Estimating Liquidity Costs in Agricultural Futures Markets using Bayesian Methods

Julieta Frank and Philip Garcia*

Selected Paper prepared for the presentation at the American Agricultural Economics Association Meeting, Long Beach, California, July 23-26, 2006

---

*Julieta Frank (jfrank1@uiuc.edu) is a Graduate Research Assistant in the Department of Agricultural and Consumer Economics at the University of Illinois at Urbana-Champaign.
Philip Garcia is a Professor in the Department of Agricultural and Consumer Economics at the University of Illinois at Urbana-Champaign.

Estimating Liquidity Costs in Agricultural Futures Markets using Bayesian Methods

Abstract

Estimation of liquidity costs in futures markets is challenging because bid-ask spreads are usually not observed. Several estimators of liquidity costs exist that use transaction data, but there is little agreement on their relative accuracy and usefulness, and their performance has been questioned. We use a Bayesian method proposed by Hasbrouck which possesses conceptually desirable properties to estimate liquidity costs of six agricultural future contracts. The method builds on Roll's model and uses Markov Chain Monte Carlo estimation. Our Bayesian estimates are lower than more traditional estimates and as anticipated decrease even more when more realistic assumptions such as discreteness are incorporated. The findings demonstrate the need for further research to clarify the usefulness and accuracy of the procedure.

Key Words: agricultural futures markets, liquidity, bid-ask spread, Bayesian estimation, Markov Chain Monte Carlo estimation, Gibbs sampler

Estimating Liquidity Costs in Agricultural Futures Markets using Bayesian Methods

The cost of liquidity, often referred to as the bid-ask spread, is the difference between the prices for immediate purchase and sale (Bryant and Haigh 2004). This difference is a significant source of transaction costs that are usually ignored in economic analysis due to a lack of relevant data. The lack of data becomes particularly important in commodity futures markets where bids and offers occur in an open outcry pit and are not recorded. To circumvent this problem, spread estimators have been proposed that use transaction data only. Some examples are serial covariance estimators (Roll 1984; Chu et al. 1996), and mean absolute price change estimators (Thompson and Waller 1988; Wang et al. 1997; Smith and Whaley 1994).

These spread estimators are simple and straightforward to implement. However, they have weaknesses. For example, the covariance between adjacent price changes can yield positive values when using the Roll estimator, making it difficult to obtain spread estimates. The Thompson-Whaley estimator can fail to distinguish between true price change volatility and volatility attributable to the bid-ask price bounce (Smith and Whaley 1994). Bryant and Haigh (2004) report a downward bias of the above estimators using coffee and cocoa spread data from the London International Financial Futures Exchange (LIFFE). Locke and Venkatesh (1997) and Smith and Whaley (1994) suggest that Roll's estimator is inadequate for futures markets and Stoll (1989) report downward biases of this estimator for stocks markets. Laux and Senchak (1992) modify the Roll estimator to correct observed bias, but their approach does not eliminate the occurrence of negative spreads in financial futures.

Recently, Hasbrouck (2004) implemented a Bayesian Markov Chain Monte Carlo (MCMC) algorithm, the Gibbs sampler, to generate estimates of liquidity costs. Within this framework he estimates unobserved spreads and trade direction indicators based on Roll's model. Bayesian techniques are attractive in this context for a number of reasons. Estimation is based on parameters' posteriors which incorporate all the information in the observed transaction prices. Despite being based on Roll's model, Hasbrouck's procedure does not contain the problem of unfeasible values (i.e. positive covariance between adjacent price changes) because the parameters are random draws from their conditional distributions. Moreover, unobserved latent variables, like the trade direction indicator, are estimated conditional on observed transaction prices rather than derived from tick rules. Another motivation for this proposed method is the ease of computation. Conventional estimation techniques require computing the whole probability space for the trade directions ($q_1, \ldots, q_T$), which involves 2^T combinations, whereas the Gibbs sampler approach needs only T updates for each iteration[1]. In addition, the proposed Bayesian method can accommodate discreteness of futures prices.[2] Discreteness needs to be accounted for because future prices' movements are limited by the tick size. Locke and Venkatesh (1997) find that the proportion of minimum tick size changes has great impact in the performance of spread estimators.

