Measuring the bid-ask spreads: a note on the potential downward bias of the Thompson–Waller estimator

Yoichi Otsubo

Institute for Monetary and Economic Studies, Bank of Japan, Chuo-ku, Tokyo 103-8660, Japan
E-mail: youichi.otsubo@boj.or.jp

The upward bias of the widely used Thompson–Waller estimator has been pointed out in the literature. In contrast, the current article provides a case the estimator would have downward bias: frequent continuous arrivals of orders in the same side associated with a small price change. The upward bias might be cancelled out by downward bias, and the estimator might perform better than the other methods such as Wang–Yau–Baptiste used by the U.S. Commodity Futures Trading Commission. The high-frequency data of the emissions market allows us to provide an empirical evidence.

Keywords: bid-ask spread; carbon emissions market; futures market; market microstructure; Thompson–Waller estimator

JEL Classification: C18; G12; G19

1. Introduction

The difficulty in estimating bid-ask spreads in futures market is that the information on trade direction is usually not provided. One of the popular methods to estimate the spread is the Thompson–Waller (TW) estimator. Applications of the measure are found in Ma et al. (1992) and Bryant and Haigh (2004) among others. The TW estimate equates to the average bid-ask spread if the expected true price change and the variance of true price change are both zero. Under the violation of this condition, it consists of two components, the bid-ask spread and the magnitude of true price changes, and it would be biased upward. While the upward bias due to the change in the true price has been pointed out in the literature, the potential downward bias has not been carefully discussed.

This study shows that the TW estimator would have downward bias when the method is used for a market with frequent continuous arrivals of orders in the same side accompanied by a small price change. The upward bias of the TW estimator would be then cancelled out by the downward bias. The estimator may then perform better than its modified version like Wang–Yau–Baptiste (WYB) (1997) used by the U.S. Commodity Futures Trading Commission.

1 This upward bias in the TW estimator has been pointed out in the literature, for example, Smith and Whaley (1994).
Bid-ask spread estimation bias

(CFTC). The application to the carbon futures data implies that its trading pattern and the price change provide the conditions reducing the bias of the TW estimator.

The remaining part of this article is organized as follows. Section II describes the spread estimation methods and discusses their possible bias. Section III discusses the bias of the spread estimators from the empirical findings and Section IV concludes.

II. Spread Estimation Methods

Spread estimation methods considered in the present study can be categorized into two types: the absolute price change measures and the trade indicator model.

Absolute price change

The methods introduced in this section are variants of average absolute change in the transaction price. The potential bias of the estimators is discussed in the last subsection.

Thompson–Waller. The idea of measuring the bid-ask spread by the average of absolute price change is first applied by Thompson and Waller (1987). The change in transaction price, $\Delta p_t$, can be expressed as

$$\Delta p_t = \frac{S}{2} I_t + \Delta m_t$$

where $S$ is the spread, $\Delta m_t$ is the change in the true price and $I_t$ is the indicator variable, $I_t = 2$ if a buy order follows a sell order, $I_t = -2$ if a sell order follows a buy order and $I_t = 0$ otherwise. Then the TW spread estimate is given by,

$$S_{TW} = \frac{1}{T^+} \sum_{t=1}^{T^+} |\Delta p_t| / T^+$$

$T^+$ is the number of nonzero changes in the transaction prices.

Modified Thompson–Waller. If the trade initiation of the executed transaction, $I_t$, is observable, the TW estimator can be modified as

$$S_{MTW} = \frac{1}{T^0} \sum_{t=0}^{T^0} |\Delta p_t| / T^0$$

where $\Delta p_t$ is the price change that moves from bid to ask (or ask to bid) and $T^0$ is the number of such changes in the transaction prices.

The estimates $S_{TW}$ and $S_{MTW}$ equate to the average bid-ask spread if the expected true price change and the variance of true price change are both zero. Under the violation of this condition, as we can see from Equation 1, they consist of two components, the bid-ask spread and the magnitude of true price changes, and it would be biased upward.

Wang–Yau–Baptiste. Wang et al. (1997) attempt to reduce the bias of the TW estimator by discarding any price change that follows another price change of the same sign. The method is used by the CFTC to estimate the actual bid-ask spread. The WYB estimator is given as

$$S_{WYB} = \frac{1}{T^0} \sum_{t=0}^{T^0} |\Delta p_t| / T^0$$

where $\Delta p_t$ is the price change that moves in a different direction from the previous change and $T^0$ is the number of such changes in the transaction prices.

Bias of the absolute price change estimators. If the assumptions of the true price changes are violated, $S_{TW}$ and $S_{MTW}$ would have upward bias. While the literature has pointed out this bias due to the true price changes, the possibility of the TW estimator having downward bias has not been carefully discussed.

Suppose that a buyer initiated trade follows a buyer initiated trade. In such case, the absolute price change captured by $S_{TW}$ is not the spread, but it is the change in the best ask price. If the two consecutive trades are executed in a short period, it is likely that the change in the lowest ask price is small. When these changes are smaller than the average spread, $S_{TW}$ would be biased downward.

