A New Spread Estimator

Michael Bleaney[*] and Zhiyong Li

University of Nottingham

forthcoming Review of Quantitative Finance and Accounting

2015

Abstract

A new estimator of bid-ask spreads is presented. When the trade direction is known, any estimate of the spread is associated with a unique series of conjectural mid-prices derived by adjusting the observed transaction price by half the estimated spread. It is shown that the covariance of successive conjectural mid-price returns is maximised (or least negative) when the estimated spread is equal to the true spread. A search procedure to maximise this covariance may therefore be used to estimate the true spread. The performance of this estimator under various conditions is examined both theoretically and with Monte Carlo simulations. The simulations confirm the theoretical results. The performance of the estimator is good.

Keywords: Bid-ask Spread, Feedback Trading, Estimation

JEL: G10

[*]Corresponding Author: Professor Michael Bleaney, School of Economics, University of Nottingham, University Park, Nottingham NG7 2RD, England; Tel +44 115 95 15464; Email: michael.bleaney@nottingham.ac.uk

1 Introduction

Bid-ask spreads are important measures of liquidity in financial markets. Developing efficient spread estimators has been a lively area of research for several decades. The development of electronic trading in the major markets has reduced spreads and made them more transparent, since quoted spreads are often recorded. Nevertheless spread estimation is still important for less liquid markets and where "price improvement" occurs, i.e. where the trader obtains a better price for the transaction than that quoted to them.

Bid-ask spread estimators can be divided into two groups: the Roll family of estimators which are based on the serial correlation of transaction returns (Roll 1984, Glosten and Harris 1988, Choi et al. 1988, Stoll 1989, George et al. 1991, Huang and Stoll 1997 and Hasbrouck 2004, 2009), and other estimators such as Lesmond et al. (1999), Holden (2009), Goyenko et al. (2009) and Corwin and Schultz (2012).

One issue that has received little attention is the effect on spread estimators of feedback trading (order flows reacting to price returns). Feedback trading has been empirically recorded in stock markets by Hasbrouck (1991) and Nofsinger and Sias (1999), and in foreign exchange markets by Daníelsson and Love (2006). De Long et al. (1990) provide a theoretical model of feedback trading. Huang and Stoll's (1997) estimator is biased in the presence of feedback trading, as is the entire family of Roll-type estimators (Bleaney and Li 2013). As Daníelsson and Love (2006) point out, feedback trading is very likely to occur in time-aggregated data, and will bias estimates of the effect of order flows on returns. Feedback trading may exist in tick-by-tick data as well.

In this essay we introduce a new trial-and-error method of estimating the spread that has the same data requirements as the HS model (i.e. it uses trade direction as well as return information) but which performs better than the HS model in the presence of feedback trading. When transaction prices and the direction of transactions are known, any estimate of the spread is associated with an estimate of the mid-price for each transaction. The proposed method relies on the fact that, under certain assumptions, the estimated covariance of current and lagged mid-price returns will be maximized (or least negative) when the estimated spread is equal to the true spread. A search to maximize this estimated covariance should therefore yield an accurate estimate of the spread. The assumptions are that mid-price returns are independent of past, current and future order flows. This happens if order flows do not react to returns (no feedback trading), and returns do not react to order flows (no price adjustment by dealers for inventory control or precautionary reasons). Once the estimate of the spread has been obtained, the accuracy of these assumptions can be tested. As with the HS estimator, information on transaction data and order flows is needed. Like the HS estimator, our new estimator is biased in the presence of feedback trading, but substantially less biased than the HS estimator.

When order flows are unknown, one may apply the method introduced in Hasbrouck (2004, 2009) to obtain order flows before using our estimator, as is the case with the HS estimator. The combination of our new estimator and Hasbrouck's method should generate satisfactory estimates.

2 Relevant Literature

Roll (1984) establishes a spread estimation model based on serial correlation of returns. The Roll model considers the order processing cost only. The spread is estimated from the auto-covariance of price returns. Glosten and Harris (1988) introduce a spread estimation model where both the order processing cost and the adverse selection cost are considered. A buy (sell) order will raise (decrease) the midpoint of bid and ask prices. For the model to work, price returns and trade indicators which represent the directions of trades are needed. Choi et al. (1988) incorporate the correlation of order flows into Roll's model. Stoll's (1989) model extends Roll's model by incorporating the probability of price reversal. Stoll's model can also be used to infer the components of the spread. George et al. (1991) relax Roll's assumption of a random walk of mid-prices. When mid-prices contain positive correlation components, the Roll model will underestimate the true spread. Huang and Stoll (1997) develop a general model which incorporates previous spread estimation and decomposition models such as Roll's, Glosten and Harris's (1988) and Stoll's (1989) model. Huang and Stoll's model requires price returns and trade indicators. Hasbrouck (2004, 2009) extends Roll's model by using the Bayesian Gibbs sampler. Hasbrouck's model performs better than Roll's model.