Accurate estimates of liquidity costs are of interest to exchanges, market participants, and researchers. For exchanges, knowing the cost of providing liquidity in their different markets can help develop strategies for the development of new products

[1] For $T = 10^3$ observations and $n = 10^5$ iterations, $2T \sim 10^{300}$ and nT=10^8!
[2] This method can also accommodate clustering, the tendency of futures prices to cluster at natural multiples of the tick size, and the influence of the size of transaction. However, here we do not study these aspects of the markets.

as well as in the regulation of market-making services of existing products. For example, liquidity costs might be useful to assess the quality of the hedging service provided by a futures contract (Pennings and Meulenberg 1997). For market participants, estimates of liquidity costs in different markets and exchanges are useful in making operational decisions. Brorsen, Buck, and Koontz (1998) suggest that wheat hedgers would maximize their utility by choosing the Chicago Board of Trade (CBOT) if they are slightly risk averse and face high liquidity cost differences, but the Kansas City Board of Trade (KCBT) is a better (utility maximizing) option if they are faced with low liquidity cost differences. For researchers, understanding the structure of liquidity costs in futures markets may provide a more comprehensive view of the pricing process. Much research has been done in stock markets, however, futures markets have been less explored due to the lack of bid-ask quotes.

The demand for accurate estimates of liquidity in futures markets is clear; however the supply of such estimates is scarce. The purpose of the research is to estimate liquidity costs in selected agricultural commodities using Hasbrouck's approach. To our knowledge, no recent studies of liquidity costs across agricultural commodities and exchanges exist. We use 2005 time-stamped price data from actively traded commodities in U.S. major exchanges, i.e. CME (hogs and live cattle), CBOT (corn, soybeans, and wheat), and KCBOT (wheat). We chose these markets because the information contained in prices of active markets is rich, a desirable property when the only source of information to estimate the unobserved variables in the Gibbs sampler are the transaction prices.

Literature review

Considerable research has been performed on market microstructure in general, and on liquidity in particular for stock and financial futures markets. Studies on commodity futures markets, however, are more scarce. Moreover, as argued by Bryant and Haigh (2004), findings from financial markets are not always directly applicable to commodity markets. Our discussion focuses on research performed in commodity futures markets.

Thompson and Waller (1987) studied coffee and cocoa contracts in the New York Board of Trade (NYBOT) over the three-year period 1981-83. Under the hypothesis of negative price correlation, they use the average absolute value of price changes to measure execution costs. Negative price correlation emerges because market makers fill buy orders at a higher price than sell orders. Their findings show lower execution costs in actively traded nearby contracts relative to thinly traded more distant contracts. On average, executions costs are 0.15 (0.12) percent of the contract value in distant coffee (cocoa) contracts, while they reduce to 0.07 percent in nearby coffee and cocoa contracts. Furthermore, they show that negative price correlations cannot be exploited to arbitrage the market, but rather represent the cost for immediate liquidity.

Thompson and Waller (1988) estimated liquidity costs for corn and oats contracts traded in the Chicago Board of Trade (CBOT) in 1984 and 1986. However, their results are mixed and in some cases not consistent with expectations when the Roll measure is used. Thompson, Eales, and Seibold (1993) compared liquidity costs for the same commodity traded in different exchanges, i.e. wheat in the CBOT and in the KCBT in 1985. The data are intra-day prices on a tick basis (i.e. zero price changes are not recorded) from January to June. Using the Roll's measure and the average absolute price changes measure, their results suggest that in Kansas City liquidity costs are significantly higher due, in part, to its lower trading volume. However, there are other variables, which are not identified, that could also help explain this difference. In both exchanges, liquidity costs are higher and more sensitive to trading volume at expiration.

Ma, Peterson, and Sears (1992) investigate the intraday behavior of selected futures contracts, including corn and soybeans. They argue that commonly used spread estimators are usually biased. Specifically, they argue that the Bhattacharya estimator (mean value for all cases where sequential price changes reverse signs) might be downward biased, the Thompson –Waller estimator (average absolute price changes) might be upward biased, and that the Roll estimator might overstate the actual spread when transaction prices are recorded on a tick basis. However, these three estimators show little variation in reflecting the U-shape price behavior during the day which suggests that market-makers require a higher premium in the form of a bid-ask spread to protect themselves from informed traders.

Smith and Whaley (1994) recognize the problems associated with the Roll and Thompson-Waller estimators. They point out that the Roll estimator becomes troublesome when the covariance between adjacent price changes is positive, and that the Thompson-Waller estimate gives an upward bias of the spread because it fails to recognize the variance of true price changes contained in the absolute value of price changes. To overcome this problem, Smith and Whaley suggest a new spread estimator based on the first two moments of absolute price change distribution. Their estimator is derived for tick-basis datasets, and is robust to different levels of serial correlation and volatility of true price changes when both simulated data and S&P500 futures data for the period 1982-1987 are used.