It is most likely that the unrealistic assumptions of the true price changes are violated and therefore TW estimator suffers from upward bias. However, if a market observes many trades in such a manner discussed earlier, the upward bias caused by the variance of the true price changes would be cancelled out by the downward bias.

---

2 This estimator was applied in, for example, Bryant and Haigh (2004), Tse and Zabotina (2004) and Chou and Chung (2006).
To calculate $S_{MTW}$, change in the price between two consecutive buyer (seller) initiated trades are discarded. Thus, the price changes that potentially offset the upward bias of the estimator are filtered out.

The WYB estimator discards the price changes that followed by another change with same direction. It is likely that a positive (negative) price change is caused by the placement of a buy (sell) order. Hence, the price changes discarded to compute $S_{WYB}$ would be similar to those discarded to calculate $S_{MTW}$. Thus, although the WYB estimator discards such price changes in order to reduce the bias, it would have greater bias than $S_{TW}$.

Therefore, although the assumptions that TW estimator is based on are unrealistic and would have bias, it would perform better than the other absolute price change estimators in practice.

**Trade indicator model**

In this section, two models of bid, ask and transaction prices are introduced. Both models describe the trade direction by trade indicator variables.

**Roll.** Roll (1984) assumes an informationally efficient market and assumes that the true price $m_t$ follows a random walk process,

$$m_t = m_{t-1} + u_t$$  

where the $u_t$ are i.i.d. zero-mean random variables with variance $\sigma_u^2$. In a competitive market, traders will set the bid $p_t^b$ and ask $p_t^a$ quotes wide enough to cover their execution cost, $c$. Namely,

$$p_t^b = m_t - c,$$  

$$p_t^a = m_t + c$$

The bid-ask spread is $p_t^a - p_t^b = 2c$, and $c$ can be interpreted as the half-spread. Denoting the trade direction by $x_t$, the transaction price $p_t$ can be represented as,

$$p_t = \begin{cases} 
 p_t^b & \text{if } x_t = -1 \\
 p_t^a & \text{if } x_t = +1 
\end{cases}$$

where trade direction of the incoming order is given by the Bernoulli random variable $x_t \in \{-1, +1\}$. $-1$ indicates a sell order and $+1$ indicates a buy order.

Orders are assumed to arrive with equal probability, serially independent. It is also assumed that the trade direction arrival is independent of the efficient price innovation $u_t$.

The Roll model has two parameters, $c$ and $\sigma_u^2$. These are estimated from the variance and first-order autocovariance of the price changes, $\Delta p_t$. The variance and covariance are

$$Var(\Delta p_t) = 2c^2 + \sigma_u^2$$

$$Cov(\Delta p_t, \Delta p_{t-1}) = -c^2$$

The spread, $S_{Roll}$, is then estimated as

$$S_{t}^{Roll} = 2\sqrt{-Cov(\Delta p_t, \Delta p_{t-1})}$$

and

$$\sigma_u^2 = Var(\Delta p_t) + 2Cov(\Delta p_t, \Delta p_{t-1})$$

**Roll OLS.** From Equations 5 to 7, the transaction price process is,

$$\Delta p_t = m_t + cx_t - (m_{t-1} + cx_{t-1})$$

$$= c\Delta x_t + u_t$$

If one can observe the trade initiations, the above model can be estimated by usual ordinary least squares (OLS) regression. Hence the Roll OLS spread estimate is

$$S_{t}^{ROLS} = 2\hat{c}$$

where $\hat{c}$ is the OLS estimate of Equation 9. Note that the OLS does not require the trade initiation variable to be serially uncorrelated.

**III. Empirical Analysis: Application to the Emissions Market Data**

The largest market for carbon trading is the European Climate Exchange (ECX). The intra-day observation used here are transaction prices for the December 2009 futures contract. The data records each trade price and direction. The sample begins on 2 January 2009 and ends on 14 December 2009.
Figure 1 plots the TW, the modified TW and the WYB estimates.

The WYB and modified TW are consistently wider than the TW. This result is discussed in Section II. The TW seem to be less upward biased. The continuous execution of orders in the same side, which is discarded to calculate $S^{MTW}$ and $S^{WYB}$, reduces the bias of $S^{MTW}$. Table 1 reports the number of such orders.

On average, 65% of the observations used in the TW are the orders continuously traded in the same side. If the price change of continuous orders is significantly smaller than the price change of other cases, TW should have downward bias, which may offset its upward bias. Figure 2 compares the monthly average of the absolute price change of the orders discarded to calculate $S^{MTW}$ and $S^{MTW}$ themselves.

The size of price changes of continuous execution in the same side is remarkably smaller than other case. Notable size of the observations used for the TW is those small price changes. The wider spread of WYB and the modified TW is explained by the elimination of the small price changes. Thus, the gap between the modified TW and the TW is considered as the magnitude of the downward bias of TW.

