When the appropriate data are available, a direct measure of per contract transaction costs may be calculated, similar to that described in the above example, based on aggregate trader data. This measure is the difference between the average purchase price and the average sale price for all futures customers, with prices weighted by transaction size. [This procedure is alluded to in Stoll (1989), and calculated by Beebower and Priest (1980) and Baesel et al. (1983). The measure is labeled θ_C, and is unbounded, in the sense that it is an ex post accounting measure over a finite interval of the actual transaction costs during this interval. This is clearly the appropriate measure that should be used to examine the costs of regulations or rules which affect market microstructure. It is the value for which the estimators, θ_{MM} and θ_R, are derived in this study.

DATA AND RESULTS

Data

The data come from six months of trading, 1/1/92 to 6/30/92, in 12 futures contracts on the Chicago Mercantile Exchange, for 126 trading days in all. For each commodity the nearby futures month is selected each day, rolling forward to subsequent months on the date when the maximum daily futures volume switches from one month to the next. Two data sources are employed. The first is the quote capture, or time and sales (T&S) record.[15] This is essentially a time-stamped transaction price series

[15]The method by which futures prices are reported differs substantially from equities. First, note that all futures trading occurs as the result of an agreement between individuals physically located in one of the futures pits, whether these individuals are agents or principals. Futures prices are reported by pit observers employed by the respective exchange who make entries into a computer terminal when prices change, from whence the prices are distributed electronically by data vendors. The futures price series is, in a sense, a filtered version of the stock ticker, with any repeated prices removed from the series and no quantities reported. When there is little trading occurring, the exchange may report unaccepted bids which are above the last price or asks which are below the last price. They are noted as such on the tape. These reported bids and offers are relatively infrequent in the more liquid contracts which are used in this analysis. To obtain a useable time series of prices for this research, reported bids and offers are removed, and any resulting repeated transaction prices are removed, to arrive at a consistent series of transaction prices *after* price changes.

collected by the exchange except that prices are reported only after a change in price. Successive trades at the same price are not reported. This data is maintained additionally by the Commodity Futures Trading Commission (CFTC), which is the source for the data used in this study. Smith and Whaley (1994a, 1994b) used the same data source. The data are used here to calculate the time-series based transaction costs estimates, θ_R and θ_{MM}.[16]

The second data source is not as widely available. This is the trade register which is created in the exchange's clearing operation. It is also maintained by the CFTC which is, again, the data source.[17] For each trade, the register includes data on the price and quantity as well as key information on the two floor traders engaging in the trade, including whether the trader is a buyer or seller, whether the trade is for the floor trader's personal account, or for a customer, for the trader's clearing member, or for another floor trader. This data set allows for the direct calculation of θ_C for each day as defined above. Thus, the sample period for all calculations is identical.

Summary Statistics

Table I presents some descriptive statistics from the sample to illustrate the salient features of marketmaking behavior in these contracts. As shown in columns (1) and (2), for this sample, these contracts have clear differences in terms of activity, which may have a bearing on the performance and accuracy of the time series estimators. Lumber and feeder cattle are, on average, the lowest volume contracts, while S&P 500 and Eurodollars are the highest volume contracts. All personal trades by all floor traders are included in the marketmaker statistics, and all customer trades by all floor traders are included in the customer trading statistics.

[16]The purpose of this study is to establish the empirical relevance of these time-series based estimators. However, at the suggestion of a referee, the θ_{MM} measure was calculated for dates surrounding the implementation of the top-step rule. For that exercise, since 77 days are available after the implementation until the October 1987 market crash, an additional 77 days prior to the implementation date are included. The average estimates are similar to those obtained by Smith and Whaley (1994b). However, the most interesting result of this lateral investigation is the finding that by using these samples, the value of θ_{MM} is *lower* after the implementation, consistent with the increased liquidity found by the CME. Smith and Whaley's results, that θ_{MM} is higher after the top-step rule, is most likely due to their use of a 5-year control sample against a 3-month test period. Note that the S&P 500 index rose about 250% over their sample period, suggesting that the dollar value of the spread would be likely to rise as well.