Locke and Venkatesh (1997) compute futures transactions costs for several commodities to assess the performance of commonly used spread estimators. Transaction costs are defined as dollar flows from customers to market-makers, and are estimated as the difference between the average purchase price and the average sale price for all futures customers, with prices weighted by transaction size. Roll's estimates for pork bellies and live cattle in the Chicago Mercantile Exchange (CME) appear to be significantly different to the measured transaction cost. A plausible explanation for such discrepancy is related to the minimum tick size. Commodities with the highest proportion of minimum price changes show the greatest bias between spread estimators and transaction costs. As pointed out by Locke and Venkatesh (1997), the transaction cost per contract may be lower than the minimum tick, and may also not equal to an integer multiple of the minimum tick. This can happen when trades occur between customers with no execution-related transaction costs, or when market-makers adopt pricing and inventory control strategies leading to positive revenue on only a fraction of contracts.

Besides Locke and Venkatesh (1997), only Bryant and Haigh (2004) contrast observed and estimated spreads in commodity futures markets. Observed bid and ask prices, as well as transaction prices, are taken from the LIFFE for cocoa and coffee. In general, absolute price change estimators perform better than serial covariance estimators when evaluated using the bias and the mean square error criteria, but the latter have lower error variances. These findings imply that spread estimators might not be reliable and alternative measures of liquidity costs are needed when observed bid and ask prices are not available, as is the case for most major US exchanges.

As noted, discreteness of futures prices might be an important issue in spread estimation. However, other than Hasbrouck (2004), there are no studies dealing with this in commodity markets. Therefore, here we built on the findings for stocks markets. As explained by Harris (1991), price discreteness is expected to reduce the costs of negotiating in a given market. This is due to the fact that discrete prices represent a smaller set of prices that limits the number of bids and offers that can be made, and therefore negotiations converge more rapidly. In effect, price discreteness minimizes negotiation costs by avoiding extended rounds of bargaining over amounts of diminishing importance. Furthermore, Harris (1994) show that significant reductions in bid-ask spreads may be obtained if traders could use minimum price variations. Hasbrouck (1999) reviews models of stock price discreteness and building on Dravid (1991) proposes a model that allows for asymmetric rounding of the bid and ask quotes. Here we use the model which is explained in the next section.

Methods

The methods used here build on the Roll model and Bayesian estimation to infer the effective bid-ask spread directly from times series of transaction prices. Roll's method is well suited to commodity futures markets not only because the bid-ask spreads are often not recorded, but also because they satisfy the assumptions of informational efficiency and stationarity of price changes more closely than other markets (Laux and Senchack 1992). Furthermore, as identified by Laux and Senchack (1992) failure to capture an asymmetric information component might not be a problem in futures markets because prices are mainly driven by macroeconomic events rather than privately produced firm-specific information. Bayesian estimation is implemented using the Gibbs sampler which is a Markov chain Monte Carlo estimator. This technique is more attractive than the conventional Roll model because i) estimation is based on parameters posteriors which incorporate all the information in the observed transaction prices, ii) it can accommodate latent (unobserved) data, iii) it does not require negative covariance between adjacent price changes, and iv) it allows for a more comprehensive and realistic model that incorporates price discreteness.

In the Roll model, markets are assumed to be efficient. The efficient price is *m(t)* which would hold in the absence of transaction costs and reflects all available public information. Over time the efficient price follows a random walk. However, futures markets operate through dealers who post bid (*b(t)*) and ask (*a(t)*) prices. Buyers buy at the price *a(t),* sellers receive the price *b(t),* and the cost of a transaction is *c*. As a result the prices that we observe when trading takes place are *p(t)* and are modeled as,

$$(1) \qquad m(t)=m(t-1) + u(t) \qquad\qquad u(t) \sim N(0, \sigma^2_u)$$

$$p(t) = m(t) + cq(t)$$

where *m(t)* is the log efficient price, $q(t)=\{-1$ for a sell , $+1$ for a buy$\}$ is the trade direction, *p(t)* is the log transaction price, $p(t) = b(t)$ if $q(t) = -1$, $p(t) = a(t)$ if $q(t) = +1$, and *c* is the half spread.

