Estimating Actual Bid-Ask Spreads in Commodity Futures Markets

Henry L. Bryant

and

Michael S. Haigh*

Paper presented at the NCR-134 Conference on Applied Commodity Price Risk Analysis, Forecasting, and Market Risk Management
St. Louis, Missouri, April 23-24, 2001

* Graduate research assistant (h-bryant@tamu.edu) and assistant professor (mshaigh@tamu.edu),
Department of Agricultural Economics, Texas A&M University

Estimating Actual Bid-Ask Spreads in Commodity Futures Markets

Abstract: Various bid-ask spread estimators are applied to transaction data from LIFFE cocoa and coffee futures markets, and the resulting estimates are compared to observed actual bid-ask spreads. Results suggest that actual bid-ask spreads, which are not reported by most open-outcry futures markets, can be reasonably estimated using readily available transaction data. This is especially important since recent research seems to indicate that efforts to estimate effective spreads using data commonly available from futures markets have not been successful. Thus estimates of actual spreads can give market participants and researchers some idea of potential transaction costs. Accurate estimates of bid-ask spreads will also be needed to assess the relative efficiency of electronic versus open-outcry trading. Results indicate that estimators using averages of absolute price changes perform significantly better at estimating actual bid-ask spreads in futures markets than estimators using the covariance of successive price changes.

Keywords: futures markets, market microstructure, bid-ask spread

Introduction

The costs associated with trading in markets have been the subject of much study in recent years. Beginning with the work of Demsetz (1968), many investigators have been interested in estimating these costs and inferring their determinants. As Demsetz carefully described, market participants often must pay a higher price to buy immediately than the price that they could receive if they wished to sell immediately. The former price is commonly referred to as the ask, and the latter the bid. The difference between these two prices is referred to as the bid-ask spread. Bid-ask spreads in futures markets have been studied extensively. A small sample of recent contributions includes Ma, Peterson and Sears (1992), Ding (1999), Shyy, Vijayraghavan, and Scott-Quinn (1996), and Bae, Chan, and Cheung (1998).

As market participants often must pay this spread, it is thus closely related to the costs associated with trading, and has received much attention from investigators. Indeed, much of the research to date has been dedicated to establishing an accurate way of estimating the bid-ask spread as neither the bid nor the ask are usually reported by most open outcry markets. While several competing bid-ask spread estimators have been proposed to date, the estimators have not been jointly evaluated by comparing their individual predictions of the bid-ask spread to observed bid-ask data. This is particularly surprising, as accurate estimates of bid-ask spreads are useful for realistically evaluating hedging, speculating, and arbitrage strategies. Furthermore, as futures exchanges move to electronic trading, researchers will be interested in measuring changes in market efficiency. One measure of market efficiency is the bid-ask spread, and research comparing open-outcry spreads with electronic spreads seems inevitable. Reliable estimates of spreads that prevailed during open-outcry trading will thus be necessary. It is the purpose and contribution of this paper therefore to carefully assess the performance of various bid-ask spread estimators, using a variety of evaluation criteria, by employing actual bid-ask data for two different types of commodities (coffee and cocoa) which were, until recently, actively traded using open outcry at the London International Financial Futures Exchange (LIFFE).

Many estimators of bid-ask spreads have been suggested (descriptions of those tested in this paper are in the following section). Many of these estimators set out to estimate the *effective* bid-ask spread. Roll (1984) defined the effective spread as the "the spread faced by the dollar-weighted average investor who actually trades at the observed prices". He believed that this would likely be less than the quoted spread, because "actual trading is done mostly within the quotes." Smith and Whaley (1994) defined the effective spread as "the difference between the price at which the market maker buys (sells) a security and the price at which he subsequently sells (buys) it". They explained that this may differ from the quoted spread, as market makers may exit some positions at zero gross profit (so-called "scratch sales"). Recent research (Locke and Venkatesh 1997) has indicated, however, that available estimators do a poor job of estimating these effective spreads.

If effective spreads cannot be estimated reliably, market participants must find other indicators of potential transaction costs. One possibility is using spread estimators to estimate *actual*, rather than effective, bid-ask spreads. Actual spreads are, of course, related to effective spreads. It seems reasonable to believe that actual, quoted price spreads may serve as an upper bound on average, effective spreads. As such, a good estimate of the actual spread may serve as a useful starting point for estimation of transaction costs (in the Locke and Venkatesh sense) using transaction data. Certainly the actual, quoted bid-ask spread represents the "worst case" cost of immediacy that any relatively small individual market customer will incur. Thus estimates of the actual spreads are useful for realistically evaluating hedging, speculating, and arbitrage strategies in the way of Bae, Chan, and Cheung (1998), and will be instrumental in evaluating the merits of electronic trading relative to those of open-outcry trading.

Thus far a direct evaluation of estimator performance in estimating actual spreads has not been undertaken for futures markets, however, as bid-ask quotes have generally not been available. This study conducts such an evaluation using transaction, bid, and ask price observations from the coffee and cocoa markets at LIFFE. Spread estimators that were intended for estimating both effective and actual spreads are evaluated. Locke and Venkatesh and Smith and Whaley both noted that estimates resulting from both types of estimators were highly correlated. It is thus possible that including estimators intended to estimate effective spreads may prove insightful.

The remainder of the paper is organized as follows. First we present a brief overview of the bid-ask spread estimators, and then introduce methods of assessing the estimator accuracy. Next we describe the data and then discuss the results. The last section concludes.