Lesmond et al. (1999) (LOT) develop a spread estimator based on Kyle (1985) and

Glosten and Milgrom (1985). The LOT estimator assumes that informed trading will move the price, and thus lead to a non-zero return, and that other trading will not move the price, and thus lead to a zero return. This model requires data on price returns.

Holden (2009) and Goyenko et al. (2009) develop a spread estimator based on price clustering. The estimator assumes that "price clustering is completely determined by the spread size". The spread is "a probability-weighted average of each possible spread size" divided by the average price. The estimator is called "effective tick". Holden (2009) develops a spread estimator which is a hybrid of "effective tick" and the Huang and Stoll (1997) model.

In Corwin and Schultz (2012), the spread is estimated by daily high-low prices. Corwin and Schultz's model assumes that the price follows a random walk and the highest price of a day is an ask-price (at which a trader makes a buy order to the market maker) and the lowest price of a day is a bid-price (at which a trader makes a sell order to the market maker). This model requires data on price returns and daily high-low prices.

3 A New Estimator

In this section, we introduce a new estimator based on conjectures about the spread. The intuition is simple. We make a conjecture about the spread and calculate conjectural mid-price returns according to the conjecture. The conjectural error influences the covariance between two adjacent conjectural mid-price returns. It will be shown that the series of true mid-price returns has the greatest covariance among all other series of conjectural mid-price returns. Therefore, after trying conjectural spreads and calculating the corresponding conjectural mid-price returns and the covariances, we use the conjectural spreads which correspond to the series with the greatest covariance as the estimate of the true spread.

The bid-ask spread is the difference between the ask price and the bid price. Let s_t be the transaction price which is the ask (bid) price if a buy (sell) order is executed. Transaction prices can be divided into two parts. One is the bid-ask spread and the other is the unobserved mid-price. Formally, the price is given by:

$$s_t = M_t + \frac{SP}{2} \cdot BS_t$$

where M_t is the mid-price. SP is the effective bid-ask spread, and BS is the trade indicator which shows the direction of the trade. $BS = 1$ if there is a buy order and $BS = -1$ if there is a sell order. Then the transaction price return is given by:

$$\Delta s_t = \Delta M_t + \frac{SP}{2}(BS_t - BS_{t-1})$$

where Δ is the first order difference operator. The spread will enlarge (reduce) the observed return when the change of the trade direction has the same (opposite) sign as the mid-price change. If the trade direction does not change ($BS_t - BS_{t-1} = 0$), the observed return is equal to the mid-price change.

We assume that returns are uncorrelated with past, current or future order flows. This means that there is no feedback trading (order flows do not react to past returns), and no adjustment of prices to order flows for inventory control or precautionary reasons. These are the ideal conditions for the estimator. If we have a conjecture about the spread, the error of the conjecture is given by:

$$\Omega_t = SP_t - \widetilde{SP}_t$$

where $\widetilde{SP}_t$ is the conjecture, and Ω_t is the conjectural error. All the symbols with $\sim$ represent conjecture values.

We assume that the spread is fixed throughout the series. Thus the spread and its conjecture and the conjectural error are fixed.

$$\Omega = SP - \widetilde{SP} \tag{1}$$

A conjectural mid-price series ($\widetilde{M}_t$) can be obtained from conjectural spreads.

$$\widetilde{M}_t = s_t - \tfrac{1}{2}\widetilde{SP}BS_t \tag{2}$$

If we re-arrange the equation above, one can find that the difference between the true mid-price and its conjecture is half the conjectural error:

$$\begin{aligned}\widetilde{M}_t &= M_t + \tfrac{1}{2}SPBS_t - \tfrac{1}{2}\widetilde{SP}BS_t \\ &= M_t + \tfrac{1}{2}BS_t(SP - \widetilde{SP}) \\ &= M_t + \tfrac{1}{2}BS_t\Omega\end{aligned} \tag{3}$$

Then the mid-price return is given by:

$$\Delta\widetilde{M}_t = \Delta M_t + \frac{1}{2}\Omega BS_t - \frac{1}{2}\Omega BS_{t-1}$$

We now show that under ideal conditions a trial-and-error method can identify the true spread.