Conventional approach

Conventionally, the spread model (1) is estimated using moment estimates. Taking price differences (i.e. $\Delta p(t) = u(t)+c\Delta q(t)$) and solving for Var ($\Delta p(t)$) and Cov ($\Delta p(t), \Delta p(t-1)$) yields the following estimators:

$$(2) \qquad \sigma^2_u = \gamma_0 + 2\,\gamma_1$$

$$(3) \qquad c = \sqrt{-\gamma_1}$$

where $\gamma_0 = Var\,(\Delta p(t))$ and $\gamma_1 = \mathrm{Cov}\,(\Delta p(t), \Delta p(t\text{-}1))$.

Bayesian approach

The Bayesian approach is facilitated by the Gibbs sampler. The Gibbs sampler is an algorithm to generate a sequence of samples from the conditional probability distributions of random variables. The algorithm is motivated because it is applicable when the joint distribution is not known but the conditional distribution of each variable is known. As a Markov chain Monte Carlo method, the Gibbs sampler generates sample values from the distribution of each variable in turn, conditional on the current values of the other variables (i.e., $x_1 \sim f(x_1|x_2,x_3,\ldots)$, $x_2 \sim f(x_2|x_1,x_3,\ldots)$, $x_3 \sim f(x_3|x_1,x_2,\ldots)$).

In the Bayesian approach the transaction cost, c, and the variance of the log efficient price changes, σ^2_u, are the unknown parameters from the regression specification

(4) $\quad \Delta p(t) = c\Delta q(t) + u(t)$

$\quad q(t)=\{-1$ for a sell $,\ +1$ for a buy$\} \qquad q(t) \sim$ Bernoulli$(1/2)$.

In the model the joint distribution $F(q,c,\sigma^2_u|p)$ is also unknown and we use the Gibbs sampler to obtain sample values $(q^{(i)},c^{(i)},\sigma^{2\,(i)}_u) \sim F(q,c,\sigma^2_u|p)$ based on known conditional distributions for a known set $p=\{p_1,p_2,\ldots,p_T\}$. This method is implemented, for the vector variable $\boldsymbol{\Theta}=(q,c,\sigma^2_u)$, using a Markov chain to make n random draws which converge in distribution to the joint distribution after a sufficiently large number of iterations. That is,

$(q^{(0)},c^{(0)},\sigma^{2\,(0)}_u),\ (q^{(1)},c^{(1)},\sigma^{2\,(1)}_u),\ldots,\ (q^{(n)},c^{(n)},\sigma^{2\,(n)}_u) \qquad \boldsymbol{\Theta}^{(n)} \sim F^{(n)}(\,q,c,\sigma^2_u|p)$

where $\boldsymbol{\Theta}^{(n)} \to^{D} \boldsymbol{\Theta}$ as $n \to \infty$. The liquidity cost c is then computed as the first moment of the marginal distribution $f(c|p)$. The efficient price is not included because, in the specification $m = p - c \bullet q$, it is redundant.

The conditional prior distribution for c is assumed to be (positive) normal, so the posterior is $c| p \sim N^{+}(\mu_c^{post}, \Omega_c^{post})$, where, N^{+} is the normal density restricted to $[0,+\infty)$, $\mu_c^{post}=Dd$, $\Omega_c^{post}=\sigma^2_u(X'X)^{-1}$, $D^{-1}=X'\sigma^{2\ -1}_u X+(\Omega_c^{prior})^{-1}$, $d=X'\sigma^{2\ -1}_u p+(\Omega_c^{prior})^{-1}\ \mu_c^{prior}$, $X=[\Delta q(t)]$, $\mu_c^{prior}= 0$, and $\Omega_c^{prior}=10^6$. The conditional posterior distribution for σ^2_u is $\sigma^2_u \mid p \sim IG(\alpha^{post}, \beta^{post})$, where $\alpha^{post}= \alpha^{prior}+T/2$, and $\beta^{post}= \beta^{prior}+\Sigma u_t^2/2$, $\alpha^{prior} =\beta^{prior} =10^{-12}$. The direction of the incoming order $q(t)$ is assumed to be a random variable distributed as Bernoulli (1/2) so that buy and sell orders are equally probable.

The implementation of the algorithm is as follows. Begin with an initial (arbitrary) guess of $(q,c,\sigma^2_u)^{(0)}$ and generate n =10,000 draw sequences (we discard the first 2,000 considered as a burning time and keep 8,000 for estimation), where each draw incorporates the most recent information from previous draws and is conditional on the set of observed transaction prices p. Specifically,

1. Draw $c^{(1)}$ from $f(c|\sigma_u^{(0)}, q^{(0)}, p)$
2. Draw $\sigma_u^{(1)}$ from $f(\sigma_u|m^{(0)}, p)$
3. Draw q $^{(1)}$ from $f(q|c^{(1)}, \sigma_u^{(1)}, p)$.