Table 1. Number of observations used in the spread estimation: Thompson–Waller, modified Thompson–Waller and Wang–Yau–Baptiste estimates

| Month    | $T^+$ | $T'$ | $T''$ | $T^+ - T''$ |
|----------|-------|------|-------|-------------|
| January  | 7073  | 2614 | 2784  | 4459        |
| February | 8746  | 3270 | 3227  | 5476        |
| March    | 10 483| 3752 | 3864  | 6731        |
| April    | 14 988| 5055 | 5825  | 9933        |
| May      | 11 658| 3672 | 4347  | 7986        |
| June     | 13 306| 4480 | 5106  | 8826        |
| July     | 10 186| 3374 | 3507  | 6812        |
| August   | 8051  | 2685 | 2937  | 5366        |
| September| 6493  | 2213 | 2061  | 4280        |
| October  | 6163  | 2083 | 1877  | 4080        |
| November | 5004  | 1831 | 1333  | 3173        |
| December | 1516  | 618  | 408   | 898         |

Notes: $T^+$, $T'$ and $T''$ are the number of observations used for Thompson–Waller estimator (Equation 2), the modified Thompson–Waller estimator (Equation 3) and the Wang–Yau–Baptiste estimator (Equation 4), respectively. $T^+ - T''$ is the number of continuous arrivals of orders in the same side.

Although the modified TW would not suffer from the downward bias, it would suffer from the upward bias due to the change in the true price. Prior to that discussion, we examine the results of the Roll...
estimates. Figure 3 plots the Roll covariance and the Roll OLS.

The Roll model assumes that the trade initiations are serially uncorrelated and are also uncorrelated with changes in the efficient price. We already confirmed that the first assumption is not appropriate with our data, from the observation that the 65% of the order is followed by the same side of the order.

The Roll OLS does not require the serial uncorrelation of the trade directions. Hence, we expect that
the Roll OLS would be a better estimate than the Roll covariance. Furthermore, by analysing the Roll OLS residuals, \( \hat{u}_t \), we examine the factor of upward bias in the absolute price change methods. Figure 4 reports the modified TW spread, the Roll OLS and the standard deviation of its residuals.

The standard deviation of the Roll OLS residuals, \( \sqrt{\text{Var}(\hat{u}_t)} \), is an ad hoc estimate of the size of the change in the true price. The figure indicates that the larger the variances of the changes in the true price, the greater the gap between the estimates. That is an empirical evidence of the upward bias in the modified TW caused by the magnitude of the true price change.

\( S^{TW} \), \( S^{MTW} \) and \( S^{WYB} \) can be upward biased due to the true price change. Our discussion in Section II provides a possible case that the upward bias of \( S^{TW} \) would be cancelled out, while the bias of \( S^{MTW} \) and \( S^{WYB} \) would not. The pattern of the order arrivals, the transaction price change and the true price change suggests that the ECX data are such a case. Thus, although the assumptions for the TW estimator are obviously violated, in practice, the TW estimator may provide better estimate than the other two absolute price change estimators.

**IV. Conclusion**

This study reveals that the TW estimator can have downward bias and may cancel out its upward bias. An application to the carbon futures trading data shows that its trading pattern and the price change provide the conditions reducing the bias. The finding suggests that when a researcher has a data set in hand which does not provide trade directions, the ordinary TW may provide better estimate than WYB and modified TW which meant to augment the TW estimator.

**Acknowledgement**

I would like to thank Bruce Mizrach for providing useful comments and kindly sharing the data.

**References**

Bryant, H. L. and Haigh, M. S. (2004) Bid-ask spreads in commodity futures markets, *Applied Financial Economics*, 14, 923–36. doi:10.1080/0960310042000284669

Chou, R. K. and Chung, H. (2006) Decimalization, trading costs, and information transmission between ETFs and index futures, *The Journal of Futures Markets*, 26, 131–51. doi:10.1002/fut.20189
Ma, C. K., Peterson, R. L. and Sears, R. S. (1992) Trading noise, adverse selection, and intraday bid-ask spreads in futures markets, *The Journal of Futures Markets*, **12**, 519–38. doi:10.1002/fut.3990120504

Roll, R. (1984) A simple implicit measure of the effective bid-ask spread in an efficient market, *The Journal of Finance*, **39**, 1127–39. doi:10.1111/j.1540-6261.1984.tb03897.x

Smith, T. and Whaley, R. E. (1994) Estimating the effective bid/ask spread from time and sales data, *The Journal of Futures Markets*, **14**, 437–55. doi:10.1002/fut.3990140406

Thompson, S. and Waller, M. (1987) The execution cost of trading in commodity futures markets, *Food Research Institute Studies*, **20**, 141–63.

Tse, Y. and Zabotina, T. (2004) Do designated market makers improve liquidity in open-outcry futures markets?, *The Journal of Futures Markets*, **24**, 479–502. doi:10.1002/fut.10117

Wang, G. H., Jot Yau, K. and Baptiste, T. (1997) Trading volume and transaction costs in futures markets, *The Journal of Futures Markets*, **17**, 757–80. doi:10.1002/(SICI)1096-9934(199710)17:7<757::AID-FUT2>3.0.CO;2-M