Definition Let A be a set of all conjectures of the true spread $A = \{\widetilde{SP}_1, \widetilde{SP}_2, \cdots, \widetilde{SP}_n\}$

Definition Let B be a set of covariances of two adjacent conjectural mid-price returns obtained according to the conjecture of the true spread $B = \{Cov_1, Cov_2, \cdots, Cov_n\}$, where $Cov_i = Cov[\widetilde{M}(\widetilde{SP}_i)_t, \widetilde{M}(\widetilde{SP}_i)_{t-1}]$.

Proposition 3.1 *If there is no feedback trading, and no inventory control or asymmetric information components of the spread, then the spread and its conjecture, and thus the conjectural error, are serially independent or are fixed. If a conjecture of the spread $\widetilde{SP}_i \in A$ corresponds to $Cov_i = max(B)$, it equals the true spread i.e. $\widetilde{SP}_i = SP$.*

Proof The full proof is in the appendix. The covariance of two adjacent conjectures of mid-price returns is:

$$\begin{aligned} & Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) \\ = & E\{[\Delta\widetilde{M}_t - E(\Delta\widetilde{M}_t)][\Delta\widetilde{M}_{t-1} - E(\Delta\widetilde{M}_{t-1})]\} \end{aligned} \tag{4}$$

Assume the expectation of the conjectural mid-prices is zero. Thus, the equation above can be written as:

$$\begin{aligned} & Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) \\ = & E[\Delta\widetilde{M}_t \cdot \Delta\widetilde{M}_{t-1}] \\ = & E[(\Delta M_t + \tfrac{1}{2}\Omega BS_t - \tfrac{1}{2}\Omega BS_{t-1})(\Delta M_{t-1} + \tfrac{1}{2}\Omega BS_{t-1} - \tfrac{1}{2}\Omega BS_{t-2})] \end{aligned} \tag{5}$$

As shown in the Appendix, the assumptions imply that BS is independent of ΔM at all dates, so many terms in (5) are zeros. The variable BS is a binary variable (1 or -1), thus $E(BS_{t-1}^2) = 1$. Then we can finally obtain:

$$\begin{aligned} & Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) \\ = & Cov(\Delta M_t \cdot \Delta M_{t-1}) + \tfrac{1}{4}\Omega^2 \cdot [2E(BS_t \cdot BS_{t-1}) - E(BS_t \cdot BS_{t-2}) - 1] \end{aligned} \tag{6}$$

The right hand side of the equation is a quadratic polynomial of the expectation of the error of the conjecture. For a given series, the first term on the right hand side is a constant. It is straightforward that when the error is zero (i.e. $\Omega = 0$), the second term is zero. Furthermore, when $\Omega = 0$, there is a global extreme for the right hand side polynomial, symmetrically, the left hand side of the equation $Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1})$ is also at the extreme value:

$$\underset{\Omega}{arg\ max}\, Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) = 0 \tag{7}$$

When the conjectural error is zero, the conjectural spread is the true spread:

$$\Omega = SP - \widetilde{SP}_i = 0 \tag{8}$$

Therefore the conjectural spread which maximises the covariance equals the true spread.

$$\underset{\widetilde{SP}_i \in A}{arg\ max}\, Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) = SP \tag{9}$$

Q.E.D

Proposition (3.1) sheds lights on the spread estimation. The most important point is that at the extreme point, the expectation of conjectural spreads equals that of true spreads. Therefore, one can apply exhaustive search to find the one which is closest to the true spread.

More specifically, the first step of the estimation is to choose a conjectural spread ($\widetilde{SP}_i$). Secondly, calculate the conjectural mid-prices ($\widetilde{M}_i$), and then the conjectural returns ($\Delta\widetilde{M}_{t,i}$) using the conjectural spread according to Equation (2). Thirdly, calculate the covariance of two adjacent returns of conjectural mid-prices ($Cov(\Delta\widetilde{M}_{t,i}, \Delta\widetilde{M}_{t-1,i})_i$). Fourthly, repeat the first three steps enough times to draw a curve of the covariance against the conjectural spread. Finally, find the maximum point of the curve and the conjectural spread corresponding to the maximum point is the estimate of the spread. Because the curve is a negative parabola, there is no need to try all possible values of the spread, instead, one can stop when the shape of a negative parabola appears.

Figure (1) presents the intuition underlying Proposition (3.1) for the simple case where the true mid-price does not change over three periods. It shows the general relationships between the returns of true mid-prices and the returns of conjectural mid-prices, when only transaction prices and the direction of transactions are known. The conjectural spread here is less than the true spread ($\omega > 0$). The conjectural mid-prices are obtained from Equation (2) using transaction prices, given a conjectural spread. The dotted lines are conjectural mid-prices and the solid lines are true mid-prices or transaction prices. A and B represent the ask and bid prices respectively; these are the prices that are actually observed. M and $\widetilde{M}$ represent the true mid-price and the conjecture of it respectively. Δ is the first-order difference operator. Ω is the error of the conjecture which is the difference between the true spread and the conjectural spread, or equivalently, between the true mid-price and the conjectural mid-price.