The proposed method can accommodate discreteness, a salient feature of futures prices that arises because bids, asks, and transaction prices are not continuous because their movements are limited by the minimum tick size. It is important that spread models incorporate such movements because both the spread and the tick size may have similar magnitudes. The inclusion of discreteness during estimation helps to draw variables that reflect more accurately the price behavior of the markets under study. Discreteness is incorporated into the model using floor and ceiling functions that round prices to the nearest tick. The model to estimate liquidity costs with discreteness is defined as,

$$(5) \qquad m(t)=m(t-1) + u(t) \qquad\qquad u(t) \sim N(0, \sigma^2_u)$$

$$q(t)=\{-1 \text{ for a sell}, +1 \text{ for a buy}\} \qquad q(t) \sim \text{Bernoulli}(1/2)$$

$$P(t)=\{\text{floor}(M(t) - C,) \text{ for } q(t)=-1, \text{ Ceiling}(M(t) + C,) \text{ for } q(t)=+1\}$$

where $P(t)$ is the observed transaction price in levels, scaled so that the tick size is unity, and $M(t)=e^{m(t)}$ is the efficient price in levels.

Here the joint distribution $F(m,q,C,\sigma^2_U|P)$ is unknown and we use the Gibbs sampler to obtain sample values $(m^{(i)},q^{(i)},C^{(i)},\sigma^2_u{}^{(i)}) \sim F(m,q,C,\sigma^2_U|P)$

1. Draw $C^{(1)}$ from $f(C/\sigma_U^{(0)}, q^{(0)}, P)$
2. Draw $\sigma_u{}^{(1)}$ from $f(\sigma_u/ m^{(0)}, P)$
3. Draw $q^{(1)}$ from $f(q/m^{(0)}, C^{(1)}, \sigma_u^{(1)}, P)$
4. Draw $m^{(1)}$ from $f(m/m^{(0)}, \sigma_u^{(1)}, q^{(1)}, P)$

for $\sigma_U^{(i)}=\underline{P}^2. \sigma_u^{(i)}$. A more exhaustive description of this procedure can be found in Hasbrouck (2004).

Data

Liquidity costs are estimated for six futures contracts traded in three major U.S. exchanges. The contracts are lean hogs (CME), live cattle (CME), corn (CBOT), soybeans (CBOT), and wheat (CBOT and KCBT). We use the *volume by tick* database from the CME, the *futures tick* database from the CBOT, and the *wheat futures tick* database from the KCBT. These datasets provide prices of trades executed during the day in the open auction with their corresponding time stamps. Because the recorded data includes a considerable number of zero price changes for CME and KCBT, our analysis does not have the restrictive assumption of all trades occurring between market-makers and traders. Following Ma, Peterson and Sears (1992) we discard actual bid and ask prices in the *futures tick* database from the CBOT since these observations may be biased because not all bid and ask prices are recorded.

For each commodity we selected the most actively traded contract during a month in 2005 with the highest trading volume for all contracts, so that we do not run into the problem of infrequent trading (Wang, Yau, and Baptiste 1997; Ma, Peterson, and Sears 1992). Preliminary analysis of the data shows that the spread estimators described in the previous section are sensitive to the time period used. Therefore, liquidity costs are estimated using samples of the above contracts that did not display large jumps in prices. Each sample consists of transaction prices showing stationary behavior observed within a week of the specified trading month.

Table 1 summarizes the trading month, contract specifications, and summary descriptive statistics of the data for each commodity. Soybeans show the greatest number of average daily transactions and lowest average time between trades, thus we anticipate this market to have the lowest liquidity costs. On the other hand, wheat traded in the KCBT has the lowest number of trades, the highest average time between trades, and a higher average price than wheat in the CBOT. Liquidity costs are expected to be higher in this market.