Bid-Ask Spread Estimators

Previous research on bid-ask spread estimators have either utilized the covariance of successive price changes or have employed averages of absolute price changes. The former type of estimator originally applied in equity research was first developed by Roll (1984). Roll made four assumptions, given which he developed a joint price distribution of price changes in a market that included market makers. First, he assumed an informationally efficient market. Second, he assumed that observed price changes had a stationary probability distribution. Third, he assumed that all customers made use of the market maker, who maintained a constant spread, s. Fourth, he assumed successive transactions would be market maker sales or purchases with equal probability. Given these assumptions, he then deduces that any non-zero price changes that are not the result of the arrival of new information will be movements between the bid and ask prices, and any price change of zero is the result of two successive transactions at either the bid or the ask. This implied a joint probability distribution for successive price changes. He then calculated variances of price movements and the covariance of successive price movements (as functions of *s*), and proved that this calculated covariance conditional on no new information arriving was equal to the unconditional covariance of successive price changes. Solving the covariance for equation for *s* resulted in Roll's estimator of the *effective* spread

$$RM = 2\sqrt{-\operatorname{cov}(\Delta p_t, \Delta p_{t-1})}\ . \tag{1}$$

Even though this estimator is intended to estimate effective spreads, it is calculated and compared to observed actual spreads in this study for purposes of comparison. This estimator has not typically been applied to futures transaction data because Roll's forth assumption is often inappropriate for such data.

Roll's estimator explicitly assumed equal probabilities (conditioned on the previous transaction type) of an observed transaction being a bid or an ask. With the U.S. time-and-sales data, however, the probability of observing a bid after a bid or an ask after an ask is zero (assuming no new information has arrived to move the bid and ask prices). Given this situation, if no new news arrives and the market maker(s) charge a constant spread s, then the covariance between successive price changes is $-s^2$. Solving for s results in

$$RM* = \sqrt{-\operatorname{cov}(\Delta p_t, \Delta p_{t-1})}\ . \tag{2}$$

It should be noted that in open-outcry futures markets one trader cannot bid lower than any other current bid, and cannot ask more than any other current asking price (Silber 1984; Frino, McInish, & Toner 1998). Thus trading cannot take place within a prevailing bid-ask spread. Roll reported that he was estimating an effective spread because stock trading can take place within the quoted spread, and so using actual transaction data implied the spread that actual investors faced. Also, with stock transaction data it is possible to observe successive prices that are equal. If no new information has arrived (i.e. transactions are only being observed at one bid price and one ask price; the underlying "true" price is not moving) and there is a constant spread *s*, then a series of price changes from which observations of zero have been removed can only have a sample covariance of successive price changes of $-s^2$. Therefore, when applied to futures transaction data that omits price changes of zero, RM^* can only be estimating the *actual* bid-ask spread. [1] Its effectiveness in this regard is evaluated here.

Chu, Ding, and Pyun (1996) suggested an estimator of the effective spread that relaxed Roll's forth assumption that any given transaction has equal probability of taking place at the bid or the ask. They developed an estimator that incorporates the probability ( ) that an observed transaction takes place at the same price as the previous transaction, and the probability ( ) that an observed transaction takes place at the same price as the next transaction. These probabilities are estimated by applying a test that attempts to identify the price (bid or ask) at which each transaction occurred. The reader is referred to Chu, Ding, and Pyun for the theoretical development of their estimator, as it is too lengthy to reproduce here. The resulting estimator is

$$CDP = \sqrt{\frac{-\operatorname{cov}(\Delta p_t, \Delta p_{t-1})}{(1-d)(1-a)}}\,. \tag{3}$$

Thompson and Waller (1988) referred to the actual bid-ask spread as “the cost of immediate liquidity incurred when entering or exiting a market” or “liquidity cost” for short. They proposed the following actual spread estimator:

$$TWM = \frac{1}{T}\sum_{t=1}^{T}\left|\Delta p_t\right|, \tag{4}$$

where Δp_t, $t = 1,\ldots,T$ is the series of non-zero price changes. They described this as being a function of the average bid-ask spread, and the magnitude and frequency of real price changes. Their estimator presumes that the average bid-ask spread component will be the primary determining factor, and no attempt is made to filter out real price changes. This estimator was applied in Thompson and Waller (1988) to study the determinants of liquidity costs in feed grain markets, and was used to compare liquidity costs between two similar markets in Thompson, Eales, and Seibold (1988). Ma, Peterson, and Sears (1992) used the *TWM* to study intraday patterns in and determinants of various Chicago Board of Trade (CBOT) contracts.

The CFTC estimator of the actual bid-ask spread was described in Wang, Yau, and Baptiste (1997). Like *TWM*, this estimator also takes an average of absolute non-zero price changes, but attempts to remove the effect of real price changes by omitting any price change that follows another price change of the same sign. That is to say, the CFTC estimator is the average, absolute, *opposite direction*, non-zero price change. This requirement that some data be omitted means that a greater quantity of data may be required to calculate a spread estimate. In thinly traded markets, “bounces” between the bid and ask prices may be fairly infrequent while real price changes may be more numerous.[2]

Smith and Whaley (1994) adopted a different strategy to account for the effects of true price changes. They made two assumptions. First, they assumed that the spread is constant over the time frame for which it is being estimated. Second, they assumed that the expected value of *true* price changes is zero. They did not assume, however, that the variance of true price changes is zero, an assumption in *TWM*. Then, taken as given that the observed price series does not include repeated observations of the same price, they derived the first and second population moments of the *observed* price changes. These are functions of both the spread and the variance of true price changes. These population moments were then set equal to the sample moments of the observed price changes, and these two equations were solved for the two variables. Hence Smith and Whaley arrived at an estimator for the spread that explicitly accounts for the effects of true price changes.