There is a sell order in period $t-2$, and, thus, the bid price is recorded. There is a buy order in period $t-1$, and thus the ask price is recorded. There is a sell order in period t, and thus the bid price is recorded. Because the conjectural spread is less than the true spread, in periods $t-2$ and t, the conjectural mid-prices are $0.5 \cdot \Omega$ less than the true ones, and in period $t-1$, the conjectural mid-price is $0.5 \cdot \Omega$ greater than the true one.

Between periods $t-2$ and $t-1$, the trade direction changes from selling to buying, so the conjectural error makes the conjectural mid-price return greater than the true mid-price return ($\Delta\widetilde{M}_{t-1} = \Delta M_{t-1} + \Omega = \Omega$).

Between periods $t-1$ and t, the trade direction switches back from buying to selling, so the conjectural error partially cancels the true mid-price return ($\Delta\widetilde{M}_t = \Delta M_t - \Omega = -\Omega$). When the trade direction does not change, the returns of the conjectural mid-prices and of the true mid-prices are the same.

In the case shown in Figure (1), the covariance of two adjacent returns of true mid-prices is zero because the mid-price is fixed. The covariance of two adjacent conjectural mid-price returns is, however, negative. This is because, when the spread is underestimated, the conjectural mid-price returns take on some of the negative serial correlation of the transaction price series induced by the spread (as in the case of Roll's (1984) analysis of the effect of the spread on the serial covariance of transaction price returns, the cases where the trade direction does not switch make no difference).

Now consider the opposite case where the spread is over-estimated (but true mid-price returns are still zero as in Figure 1). Then the conjectural mid-price would be below the true mid-price in periods $t-2$ and t, and above it in period $t-1$, so in this case also the conjectural mid-price series has negative serial covariance that is not present in the true mid-price series.

[Figure 1 near here]

4 Errors of the New Estimator

4.1 Feedback Trading and the New Estimator

In this section, we discuss the impact of feedback trading on the performance of the estimator.

We assume that the mid-price returns can be written as follows:

$$\Delta M_t = \epsilon_t$$

where ϵ_t is a shock which is not influenced by order flows.

If feedback trading exists, order flows are influenced by the shocks. Formally, the covariance between order flows and the shocks is not zero:

$$\begin{aligned} Cov(\epsilon_{t-1} \cdot BS_{t-1}) &= E(\epsilon_{t-1} \cdot BS_{t-1}) \neq 0 \\ Cov(\epsilon_{t-1} \cdot BS_t) &= E(\epsilon_{t-1} \cdot BS_t) \neq 0 \end{aligned} \tag{10}$$

We assume a quote-driven market in which traders receive price quotes, and therefore observe the shock (ϵ_{t-1}) before placing the order (BS_{t-1}). Then we define the covariance between ϵ_{t-1} and BS_{t-1} as current feedback trading (one-period feedback trading). The influence of the shock may persist in the next period, i.e. the shock ϵ_{t-1} may influence the order flow in period t as well. We define the covariance between ϵ_{t-1} and BS_t as lagged feedback trading (two-period feedback trading).

With the existence of feedback trading, the covariance between two adjacent conjectures of mid-price returns is (the detail of the deduction is presented in appendix 9 and 10.1):

$$\begin{aligned} & Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) \\ &= E(\Delta\widetilde{M}_t \cdot \Delta\widetilde{M}_{t-1}) \\ &= E(\epsilon_t \cdot \epsilon_{t-1}) + \Pi_1 \cdot \Omega^2 + \Pi_2 \cdot \Omega \end{aligned} \tag{11}$$

where

$$\begin{aligned} \Pi_1 &= \tfrac{1}{4} E\left[2BS_{t-1} \cdot BS_{t-2} - BS_t \cdot BS_{t-2} - 1\right] \\ \Pi_2 &= \tfrac{1}{2} E\left[-BS_{t-1} \cdot \epsilon_{t-1} + BS_t \cdot \epsilon_{t-1}\right] \end{aligned} \tag{12}$$

Equation (11) suggests that when there is feedback trading, the covariance of the conjectural mid-price returns ($Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1})$) contains a linear term in the conjectural errors (Ω) as well as a quadratic one, so the value of Ω which maximises the polynomial is no longer zero.In other words, the estimator is biased.