Results

Liquidity cost estimates for the six contracts analyzed are presented in table 2. All estimates are computed using scaled transaction prices (i.e. a one-unit change represents one tick size). The table shows the log half-spread, c, and the half-spread in levels in ticks, C. Care must be taken in interpreting the absolute size of the spreads and in making comparisons of liquidity costs across commodities. While we use prices that are stationary, the levels and volatility of the individual series differ as can be seen by their coefficients of variation. The coefficients of variation also are quite small relative to those often calculated over longer time periods, suggesting that the absolute values of the spreads may also be quite small. Comparisons across commodities are further complicated by the fact that examination of the data revealed large differences in the number of zero price changes reported by exchange and commodity. Prices at the CME and KCBT included a high proportion of zero price changes (live cattle—32%; lean hogs—30%; wheat—41%), while the commodities traded at the CBOT (corn—0.56%; soybeans—0.28%; wheat—0.44%) did not. Clearly, these related factors can be expected to influence the absolute and relative transaction costs, and can make it difficult to draw conclusions across markets. Nevertheless, examination of the performance of estimation techniques for each market is valid.

For all six contracts, as anticipated the spreads regardless of the procedures used are small, reflecting the limited volatility of the data. For example, the moment estimate of C based on tick scaled data is 1.34, indicating the half spread is 1.34 times the tick size or in cents/bu, 0.335. Further, moment estimates are large relative to those generated with the Bayesian method, falling outside of two standard deviations of the Bayesian estimates. These findings are similar to those reported by Hasbrouck (2004) for pork bellies and other financial futures. As identified by Hasbrouck, these differences can be explained by the way each estimator is derived. The moment estimate in (3) is derived assuming independent $u(t)$. However, in the Bayesian approach, such independence is assumed but not imposed. When $E[u(t)u(t-1)]$ is negative, the moment estimate is inflated because by construction it does not account for this term. In contrast, for the Bayesian estimate a negative $E[u(t)u(t-1)]$ can be viewed as a small sample effect which is translated into an exact small sample distribution from where the parameters are drawn.

Table 2 also shows that *C* estimates are less than one for all six contracts when the Bayesian approach is used, indicating that cost estimates are less than one tick size. Interestingly, the largest percentage reductions in *C* from use of the Bayesian approach occur in the live cattle and lean hogs markets which had a high proportion of zero price changes. This finding is consistent with the notion that the Bayesian approach uses the information in the zero price changes to reflect when trades occur between customers with no execution-related transaction costs, or when market-makers adopt pricing and inventory control strategies leading to positive revenue on only a fraction of contracts. The best opportunity of a direct comparison across the markets is between wheat in Chicago (C) and Kansas City (K) which has been studied extensively. Here, we find lower transaction costs in Kansas City which is at odds with most of the reported research, but may reflect the differences in zero price changes reported.

The results for the Bayesian model with discreteness as specified in (5) are shown in the last two columns of table 2. As it can be seen, liquidity costs reduce significantly when discreteness is incorporated for all contracts. These findings also are consistent with our expectations of lower negotiation costs associated with price tick movements and with the results reported by Hasbrouck (2004).

Table 3 reports the dispersion of the residual $u(t)$, which provides a measure of volatility generated by the models. Again, moment estimates are lower than two standard deviations of the Bayesian estimates. For lean hogs and corn, the moment estimates are further apart from their corresponding Bayesian estimates. The estimates coming from the model incorporating price discreteness appear to be significantly lower than those coming from the simple Bayesian model.

Liquidity costs are expected to be higher for more volatile markets due to market-makers increased risk premium. For lean hogs this relationship appears to hold as we find highest dispersion (8.29—Bayes, discreteness in table 3) and highest liquidity costs (0.41—Bayes, discreteness in table 2). However, the other markets do not show an exact match between the two parameters. This might be due to other factors affecting liquidity costs and not reflected in the volatility, like specific characteristics of the contracts, the exchanges, or the markets.

Conclusions

Computing liquidity costs in commodity futures markets is not an easy task. Numerous estimators of the bid-ask spread that use transaction prices only have been proposed in the literature to overcome the problem of lack of bid-ask quotes in futures markets. However, the performance of these estimators has been questioned. Here we use a Bayesian estimator with conceptually more desirable properties proposed by Hasbrouck (2004) to compute liquidity costs of six agricultural commodity contracts.

Within each market a consistent pattern emerges in the behavior of the Bayesian methods relative to the conventional Roll spread estimator. The estimates of the spread when using the Bayesian method are systematically smaller than the moment estimate. Incorporation of discreteness lowers the spread estimate even further. This behavior is consistent with Hasbrouck's findings and confirms the notion that the Bayesian method allows for the zero price changes and lower negotiation costs in the presence of discrete price units to be interpreted in an apparently meaningful economic context.