Assessing Estimator Accuracy

The role of the bid-ask spread can be incredibly important in any hedging, speculating or arbitrage activity. For instance, as illustrated Bae, Chan, and Cheung (1998) ignoring the role of the bid-ask spread might in fact lead to a decision to undertake what appears to be a profitable trading strategy when, after accounting for the bid-ask spread is, in actuality, unprofitable. However, given that most bid-ask quotes are not recorded and hence observed by many exchange participants the key question is therefore: "How well then does each estimator perform relative to the other estimators, and which estimator which relies on observed price data should a market participant actually use?" Given the availability of actual bid-ask spread data one simple method might be to test the equality of mean squared errors or some measure of economic loss using a simple *t*-test procedure. However, in order to get a better descriptive evaluation of the performance of each estimator in this paper we initially test for differences in the biases, variances and mean squared errors of the estimators by employing a procedure originally developed by Ashley et. al (1980).

Specifically, from the definition or mean squared error, it is simple to show that for two forecasts with errors e_1 and e_2 that:

$$MSE(e_1) - MSE(e_2) = \left[s^2(e_1) - s^2(e_2)\right] + \left[m(e_1)^2 - m(e_2)^2\right], \tag{5}$$

where *MSE* is the sample mean square error, s^2 is the sample variance, and m is the sample mean error. Defining:

$$\Delta_n = e_{1n} - e_{2n} \text{ and } \qquad \Sigma_n = e_{1n} + e_{2n}, \tag{6}$$

then equation (5) can be rewritten as:

$$MSE(e_1) - MSE(e_2) = \left[\mathrm{cov}(\Delta, \Sigma)\right] + \left[m(e_1)^2 - m(e_2)^2\right]. \tag{7}$$

The null hypothesis that there is no difference in the mean squared error of two estimators is then equivalent to the null hypothesis that both terms on the right hand side of (7) are zero. This can be tested by regressing:

$$\Delta_i = b_0 + b_1\left[\Sigma_i - m(\Sigma_i)\right] + u_i\,. \tag{8}$$

This results in least squares estimates:

$$\hat{b}_0 = m(e_1) - m(e_2), \tag{9}$$

and

$$\hat{b}_1 = \left[s^2(e_1) - s^2(e_2)\right] / s^2(\Sigma)\,. \tag{10}$$

Testing the both terms on the right hand side of (7) are zero is equivalent to testing $b_0 = b_1 = 0$. If either of the two least squares coefficient estimates is significantly negative, the null hypothesis that the *MSE*'s are equal is not rejected. If one coefficient estimate is negative but not significantly so, a one-tailed t-test on the other estimate can be used. If both estimates are positive, then an *F*-test that both coefficients are zero can be performed, but a significance level equal to half of the usual level must be used (Ashley, et al. 1980).

In addition to allowing a test of the null hypothesis that two *MSE*'s are equal, estimating (8) also facilitates testing whether or not the biases and variances of two estimators are equal. From (9), it is obvious that an estimate of b_0 that is significantly different from zero implies that two biases are different. Similarly, an estimate of b_1 significantly different from zero implies that that the two variances are different.

The methodology laid out above was applied by Brandt and Bessler (1983) to compare the relative performance of various hog price forecasting methods, and was applied by Bessler and Brandt (1992) to compare the performances of futures market and expert opinion meat price forecasts. Equation (8) is estimated for each combination of two estimators for each commodity in this study to test for equality of their *MSE*'s, biases, and variances.

Moving beyond the Ashley et al. (1980) style of testing procedure, perhaps an even more accurate test would be that competing bid-ask spread estimators embody no useful information absent in the more preferred bid-ask spread estimator. This is essentially the idea behind encompassing which is closely related to conditional misspecification analysis and composite forecasting. In particular, Granger and Newbold (1973) suggested the use of a composite estimator

$$E_{cn} = (1 - l)E_{1n} + lE_{2n}, \tag{11}$$

where E_{1n} and E_{2n} are two component estimators and $l \in [0,1]$ is a parameter to be estimated. The error of this composite estimator is equal to the error of the first component estimator plus λ multiplied by the difference of the errors of the two components. Thus the equation:

$$e_{1n} = l(e_{1n} - e_{2n}) + u_n, \tag{12}$$

can be estimated to determine if estimator 2 contains information not present in estimator 1 (Harvey et al. 1998). If $l = 0$ cannot be rejected, then estimator 2 does not contain any additional useful information, and estimator 1 is said to "encompass" estimator 2. Therefore, in this study, equation (12) is estimated for each permutation of two estimators for each commodity, to determine if any of the estimators are completely useless for this application. As suggested by Harvey et al. (1998), White's heteroskedaticity-consistent variance of the estimate of l is used, as the error series e_{in} exhibits skewness and kurtosis that strongly suggest a non-normal distribution for each estimator *i*.[3]

Data

In open-outcry trading, traders continuously cry out the prices at which they are willing to buy (bids) and prices at which they are willing to sell (asks), though not necessarily both prices simultaneously. Other traders can then accept these offers to buy and sell, resulting in a transaction. On November 27th 2000 the open outcry system used for most of LIFFE's commodity products was replaced by the electronic trading system, LIFFE CONNECT™. As such, all bid/asks and transaction volumes are now available on a real time basis. Before this date (from 1996 onwards), LIFFE did record some bid and ask data from the open outcry system, but prices were only available on a per minute basis. This stands in contrast to the major U.S. futures exchanges, where transactions at price of the previous transaction are not reported, and bids and asks are only reported when little actual trading is occurring (Locke and Venkatesh 1997).

In anticipation of the move to the electronic platform in November 2000 the reporting system in the open outcry trading pit at LIFFE was changed. Specifically, from July 3rd 2000 all bids and asks and transaction data were recorded and made available to the public via the order transit and registration system. This period of time thus provides a unique data set facilitating an accurate empirical research on the microstructure of futures markets in an open outcry environment.