[17]The trade register data has been used by Fishman and Longstaff (1992); Park and Sarkar (1991); and Manaster and Mann (1995, 1996). In particular, the results in Manaster and Mann's work support the finding of relatively low income per contract for marketmakers relative to the minimum price change.

TABLE I

Daily Futures Market Statistics

Commodity	*(1) Customer Volume*	*(2) Marketmaker Volume*	*(3) Number of Marketmakers*	*(4) Participation*	*(5) Inventory per trader*	*(6) Percent Customer Cross*
(1) Live hogs	1591	1257	34.5	56.9	0.17	33.8
(2) Pork bellies	916	922	33.5	64.9	0.23	27.9
(3) Live cattle	3793	2947	58.0	53.3	0.15	38.1
(4) Lumber	223	211	18.2	76.6	0.09	22.0
(5) Feeder cattle	339	265	13.9	64.2	0.35	28.7
(6) Canadian dollar	2262	942	19.5	30.1	−0.59	49.5
(7) Swiss franc	8130	7146	76.6	50.8	−0.31	29.5
(8) Deutsche mark	18669	13915	117.4	40.5	0.12	33.8
(9) Pound sterling	5697	3901	43.8	44.1	−0.47	35.6
(10) Japanese yen	9305	6159	61.0	40.4	0.83	37.3
(11) Eurodollars	31653	36168	155.0	51.0	−1.44	26.6
(12) S&P 500	20725	20688	281.6	57.4	−0.44	33.9

Sample size is 126 trading days in 1991. Customer and marketmaker volumes are the averages of sales and buys. The number of marketmakers is the average number of floor traders who traded at least one contract for their personal account. Inventory is the average daily trading imbalance, buys minus sells, of all floor traders trading for their own account on a day. Participation is the percent of customer trades for which locals are on the opposite side. Percent customer cross is the percent of customer orders for which another customer is on the opposite side of the trade.

Column (3) shows that the number of floor traders engaged in personal trading appears to vary with volume across contract markets, as would be expected. Note that the daily figures used to calculate these averages potentially will include many traders who only trade 1 or 2 contracts on a given day for their own account. See Kuserk and Locke (1994) for a detailed description of the concentration of personal trading among floor traders in futures markets. Columns (1) and (2) also indicate that, across contracts, the mean customer volume is roughly equal to the mean volume of marketmakers. At a glance, this might suggest that marketmakers participate in nearly all customer trades. However, column (4), which gives the percent of customer trades (weighted by volume) for which locals are on the other side, reveals that such is not the case. This measure varies from a low of 30.14% for Canadian Dollars to a high of 76.6% for lumber.

In column (5), the average daily closing position per marketmaker is presented. For 11 out of 12 contracts, the number of contracts held overnight averages less than 1 in absolute value. The fact that the daily inventory per trader is generally near zero across contract markets is further confirmation that futures floor traders do appear to be acting as intraday marketmakers.[18] Finally, complementing the finding that floor traders do not participate in all customer trades, column (6) shows that approximately one-third of customer orders (weighted by volume) are executed against other customer orders.[19]

These facts contrast markedly with the assumptions underlying the traditional microstructure models. While previous work establishes clearly that futures floor traders trading for their personal accounts tend to behave like marketmakers, the nature of the competition and the ability of customer orders to cross have important implications for future microstructure modeling. In these markets, marketmakers are not guaranteed, nor do they appear to receive, the entire customer order flow. The time series of prices associated with this more complex microstructure need not necessarily be useful for transactions cost estimation.