While the Bayesian method performed much as expected, several challenges emerged in performing the analysis which may point the direction for further research. First, the Bayesian method appears to generate spread estimates that are smaller than the moment estimates which previous researchers have found to be downwardly biased. This is somewhat disturbing, but may be a reflection of the more conceptually appropriate framework and the limited number of studies that have been performed with markets where actual bid-ask spreads are available for comparison. Clearly, more work needs to be performed similar to Bryant and Haigh (2004) study comparing the performance of alternative procedures including the Bayesian method in the presence of known bid and ask quotes. Second, the spread estimator computed with the Bayesian method appears to be sensitive to the time period used. For example, we found that the spread estimates for the Bayesian method differed greatly when using different samples from our monthly observations. To some extent, this may be a reflection of the economic conditions and trading activity that varies over time, and would not necessarily be a limitation if the computational burden of generating the Bayesian estimates can be reduced. One method to assess the sensitivity of the procedure would be to simply estimate the Bayesian model for pronounced periods of economic and price behavior. Alternatively, following Hasbrouck (2004) the analysis can be expanded to permit liquidity costs to vary as a function of the volume of the stochastic nature of market transactions. As suggested by O'Hara (1995), spreads may vary with the volume traded so that the spread for large trades may be significantly larger than the small trade spread. This suggests that traded volume which presumably reflects new information in the market should be incorporated into the analysis so that large trade spreads are not confused with large liquidity costs. Unfortunately, appropriate trade volume is not available for all markets, but where available should be included in the analysis. Finally in a related context, further assessment of the Bayesian methods should identify the relationship between stochastic volatility and liquidity costs to capture more realistic market conditions and to further assess the effects of how new information which can cause prices to cluster influence transaction costs.

References

Brorsen, W., D. Buck, and S. Koontz. 1998. "Hedging Hard Red Winter Wheat: Kansas City versus Chicago." *Journal of Futures Markets* 18(4): 449-466

Bryant, H., and Michael Haigh. 2004. "Bid-Ask Spreads in Commodity Futures Markets." *Applied Financial Economics* 14: 923-936

Chu, Q. C., Ding, D. K. and C. S. Pyun 1996. "Bid-ask and Spreads in the Foreign Exchange Market." *Review of Quantitative Finance and Accounting* 6: 19-37

Harris, L. 1991. "Stock Price Clustering and Discreteness." *Review of Financial Studies* 4(3): 389-415

Harris, L. 1994. "Minimum Price Variations, Discrete Bid-Ask Spreads, and Quotation Sizes." *Review of Financial Studies* 7(1): 149-178

Hasbrouck, J. 1999. "The Dynamics of Discrete Bid and Ask Quotes." *Journal of Finance* 54(6):2109-2142

Hasbrouck, J. 2004. "Liquidity in the Futures Pits: Inferring Market Dynamics from Incomplete Data." *Journal of Financial and Quantitative Analysis* 39(2): 305-326

Laux, P., and A.J. Senchack. 1992. "Bid-Ask Spreads in Financial Futures." *Journal of Futures Markets* 12(6): 621-634

Locke, P., and P.C. Venkatesh. 1997. "Futures Markets Transaction Costs." *Journal of Futures Markets* 17(2): 229-245

Ma, C., R. Peterson, and S. Sears. 1992. "Trading Noise, Adverse Selection, and Intraday Bid-Ask Spreads in Futures Markets." *Journal of Futures Markets* 12(5): 519-538

O'Hara, M. 1995. *Market Microstruture Theory*. Blackwell Publishers.

Pennings, J.M.E., and M. Meulenberg. 1997. “Hedging Efficiency: A Futures Exchange Management Approach.” *Journal of Futures Markets* 17(5): 599-615

Roll, R. 1984. “A Simple Implicit Measure of the Effective Bid-ask Spread in an Efficient Market.” *Journal of Finance* 23: 1127-1139

Smith, T., and R. Whaley. 1994. “Estimating the Effective Bid-ask Spread from Time and Sales Data.” *Journal of Futures Markets* 14: 437-456

Stoll, H.R. 1989. “Inferring the Components of the Bid-Ask Spread: Theory and Empirical Tests.” *Journal of Finance* 44: 115-134

Thompson, S., and M. Waller. 1987. “The Execution Cost of Trading in Commodity Futures Markets.” *Food Research Institute Studies* 20(2): 141-163

Thompson, S., and M. Waller. 1988. “Determinants of Liquidity Costs in Commodity Futures Markets.” *Review of Futures Markets* 7: 110-126