Bid, ask and transaction data for cocoa and coffee futures contracts, time-stamped to the second, are provided by LIFFE on the "LIFFEstyle 2000" data CD. The LIFFE cocoa contract calls for delivery of 10 tonnes (metric tons) of cocoa, with a minimum price fluctuation of one pound sterling per tonne. Delivery months are March, May, July, September, and December. The daily volume of trading in the nearby futures averages about 2,500. LIFFE coffee futures contracts call for delivery of 5 tonnes of robusta coffee. The minimum price fluctuation is one U.S. dollar per tonne, and delivery months are January, March, May, July, September, and November. Daily trading volume in the nearby futures is roughly 2,400 contracts. Examples of the data reported for November 2000 coffee futures on 27 September 2000 are provided in Table 1.[3]

As previously noted, bid and ask prices are not necessarily called out simultaneously by a single trader. Observations of the bid-ask spread for each market are thus constructed by matching a bid or ask price with a price of the opposite type that occurred within a chosen time interval. Bid and ask prices called out in open-outcry futures trading are only required to be honored if they are immediately accepted by another trader, although it has been noted that in practice traders (especially scalpers) let their bids and offers "live" (Silber 1984).[4] Thus the choice of the time interval used to construct spread observations presents a tradeoff. Relatively restrictive criteria naturally result in fewer spread observations, but one can be more assured that these observations represent a valid actual spread. Less restrictive criteria result in more observations, but some of these observations may be too far apart in time to have constituted an actual spread.

A second, related criteria must be considered. The resulting spread observations are then used to calculate daily average spread observations. In order to ensure that a given daily average is in fact representative of the spreads that prevailed on that day, some minimum number of spreads used to calculate a daily average.

In this research the highest quality of observations (shorter time interval for spreads, more spreads per day when constructing a daily average) was used that still allowed an acceptable quantity of observations for reliable statistical analysis. The chosen criteria were a 10-second time interval for constructing a spread, and a minimum of 20 spreads for a daily average. [5] Varying these criteria somewhat did not result in significant changes to the qualitative results reported below. Applying the 10-second criterion to the data in Table 1, bid-ask spreads of $1 per tonne are observed at 10:04 a.m. and 10:18 a.m.

The average daily spread for a contract typically follows a "u-shaped" pattern in which it is higher when the delivery date is distant, decreases as time passes, and eventual increases as the delivery date approaches. This is consistent with previous research. As an example, spreads for the November 2000 coffee contract are plotted over time in Figure 1.

The transaction observations provided by LIFFE include consecutive transactions at equal prices. From this data, a "raw" series of price changes is constructed, which is then used in the calculation of *RM*. It should be noted that this type of transaction price series is not reported by the major U.S. exchanges, and so the *RM* estimator could not be applied to U.S. data in the way that it is applied here. A series consisting of strictly non-zero price changes is constructed, which is then used to calculate RM^*, *CDP*, *TWM,* and *SW*. This second price change series is thus like that which would be reported by a U.S. futures exchange. Lastly, a series of only opposite-direction price changes is assembled for use in calculating *CFTC*. This last price change series typically contains about half as many price changes as the strictly non-zero price change series, which in turn usually contains about half as many price changes as the unrestricted price change series.

Results

The daily average bid-ask spread is estimated for each day of each delivery over the time period from 3 July 2000 through 24 November 2000. Some difficulties arise in applying the spread estimators. First, the serial covariance-type estimates, *RM*, RM^*, and *CDP* cannot be calculated due to price changes that exhibit positive serial covariance. This occurs relatively more often for cocoa (about 44% of observations) than for coffee (about 20% of observations). Within each commodity, the problem occurs more often for the serial covariance estimators using only price-changing observations (*RM** and *CDP*). This problem with serial covariance estimators has been noted by many other researchers. For instance, Chu, Ding, and Pyun noted that positive serial covariance in price changes could be due to sequential information arrivals and Roll himself suggested that markets may exhibit inefficiencies over shorter time frames, which could be manifested as positive serial covariance in price changes. Observations where *RM*, RM^*, and *CDP* encounters the problems described above are omitted from the analysis.

Correlations between the daily average spreads and estimated average spreads for each commodity are given in Table 2. All of the estimates are more highly correlated with the daily average spreads for coffee than for cocoa, with the exception of *RM*. The correlations between the serial covariance estimates and the average spreads are positive, but not especially high, ranging between 0.10 and 0.32. Correlations between the remaining estimates and average spreads are more impressive, falling in the 0.47 to 0.85 range. In this respect, *TWM*, *SW*, and *CFTC* appear to do a much better job than *RM*, RM^*, and *CDP*. Also, *TWM*, *SW*, and *CFTC* are highly correlated with one another, and *RM*, *RM**, and *CDP* are relatively highly correlated with one another. Thus estimators of the same type (serial covariance-type estimators or absolute price change-type estimators) seem to be highly correlated with one another, and noticeably less correlated with estimators of the other type. Interestingly, estimators that are trying to estimate actual spreads (*RM**, *TWM*, and *CFTC*) are not necessarily highly correlated with one another, and are not necessarily more highly with the actual spread than the estimators that are trying to estimate effective spreads (*RM*, *SW*, and *CDP*).

Performance of the estimators using various measures for all observations are given for each commodity individually in Table 3. The performance of the estimators *relative* to one another is similar within each commodity. The absolute price change-type estimators seem to perform much bettor than the serial covariance type estimators by each of the performance measures. Among the absolute price change estimators, relative performance is very similar for cocoa. However the *SW* estimator performs somewhat worse than *TWM* and *CFTC* when estimating coffee spreads. Thus the relative performance *SW* estimator may be somewhat inconsistent across commodities.