Performance of the Estimators

As already noted, the average transaction cost per contract and the two estimators are computed on a daily basis, for each of the 126 days in the sample period. The transaction cost is the trading profit per contract (round trip) for all customer trades, with daily imbalances marked-to-market at the settlement price. Table II presents the averages of these daily values over the entire sample period. Note that the sample period is the same for the transaction costs and the estimators. In addition, this table presents the minimum allowable price change in each market, and the customer–marketmaker spread. The customer–marketmaker spread is the difference between the average price at which customers buy from marketmakers (floor traders trading for their own account) and the average price at which customers sell to marketmakers. There are several points worth noting in these results. First, the average transaction cost varies quite a bit across contracts. Estimates of θ_C range from \$3 per contract (for Live Cattle) to \$20 per contract (Swiss franc). These values are similar to the marketmaker revenues calculated by Kuserk and Locke (1993). Overall, transaction costs appear to be relatively low for futures trading, as these numbers translate to between 0.0004% and 0.033% of notional value, much less than the 1% or so often cited for equities. Second, transactions costs are generally smaller than the customer–mar-

[18]Even if the floor traders' daily inventories were not found to be close to zero, this would not, in and of itself, negate the assumption that they are marketmaking. In fact, Madhavan and Smidt (1991) offer evidence of equities marketmakers having moving levels of optimal inventory.

[19]Clearing members also account for some trading in these markets, so that the sum of the percentage customer cross and local participation is less than 1.

TABLE II

Direct and Estimated Futures Transactions Costs

Commodity	*(1)* *Minimum Tick*	*(2)* *Customer vs. Marketmaker*	*(3)* θ_C	*(4)* θ_R	*(5)* θ_{MM}
(1) Live hogs	10.00	12.82	7.30	7.94	13.34^{t,s}
(2) Pork bellies	10.00	14.38	10.32	9.95^{s}	19.54^{t,s}
(3) Live cattle	10.00	6.65	3.07	5.89^{t,s}	10.92^{t,s}
(4) Lumber	16.00	21.48	15.92	17.10	38.56^{t,s}
(5) Feeder cattle	11.00	13.85	9.42	7.82	19.21^{t,s}
(6) Canadian dollar	10.00	9.08	6.05	5.92	12.19^{t,s}
(7) Swiss franc	12.50	20.65	20.89	3.30^{t,s}	14.28^{t}
(8) Deutsche mark	12.50	13.39	14.28	4.07^{t,s}	12.81
(9) Pound sterling	12.50	19.57	18.12	2.65^{t,s}	16.25
(10) Japanese yen	12.50	16.27	17.47	5.72^{t,s}	13.64
(11) Eurodollars	25.00	7.82	4.81	20.23^{t,s}	25.01^{t,s}
(12) S&P 500	25.00	21.46	16.58	9.94^{t}	29.49^{t,s}

Transactions costs, θ_C, are calculated from customer trading from the futures trade register. The Roll estimator, θ_R, and the method of moments estimator, θ_{MM}, are calculated from the time and sales record. The numbers in column (2) indicate average customer costs conditional on the fact that a local is on the opposite side of the trade. ^{t}Indicates that the mean value of the estimator is significantly different than the mean value of transactions costs at the 95% level of confidence using the standard t-test. ^{s}Indicates that the sign rank test statistic for the estimator and measured transactions costs is significantly different from zero at the 95% confidence level.

ketmaker spread, with the exception of three of the currencies.[20] Third, the average estimates of θ_R are smaller than θ_{MM} for each contract. The average value of θ_C is also smaller than θ_{MM} for every contract except three currencies (the Deutsche mark, Pound sterling, and Japanese yen). This is not surprising since θ_{MM} is bounded below by the minimum price change, whereas no such constraint applies to θ_C and θ_R. The average value of θ_R is smaller than the average value of θ_C for 4 of the 12 contracts.