Thompson, S., J. Eales, and D. Seibold. 1993. “Comparison of Liquidity Costs between Kansas City and Chicago Wheat Futures Contracts*.” Journal of Agricultural and Resource Economics* 18: 185-197

Wang, H.K., Yau, J., and T. Baptiste. 1997. “Trading Volume and Transaction Costs in Futures Markets.” *Journal of Futures Markets* 17: 757-780

Table 1: Contracts Description and Summary Statistics

Commodity	Lean Hogs	Live Cattle	Corn	Soybeans	Wheat	Wheat
Exchange	CME	CME	CBOT	CBOT	CBOT	KCBT
Trading month	Jul	Jan	Jul	Jul	Aug	Aug
Expiration month	Aug	Feb	Dec	Nov	Dec	Dec
# trading days in sample	5	4	4	4	3	3
Total # of trades	2,023	1,807	4,256	4,601	1,814	1,031
Price units	Cents/lb.	Cents/lb.	Cents/bu	Cents/bu	Cents/bu	Cents/bu
Avg. price	67.57	88.80	250.02	693.58	328.21	351.61
Standard deviation	0.36	0.25	2.88	6.50	1.50	2.12
Coeff. variation (%)	0.53	0.27	1.15	0.93	0.45	0.60
Min price	66.90	88.15	241.00	681.00	325.25	346.50
Max price	68.65	89.45	255.00	711.00	332.50	356.00
Tick	0.025	0.025	0.25	0.25	0.25	0.25
Size of contract	40,000 lb	40,000 lb	5,000 bu	5,000 bu	5,000 bu	5,000 bu
Avg. daily trades	405	452	1064	1150	605	344
Avg. time between trade	34.44 sec	31.48 sec	13.25 sec	12.27 sec	23.24 sec	40.49 sec

Table 2: Liquidity Cost Estimates

		Moment	Bayes [b]		Bayes, discreteness	
		estimate [a]	Post. mean	SD	Post. mean	SD
Live cattle	c x 10^4	4.07	0.59	0.32	0.45	0.12
	C (ticks)	1.45	0.21	0.12	0.16	0.04
	C (cents/lb)	0.036	0.005	0.003	0.004	0.001
Lean hogs	c x 10^4	9.16	2.77	0.39	1.51	0.20
	C (ticks)	2.48	0.75	0.11	0.41	0.05
	C (cents/lb)	0.062	0.019	0.003	0.010	0.001
Corn	c x 10^4	8.16	5.19	0.13	0.32	0.09
	C (ticks)	0.82	0.52	0.01	0.03	0.01
	C (cents/bu)	0.204	0.130	0.003	0.008	0.002
Soybeans	c x 10^4	4.84	3.51	0.08	1.88	0.05
	C (ticks)	1.34	0.97	0.02	0.52	0.01
	C (cents/bu)	0.335	0.243	0.006	0.131	0.004
Wheat (C)	c x 10^4	5.21	4.26	0.10	0.39	0.10
	C (ticks)	0.68	0.56	0.01	0.05	0.01
	C (cents/bu)	0.171	0.140	0.003	0.013	0.003
Wheat (K)	c x 10^4	4.39	3.03	0.31	0.07	0.06
	C (ticks)	0.62	0.43	0.04	0.01	0.01
	C (cents/bu)	0.154	0.106	0.011	0.002	0.002

[a] Moment estimates are higher than two standard deviations of the Bayes estimates in all six contracts.
[b] The Wilcoxon sign test between the simple Bayes and Bayes with discreteness estimates reveals significantly different medians at the 1% level in all six contracts.

Table 3: Model Residual Dispersion (σ_ux10^4)

	Moment	Bayes [b]		Bayes, discreteness	
	estimate [a]	Post. mean	SD	Post. mean	SD
Live cattle	3.73	6.80	0.13	4.40	0.11
Lean hogs	5.40	13.47	0.25	8.29	0.19
Corn	4.65	9.71	0.16	7.23	0.13
Soybeans	4.34	6.22	0.10	5.21	0.08
Wheat (C)	4.78	5.96	0.13	5.21	0.15
Wheat (K)	6.21	7.66	0.26	6.66	0.21

[a] Moment estimates are higher than two standard deviations of the Bayes estimates in all six contracts.
[b] The Wilcoxon sign test between the simple Bayes and Bayes with discreteness estimates reveals significantly different medians at the 1% level in all six contracts.