Comparing the *absolute* performance of the estimators across commodities using the mean absolute percent error measure, the absolute price change estimators seem to perform worse when estimating coffee spreads than when estimating cocoa spreads. This suggests that the results regarding the absolute magnitudes of the performance measures of these estimators should not be extrapolated to markets for which testing has not been performed.

The results from the estimation of equation (8) for each combination of commodities are presented in Table 4. In almost all cases, the null hypotheses that $_0 = 0$ is rejected at the 5% level of significance, meaning that for the most part the differences in the biases (mean errors) reported in Table 3 are significant. The sole exception is that the difference in the biases of *TWM* and *CFTC* for cocoa are not significantly different. In most cases the null hypothesis $_1 = 0$ also cannot be rejected, with the interesting exceptions being that the error variances of *TWM* and *SW* are not significantly different for cocoa, and the error variances of *CFTC* and *TWM* are not significantly different for coffee.

It should be noted at this point that all results reported thus far are base on all data for all contracts. The u-shaped pattern in Figure 1 suggests that conditions over the life of a contract vary, and thus performance of spread estimators may thus vary by time to delivery. However, only the aggregate results are only presented as separating the data into nearby and distant groups revealed only a single interesting difference in performance. This difference is that for cocoa, the bias of the *CFTC* estimator improved to be significantly better than the *TWM* estimator, and the variance of the *CFTC* estimator improved to be not significantly different from the *SW* and *TWM* estimators. Thus the performance of the *CFTC* estimator may be somewhat better when estimating spreads for a nearby delivery.

Analyzing the signs of the coefficient estimates in Table 4, the biases of the serial covariance estimators are greater than the absolute price change estimators (significantly positive b_0 estimates), while the variances of the absolute price change estimators are greater (significantly negative b_1 estimates). This naturally suggests one to question which class of estimators generally has lower means of squared errors. As discussed earlier, in some cases an F-test can be used to test the null hypothesis that both b_0 and b_1 from equation (8) are zero for a pair of commodities estimators, implying that the mean squared errors of the two estimators are not significantly different. However if one of the two coefficient estimates is significantly negative, this null hypothesis automatically cannot be rejected. This is the case for most of the possible pairs of estimators in this study, and thus the Ashley methodology is largely powerless for finding differences in the mean squared errors here. Although the statistical methodology available cannot prove that the means of the squared errors of the serial covariance estimators are greater than those of the absolute price change estimators, the relative magnitudes reported in Table 3 strongly suggest that this is the case. Still, those interested in minimizing error variance (at the expense of significantly higher error bias) may wish to consider the serial covariance estimators.

The other criteria employed here to evaluate the bid-ask spread estimator performances is the forecast encompassing testing procedure described previously. Probability values for the tests that $l = 0$ from equation (12) for each permutation of two estimators are presented in Table 5. In most cases, the null hypothesis that one estimator encompasses another is rejected. In only one case is this hypothesis not rejected across both commodities: we cannot reject that *CDP* encompasses *RM*. Since encompassing is generally rejected, it is quite possible that a composite estimator could provide superior estimates of actual bid-ask spreads. In particular, one might speculate that combining a lower variance serial covariance estimator and a lower bias absolute price change estimator might prove fruitful.

Conclusion

Estimates of bid-ask spreads are calculated using transaction data from LIFFE coffee and cocoa futures markets. These estimates are then compared to actual spreads observed in those markets during the same period, and the performances of the estimators are then evaluated using various criteria.

Results suggest that actual bid-ask spreads, which are not reported by most open-outcry futures markets, can be reasonably estimated using readily available transaction data. This is especially important since recent research seems to indicate that efforts to estimate effective spreads using data commonly available from futures markets have not been successful. Thus estimates of actual spreads can give market participants and researchers some idea of potential transaction costs.

The mean absolute price change estimators, *TWM*, *CFTC,* and *SW*, perform better at estimating daily average bid-ask spreads than the serial covariance estimators, *RM*, *RM**, and *CDP*, by the bias and mean square error criteria (although statistical differences between the means of squared errors could not be found here). The serial covariance estimators have lower variances than the absolute price change estimators, however. Encompassing test results generally confirm that the estimators do not encompass one another, and there may be gains from combining estimates.

This research should not only be of academic interest as a contribution to the market microstructure literature, but should also be of interest to futures market practitioners, as the effect of the bid-ask spread on a trading strategy can be extremely important, even though it is rarely observed in practice. Understanding the magnitude of the bid-ask spread using the appropriate estimator is therefore important for any successful trading endeavor. While this paper has analyzed the performance of various estimators using open outcry data from LIFFE, it would be of interest to analyze the behavior of the spreads now that the trading system has changed. This and other interesting issues are left for future research.

Bibliography

Ashley, R., Granger, C. W. J., & Schmalensee, R. (1980): Advertising and aggregate consumption : An analysis of causality. Econometrica, 48, 1149-1167.

Bae, K., Chan, K., & Cheung, Y. (1998): The Profitability of Index Futures Arbitrage: Evidence From Bid-Ask Quotes. The Journal of Futures Markets, 18, 743-763.

Bhattacharya (1983): Transactions Data Tests of Efficiency of the Chicago Board of Options Exchange. Journal of Financial Economics, 12, 161-186.

Bessler, D. A. & Brandt, J. A. (1992): An analysis of forecasts of livestock prices. Journal of Economic Behavior and Organization, 18, 249-263.

Brandt, J. A. & Bessler, D. A. (1983): Price Forecasting and Evaluation: An Application in Agriculture. Journal of Forecasting, 2, 237-248.

Chu, Q. C., Ding, D. K., & Pyun, C. S. (1996): Bid –Ask and Spreads in the Foreign Exchange Market. Review of Quantitative Finance and Accounting, 6, 19-37.