The statistical significance of these differences is assessed by the t-test and the signrank test. Values of θ_R (the Roll estimator) or θ_{MM} (the method of moments estimator) which are significantly different from θ_C

[20]The currency futures markets have larger clearing member proprietary trading, which is likely to take the form of marketmaking for larger trades. The transactions cost measure includes all customer trades, whether against floor traders' proprietary accounts, other customers, or clearing member accounts. In fact, the analysis (and correspondingly, previous theoretical and empirical research) has basically ignored the role of clearing member trading for liquidity and transaction costs purposes. Clearing members may be acting as customers, who use their clearing member status to reduce their trading costs, or they may be acting as marketmakers. There are likely both types of clearing members. One mitigating factor to keep in mind is that their trading is typically less than 10% of the volume on any market. At any rate, the modeling of a microstructure with marketmakers, customers, and this third type of trader, the clearing firm with both types of characteristics, is a subject for future research.

are indicated by a "t" (for the t-test) and/or an 's' (for the sign-rank test); a 5% significance level (two-tailed) is used in both cases. θ_{MM} is significantly greater than θ_C for every contract other than the three currencies. θ_R is significantly greater than θ_C for two contracts and significantly smaller than θ_C for six contracts. These results suggest that, for these futures markets, the method of moments estimator tends to overestimate the true transaction cost, while Roll's estimator tends to underestimate the true transaction cost.

Possible Explanations for the Discrepancies

These discrepancies are possibly due to violations of the assumptions underlying the estimators. In the introduction, attention was drawn to two possibilities in particular: (i) inter-customer trading, and (ii) effects arising from the minimum tick. Table III provides data to gauge the influence of these factors on the estimators. It shows: (i) the average percent deviation of the two estimators from measured costs; (ii) the average absolute price change; (iii) for convenience, the proportion of inter-customer trading from Table I; and (iv) the proportion of (the absolute value) of price changes that are of a minimum tick.

The first two columns present the average percent biases of the two estimators across contracts. These numbers range (in absolute terms) from -2.15 for Rolls estimate for Canadian Dollars to over 400% for the method of moments estimator for Eurodollars. Overall, the method of moments estimator tends to have a larger absolute estimated bias than Roll's. The third column presents the percent of customer orders (volume weighted) which cross against other customer orders. While this high degree of inter-customer trading lowers transaction costs below a quoted bid–ask spread, there is no obvious cross-sectional relationship between the observed bias and the percent customer cross. Finally, notice in columns (4) and (5) the relatively high percentage of price changes which are a minimum tick, and the corresponding relationship between the minimum tick and the average absolute price change. These give a rough relative measure of the extent to which the minimum tick is "binding" in a market. For example, Eurodollars has the highest proportion of minimum price changes, with the average absolute price change practically equal to the minimum tick and reveals the greatest bias in the θ_{MM} estimate.

The results in Table II and III show that the method of moments and serial covariance estimators for transaction costs perform poorly as

TABLE III

Estimator Bias and Potential Causes

Commodity	(1) θ_R^B	(2) θ_{MM}^B	(3) Percent Customer Cross	(4) Mean Price Change	(5) Minimum Tick (%)
(1) Live hogs	8.92	82.99	33.8	13.38	10.00 69.96
(2) Pork bellies	−4.42	87.70	27.9	19.67	10.00 27.13
(3) Live cattle	91.86	255.70	38.1	10.92	10.00 91.32
(4) Lumber	7.55	142.52	22.0	42.22	16.00 30.70
(5) Feeder cattle	−16.45	105.24	28.7	19.53	11.00 45.80
(6) Canadian dollar	−2.15	101.49	49.5	12.25	10.00 81.34
(7) Swiss franc	−83.99	−30.71	29.5	14.46	12.50 88.95
(8) Deutsche mark	−70.99	−8.70	33.8	12.90	12.50 98.29
(9) Pound sterling	−85.21	−9.32	35.6	16.75	12.50 78.00
(10) Japanese yen	−66.13	−19.24	37.3	13.66	12.50 92.28
(11) Eurodollars	320.58	419.96	26.6	25.01	25.00 99.97
(12) S&P 500	−40.05	77.86	33.9	29.51	25.00 83.37

The biases for the estimators, defined by $\theta_R^B = (\theta_R - \theta_C)/\theta_C \cdot 100$ and $\theta_{MM}^B = (\theta_{MM} - \theta_C)/\theta_C \cdot 100$, respectively, are calculated, and the averages formed over the sample period. In addition, the average absolute price change, and average percentages of minimum tick price changes are presented. The minimum ticks are established by the exchanges. The percent of customer crosses is reprinted from Table I for expository purposes.

point estimates when applied to futures markets. The estimators may prove useful if they are shown to move with the underlying transaction costs. In other words, it could be the case that the estimates are somehow monotonically related to transaction costs, so that their use as a proxy could be maintained. To assess this, the correlation between each pair of estimators is calculated from the time-series of daily estimates. The mean absolute price change is also included in this analysis. The correlations for each commodity are then averaged and presented in Table IV. The individual correlations with measured transaction costs are generally low, and for some commodities are negative. Overall, the method of moments estimator is most closely aligned with the mean absolute price change, with an average correlation coefficient of .9408 across commodities. Neither estimator is highly correlated with measured transaction costs for

TABLE IV

Correlation of Direct and Estimated Futures Transaction Costs

		$-\theta_C$	θ_{MM}	θ_R
θ_{MM}	Pearson	.1173	—	—
	Spearman	.1201	—	—
θ_R	Pearson	−.1168	−.2315[a]	—
	Spearman	−.0799	−.1852[a]	—
θ_μ	Pearson	.1593	.9408[a]	.2487[a]
	Spearman	.1296	.9829[a]	−.1852[a]

For each of 12 commodities, the Roll, θ_R, method of moments, θ_{MM}, and mean absolute price change, θ_μ, estimates are calculated on a daily basis for the lead month. In addition, for each day, the direct transactions cost, in terms of average cost per round trip for customers, θ_C is calculated. Then the correlation of these values is formed for each commodity, with the means of these correlation coefficients across commodities presented. [a]Indicates that, on average, the correlation coefficient is significantly different from zero at the 95% level of confidence for almost all commodities.

any commodity. These low correlations offer little hope for the fruitful application of these estimators to futures markets.

Various regression specifications are attempted to try to relate the bias and the average transaction costs on both pooled and time series data to inter-customer trading or the percent of minimum price changes. Without a microstructure model incorporating customer-to-customer trading and binding discrete prices, there is no particular constraint on the regression specification. The search for a good fit is restrained only by time. However, none of the many specifications which are estimated prove fruitful in explaining the biases. Clearly, inter-customer trading and discrete price changes will interfere with the use of Roll's and the method of moments estimators to infer transactions costs. However, the nature of the interrelationship of these facts with the price series and resulting estimators is not readily apparent.

Interestingly, Smith and Whaley (1994a) find that for the S&P 500, the method of moments estimator is correlated across time with volume and implied volatility. Since the method of moments estimator appears to be highly correlated with the mean absolute price change, it is easy to see how an increase in volume and volatility, with a likely increase in the proportion of two tick or more moves, will lead to an increase in the calculated value of the method of moments estimator.

CONCLUSIONS

Transaction costs are a critical component of asset market regulatory evaluation. In the extreme case, the asserted benefits from a regulation such as the prohibition of dual trading could possibly be outweighed by increases in transaction costs for end-users. Thus, research in the area of the estimation of asset market transaction costs is clearly warranted. This article shows that common time series estimators of transaction costs appear to be inappropriately applied, at least with respect to futures markets. The application of these estimators to other assets markets such as equities remains an open question, although there should be some concern where the actual market microstructure does not correspond to the assumptions underlying the estimators.

Several important features of futures market microstructure are documented which should prove fruitful in the development of microstructure models. Many futures transactions occur directly between customers, eliminating the leakage on these trades. These transactions between customers (which have zero social costs by Demsetz' definition) lower the average costs per transaction for customers, in the aggregate, below the quoted bid–ask spread. (Some of the quotes will be the vocalization of resting customer limit orders.) In addition, futures transaction costs appear low relative to the minimum price change, or at least are not an integer multiple of the minimum price change, possibly reflecting elaborate pricing and inventory control strategies by marketmakers. Such strategic behavior by marketmakers and customers substantially complicates the drawing of inferences about market microstructure from the observed time series of prices.