Ding, D. K. (1999): The Determinants of Bid-Ask Spreads in the Foreign Exchange Futures Market: A Microstructure Analysis. The Journal of Futures Markets, 19, 307-324.

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

Frino, A., McInish, T. H., & Toner, M. (1998): The Liquidity of Automated Exchanges: New Evidence from German Bund Futures. Journal of International Financial Markets, Institutions, and Money, 8, 225-241.

Followill, R. A., & Helms, B. P. (1990): Put-Call-Futures parity and Arbitrage Opportunities in the Market for Options on Gold Futures. The Journal of Futures Markets, 10, 339-352.

Granger, C. W. J. & Newbold, P. (1973): Some Comments on the Evaluation of Economic Forecasts. Applied Economics, 5, 35-47.

Gwilym, O., Clare, A., & Thomas, S. (1998): Price Clustering and Bid-ask Spreads in International Bond Futures. Journal of International Financial Markets, Institutions, and Money, 8, 337-391.

Harvey, D. I., Leybourne, S. J., & Newbold, P. (1998): Test for Forecast Encompassing. Journal of Business & Economic Statistics, 16, 254-259.

Locke P. R., & Venkatesh, P. C. (1997): Futures Market Transaction Costs. The Journal of Futures Markets, 17, 229-245.

Ma, C. K., Peterson, R. L., & Sears, R. S. (1992): Trading Noise, Adverse Selection, and Intraday Bid-Ask Spreads in Futures Markets. The Journal of Futures Markets, 12, 519-538.

Roll, R. (1984): A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market. The Journal of Finance, 23, 1127-1139.

Shyy, G., Vijayraghavan, V., & Scott-Quinn, B. (1996): A Further Investigation of the Lead-Lag Relationship Between the Cash Market and Stock Index Futures Market With the Use of Bid-Ask Quotes: The Case of France. The Journal of Futures Markets, 16, 405-420.

Silber, W. (1984): Marketmaker Behavior in an Auction Market: An Analysis of Scalpers in Futures Markets. The Journal of Finance, 4, 937-953.

Smith, T., & Whaley, R. E. (1994): Estimating the Effective Bid/Ask Spread from Time and Sales Data. The Journal of Futures Markets, 14, 437-456.

Thompson, S., Eales, J. S., & Siebold, D. ((1993): Comparison of Liquidity Costs Between the Kansas City and Chicago Wheat Futures Contracts. Journal of Agriculture and Resource Economic, 18, 185-197.

Thompson, S. R., & Waller, M. (1988): Determinants of Liquidity Costs in Commodity Futures Markets. Review of Futures Markets, 7, 110-126.

Wang, H. K. W., Yau, J., & Baptiste, T. (1997): Trading Volume and Transaction Costs in Futures Markets. The Journal of Futures Markets, 17, 757-780..

Figure 1: Daily average bid-ask spread for November 2000 coffee futures (dollars per tonne)

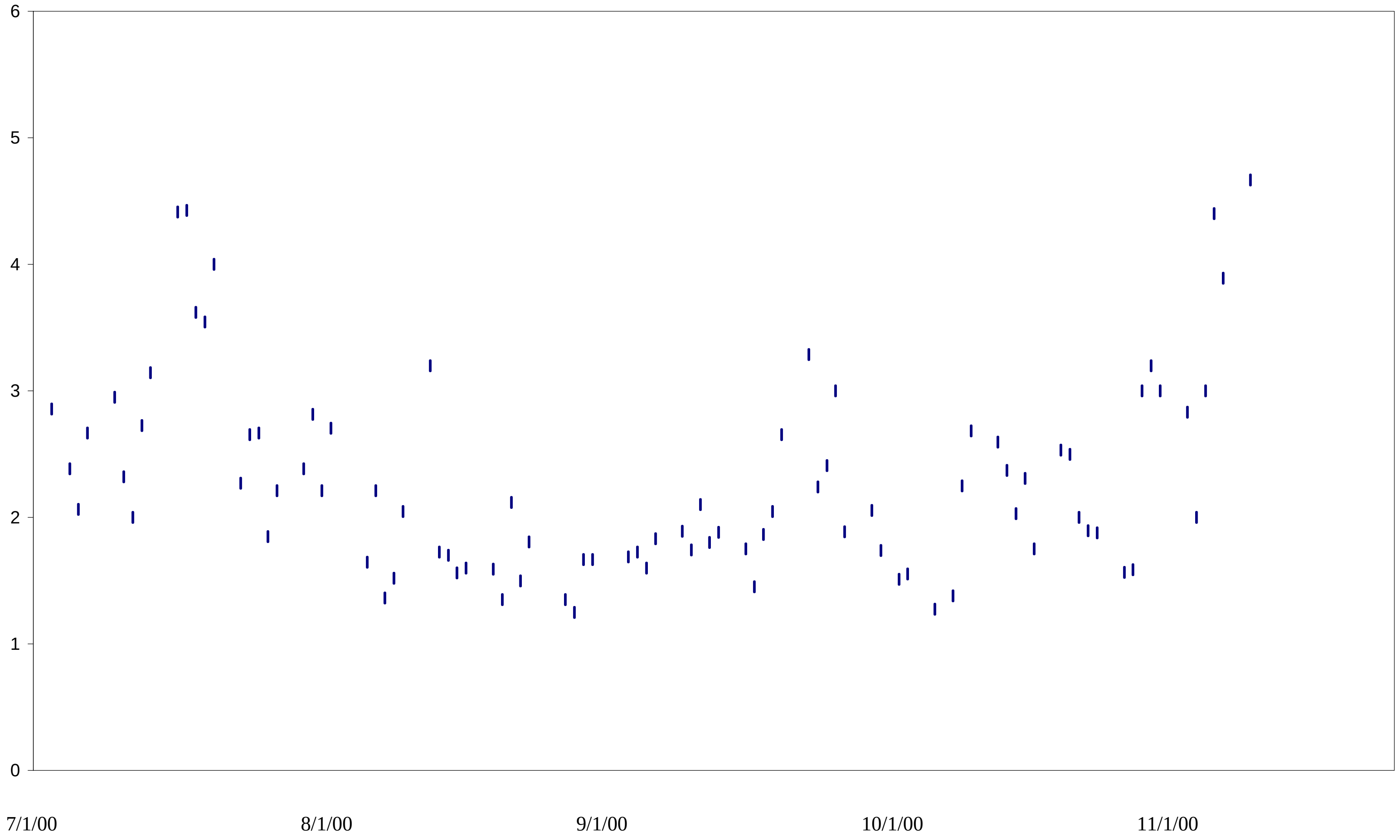

Table 1: Example of LIFFE data

Date	*Time*	*Delivery*	*Type*	*Volume*	*Price*
10/27/00	10:03:50	Nov-00	Bid	0	701
10/27/00	10:04:12	Nov-00	Bid	0	702
10/27/00	10:04:49	Nov-00	Ask	0	702
10/27/00	10:04:50	Nov-00	Bid	0	701
10/27/00	10:04:51	Nov-00	Trd	3	702
10/27/00	10:05:16	Nov-00	Ask	0	703
10/27/00	10:05:31	Nov-00	Trd	5	701
10/27/00	10:05:45	Nov-00	Trd	5	701
10/27/00	10:07:09	Nov-00	Trd	20	703
10/27/00	10:08:18	Nov-00	Bid	0	702
10/27/00	10:11:12	Nov-00	Trd	20	702
10/27/00	10:11:24	Nov-00	Trd	1	703
10/27/00	10:18:15	Nov-00	Ask	0	702
10/27/00	10:18:16	Nov-00	Bid	0	701
10/27/00	10:19:37	Nov-00	Trd	1	702
10/27/00	10:19:38	Nov-00	Trd	1	702
10/27/00	10:19:41	Nov-00	Trd	1	701

Source: London International Financial Futures and Options Exchange (LIFFE). “Type” refers to type of price observation. “Trd” denotes a trade observation.

Table 2: Correlations of daily average spreads and estimates of daily average spreads

Cocoa	*RM*	*RM**	*CDP*	*TWM*	*CFTC*	*SW*	*Spread*
RM	1.00	0.71	0.71	0.49	0.50	0.46	0.32
*RM**		1.00	0.90	0.40	0.44	0.35	0.10
CDP			1.00	0.57	0.60	0.51	0.20
TWM				1.00	0.85	0.96	0.60
CFTC					1.00	0.84	0.47
SW						1.00	0.59

Coffee	*RM*	*RM**	*CDP*	*TWM*	*CFTC*	*SW*	*Spread*
RM	1.00	0.72	0.70	0.41	0.43	0.20	0.12
*RM**		1.00	0.93	0.41	0.43	0.15	0.11
CDP			1.00	0.55	0.63	0.23	0.24
TWM				1.00	0.93	0.93	0.85
CFTC					1.00	0.86	0.82
SW						1.00	0.80

RM: Roll's measure; *RM*:* Modified Roll's measure; *TWM:* Thompson-Waller measure; *CFTC:* Commodity Futures Trading Commission estimator; *SW:* Smith and Whaley estimator.

Table 3: Performance of estimators by commodity

	RM	RM*	CDP	TWM	CFTC	SW	RM	RM*	CDP	TWM	CFTC	SW
Cocoa	*Pounds per tonne*						*Pounds per contract*					
Mean error	-0.77	-0.94	-0.52	-0.18	-0.17	-0.21	-7.74	-9.38	-5.19	-1.84	-1.65	-2.14
Mean squared error	0.73	1.00	0.52	0.08	0.10	0.09	7.26	9.96	5.25	0.84	0.95	0.91
Root mean squared error	0.85	1.00	0.72	0.29	0.31	0.30	8.52	9.98	7.24	2.89	3.08	3.02
Mean absolute error	0.78	0.94	0.62	0.23	0.23	0.24	7.82	9.38	6.18	2.26	2.32	2.39
Mean absolute percent error	51.72	61.77	40.80	14.15	14.45	14.84						
Total number of observations	111	100	100	149	149	148						
Serial correlation errors	38	49	49	N/A	N/A	N/A						
Coffee	*Dollars per tonne*						*Dollars per contract*					
Mean error	-1.02	-1.22	-0.76	-0.47	-0.44	-0.55	-5.10	-6.12	-3.82	-2.34	-2.19	-2.74
Mean squared error	1.33	1.75	0.91	0.31	0.30	0.44	6.67	8.74	4.53	1.57	1.52	2.19
Root mean squared error	1.15	1.32	0.95	0.56	0.55	0.66	5.77	6.61	4.76	2.80	2.76	3.31
Mean absolute error	1.02	1.22	0.78	0.48	0.45	0.55	5.10	6.12	3.91	2.40	2.27	2.75
Mean absolute percent error	51.00	62.62	39.03	22.37	20.97	25.03						
Total number of observations	123	117	117	143	143	137						
Serial correlation errors	20	26	26	N/A	N/A	N/A						

RM: Roll's measure; *RM*:* Modified Roll's measure; *TWM:* Thompson-Waller measure; *CFTC:* Commodity Futures Trading Commission estimator; *SW:* Smith and Whaley estimator.