BIBLIOGRAPHY

Baesel, J. B., Shows, G., and Thorp, E. (1983): "The Cost of Liquidity Services in Listed Options: A Note," *Journal of Finance,* 38:989–995.

Beebower, G., and Priest, W. (1980): "The Tricks of the Trade: How Much Does Trading Really Cost?" *Journal of Portfolio Management,* 6:36–42.

Brorsen, B. W. (1989): "Liquidity Costs and Scalping Returns in the Corn Futures Market," *The Journals of Futures Markets,* 9:225–236.

Chakravarty, S., and Holden, C. (1995): "An Integrated Model of Market and Limit Orders," *Journal of Financial Intermediation,* 4:213–224.

Chang, E.-C., and Locke, P. R. (1996): "The Performance and Market Impact of Dual Trading: CME Rule 552," *Journal of Financial Intermediation,* 5:23–48.

Chang, E. C., and Locke, P. R., and Mann, S. C. (1994): "The Effect of CME Rule 552 on Dual Traders," *The Journal of Futures Markets,* 14:493–510.

Chordia, T., and Subrahmanyam, A., (1994): "Off-floor Market Making, Payment for Order Flow, and the Tick Size," UCLA Working Paper.

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

Fishman, M., and Longstaff, F. (1992): "Dual Trading in Futures Markets," *Journal of Finance,* 47:643–672.
Fremault Vila, A., and Sandmann, G. (1995): "Floor Trading Versus Electronic Screen Trading: An Empirical Analysis of Market Liquidity and Information Transmission in the Nikkei Stock Index Futures Market," London School of Economics Discussion Paper No. 218.
Glosten, L. (1987): "Components of the Bid/Ask Spread and the Statistical Properties of Transactions Prices," *Journal of Finance,* 42:1293–1308.
Glosten, L., and Harris, L. (1988): "Estimating the Components of the Bid/Ask Spread," *Journal of Financial Economics,* 21:123–142.
Kuserk, G. J., and Locke, P. R. (1993), "Scalper Behavior in Futures Markets: An Empirical Analysis," *The Journal of Futures Markets,* 13:409–431.
Kuserk, G. J., and Locke, P. R. (1994), "Marketmaker Competition on Futures Exchanges," *Journal of Derivatives,* 1:56–66.
Laux, P., and A. Senchak 1992: "Bid–Ask Spreads in Financial Futures," *The Journal of Futures Markets,* 12:621–634.
Madhavan, A., and Smidt, S. 1991: "A Bayesian Model of Intraday Specialist Pricing," *Journal of Financial Economics,* 30:99–134.
Manaster, S., and Mann, S. (1996): "Life in the Pits: Competitive Marketmaking and Inventory Control," *Review of Financial Studies,* 9:953–976.
Manaster, S., and Mann, S. C. (1995): "Sources of Market Making Profits: Man Cannot Live by Spread Alone" Texas Christian University, Working Paper.
Neiderhofer, V., and Osborne, M. F. M. (1966): "Market Making and Reversal on the Stock Exchange," *Journal of the American Statistical Association,* 61:897–916.
Roll, R. (1984): "A Simple Implicit Measure of the Bid/Ask Spread in an Efficient Market," *Journal of Finance:* 39:1127–1139.
Silber, W. (1984): "Marketmaking Behavior in an Auction Market: An Analysis of Scalpers on Futures Markets," *Journal of Finance,* 38:937–953.
Smith, T., and Whaley, R. E. (1994a): "Estimating the Effective Bid/Ask Spread from Time and Sales Data," *The Journal of Futures Markets,* 14:437–456.
Smith, T., and Whaley, R. E. (1994b): "Assessing the Costs of Regulation: the Case of Dual Trading," *The Journal of Law and Economics,* 37:215–246.
Stoll, H. R. (1989): "Inferring the Components of the Bid/Ask Spread: Theory and Empirical Tests," *The Journal of Finance,* 44:115–134.
Working, H. (1967): "Tests of a Theory Concerning Floor Trading on Commodity Exchanges," *Food Research Institute Studies: Supplement.*