Table 4: Coefficient estimates and *p*-value for differences in bias and variance components for each pair of bid-ask spread estimators

	0					*1*				
Cocoa	*RM**	*CDP*	*TWM*	*CFTC*	*SW*	*RM**	*CDP*	*TWM*	*CFTC*	*SW*
RM	-0.201	0.230	0.586	0.624	0.563	-0.025	0.205	-0.319	-0.219	-0.358
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.532)	(0.000)	(0.000)	(0.000)	(0.000)
*RM**		0.420	0.743	0.798	0.718		0.211	-0.261	-0.181	-0.282
		(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)
CDP			0.324	0.378	0.300			-0.475	-0.397	-0.509
			(0.000)	(0.000)	(0.000)			(0.000)	(0.000)	(0.000)
TWM				0.019	-0.032				0.084	-0.020
				(0.084)	(0.000)				(0.001)	(0.137)
CFTC					-0.052					-0.103
					(0.000)					(0.000)
Coffee	*RM**	*CDP*	*TWM*	*CFTC*	*SW*	*RM**	*CDP*	*TWM*	*CFTC*	*SW*
RM	-0.245	0.219	0.566	0.615	0.522	0.011	0.077	-0.340	-0.301	-0.275
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.640)	(0.022)	(0.000)	(0.000)	(0.000)
*RM**		0.459	0.775	0.828	0.732		0.064	-0.312	-0.277	-0.247
		(0.000)	(0.000)	(0.000)	(0.000)		(0.002)	(0.000)	(0.000)	(0.000)
CDP			0.316	0.369	0.290			-0.385	-0.329	-0.338
			(0.000)	(0.000)	(0.000)			(0.000)	(0.000)	(0.000)
TWM				0.031	-0.079				0.048	0.106
				(0.043)	(0.000)				(0.054)	(0.000)
CFTC					-0.096					0.076
					(0.000)					(0.014)

RM: Roll's measure; *RM*:* Modified Roll's measure; *TWM:* Thompson-Waller measure; *CFTC:* Commodity Futures Trading Commission estimator; *SW:* Smith and Whaley estimator. $_0$ > 0 implies that the bias of the estimator in the row is greater than the bias of the estimator in the column. $_0$ < 0 implies the opposite. $_1$ > 0 implies that the variance of the estimator in the row is greater than the variance of the estimator in the column. $_1$ < 0 implies the opposite. *P* – values close to zero suggest that the bias and or/variance of two estimators is statistically different

Table 5: P-values for encompassing tests

Cocoa	*RM*	*RM**	*CDP*	*TWM*	*CFTC*	*SW*
RM		0.000	0.000	0.000	0.000	0.000
*RM**	0.000		0.000	0.000	0.000	0.000
CDP	0.583	0.003		0.000	0.000	0.000
TWM	0.000	0.000	0.000		0.217	0.429
CFTC	0.000	0.000	0.000	0.000		0.001
SW	0.000	0.000	0.000	0.000	0.007	

Coffee	*RM*	*RM**	*CDP*	*TWM*	*CFTC*	*SW*
RM		0.000	0.000	0.000	0.000	0.000
*RM**	0.000		0.000	0.000	0.000	0.000
CDP	0.059	0.000		0.000	0.000	0.000
TWM	0.000	0.000	0.000		0.010	0.000
CFTC	0.000	0.000	0.000	0.144		0.011
SW	0.000	0.000	0.000	0.000	0.000	

RM: Roll's measure; *RM*:* Modified Roll's measure; *TWM:* Thompson-Waller measure; *CFTC:* Commodity Futures Trading Commission estimator; *SW:* Smith and Whaley estimator. *P*-values are for the test of H_0: the estimator in a row encompasses the estimator in a column. A *p*-value close to zero suggests that the estimator in a particular row does not encompass an estimator in a particular column.

Endnotes

1. This is another possible explanation why Locke and Venkatesh found that estimators did a poor job of estimating *effective* spreads (the average net income of market makers per trade). If there are no transactions taking place between the best bid and best ask, effective spreads will differ from actual spreads only to the extent that traders are entering or exiting positions using limit orders rather than market orders. It is hard to imagine how this information might be conveyed by the transaction price series. The astute reader will notice that the formula for *RM** is the same as that for the spread estimator of Followill and Helms (1990). However their estimator was applied to data in which trading between a formal spread was possible, and they thus reported that they were estimating effective spreads. For the reasons outlined above, this formula is estimating actual spreads when applied in this situation, and thus there is a subtle difference.

2. Another estimator, proposed by Bhattacharya (1983), is the average of an even smaller subset of absolute price changes. Because the markets considered here have fairly low volumes except in the contracts nearest delivery, this estimator would have frequently not produced an estimate. Those interested in estimating actual spreads for higher volume commodities or contracts may wish to consider this estimator.

3. All data are subjected to a screening algorithm and obviously erroneous observations are removed.

4. This stands in contrast to electronic data whereby any bid or ask that are reported by the exchange as standing limit orders and will exist until the trader actively withdraws the bid or ask. As such, the bid-ask data series from an electronic trading environment looks very different than that from an open outcry environment.

5. Prices must be successive. For example, suppose a bid occurs at 10:00:00, and another, different bid occurs at 10:00:03. Then, an ask is observed at 10:00:07. This ask would not be mated with the first bid, even though they both occurred within 10 seconds of one another. In an earlier version of this paper the same analysis was conducted on the open outcry trade data provided by LIFFE from 1996 to July 2000 (before the reporting system changed). As mentioned previously, this data series was comprised of bid and ask quotes on a per minute basis. Consequently, this data series meant that many of the bids and asks reported within the same minute did not represent a valid spread (e.g., non positive spreads) and so did not represent the true course of events within that minute. Results from this analysis, that excluded these non-positive spreads were not entirely dissimilar to the results presented in this paper and are excluded to conserve space. They are, however, available from the authors upon request.