We now discuss possible errors that arise if we still estimate the spread by maximising the covariance between two adjacent conjectures of mid-price returns $Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1})$, as suggested in the previous section. The estimate is given by:

$$\widehat{SP} = \left[\frac{2\Pi_1 SP + \Pi_2}{2\Pi_1}\right] \tag{13}$$

where $\widehat{SP}$ is obtained when Ω maximises $Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1})$, and is the estimate of the true spread.

In the presence of feedback trading, the estimated spread should be (the detail of the deduction is presented in appendix 10.1):

$$\begin{aligned} \widehat{SP} &= SP - 4\Pi_2 \\ &= SP + E(\epsilon_{t-1} \cdot BS_{t-1}) - E(\epsilon_{t-1} \cdot BS_t) \end{aligned} \tag{14}$$

Equation (14) suggests the estimator will overestimate the spread if there is positive feedback trading and vice versa. It is of interest that, unlike the other estimators, the total influence of feedback trading on the estimator includes two period-feedback trading ($E(\epsilon_{t-1} \cdot BS_t)$). If order flows do not exhibit serial autocorrelation, which may be because the influence of the mid-price shocks is persistent, there is only current feedback trading ($E(\epsilon_{t-1} \cdot BS_{t-1})$). One may define the difference between current and lagged feedback trading net feedback trading. Because the signs of current feedback trading and lagged feedback trading in Equation (14) are different, the positive autocorrelation of order flows can reduce the influence of feedback trading and vice versa. Because of hot-potato trading, order flows, especially in the tick-by-tick case, tend to be positively autocorrelated. Thus, the influence of net feedback trading on our estimator is not as big as on the others.

When there is no feedback trading ($\Pi_2 = 0$), and no autocorrelated order flows ($E(BS_{t-1} \cdot BS_{t-2}) = 0$ and $\Pi_1 = -1$), the equation above becomes:

$$\widehat{SP} = SP \tag{15}$$

and thus,

$$\Omega = 0$$

The equations above suggest that under the ideal conditions, Equation (11) reduce to the simple version of the estimator and in this circumstance, the estimator is unbiased.

Consider now the impact of feedback trading in the HS model. The HS model is given by:

$$\Delta s_t = \tfrac{SP}{2} BS_t - \beta \tfrac{SP}{2} BS_{t-1} + \epsilon_t$$

When there is feedback trading, the estimated value from the HS model $\widehat{\frac{SP}{2}}$ is as follows:

$$\frac{\widehat{SP}}{2} = \frac{SP}{2} + E(BS_t, \epsilon_t) \tag{16}$$

Then, the HS model error will be

$$Error = 2 \cdot E(BS_t, \epsilon_t) \tag{17}$$

Equation (17) suggests that when there is positive feedback trading, the HS estimator overestimates the true spread and vice versa. In particular, when there is only current feedback trading, the bias in the HS estimator is twice as large as that of our new estimator.

4.2 Price impact and the New Estimator

In this section, we discuss the impact of inventory holding costs (IC) and asymmetric information costs (AS) on the performance of the estimator. Under these conditions, the HS estimator performs well, because it explicitly incorporates these effects.

When there are IC&AS components in the spread, the mid-price return is given by:

$$\Delta M_t = \frac{1}{2} \varrho SP BS_{t-1} + \epsilon_t$$

where mid-price returns are influenced by the past order flow (Evans and Lyons 2002), ϱ is the fraction of the components of the spread and ϵ_t is a shock which is not influenced by order flows. And the conjecture of the mid-price return is given by:

$$\begin{aligned}\Delta \widetilde{M}_t &= \Delta M_t + \tfrac{1}{2}\Omega BS_t - \tfrac{1}{2}\Omega BS_{t-1} \\ &= \tfrac{1}{2}(\varrho - 1)\Omega BS_{t-1} + \tfrac{1}{2}\Omega BS_t + \tfrac{1}{2}\varrho(SP - \Omega) BS_{t-1} + \epsilon_t\end{aligned}$$

Then the covariance between two adjacent conjectures of mid-price returns is (the detail of the deduction is presented in appendix 9 and 10.2):

$$\begin{aligned} & Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) \\ &= E(\Delta\widetilde{M}_t \cdot \Delta\widetilde{M}_{t-1}) \\ &= E(\epsilon_t \cdot \epsilon_{t-1}) + \Pi_1 \cdot \Omega^2 + \varrho\Pi_3 \cdot (SP - \Omega) \cdot \Omega \\ & +\tfrac{1}{4}\varrho^2 (BS_{t-1} \cdot BS_{t-2}) \cdot [(SP - \Omega)^2] \end{aligned} \tag{18}$$

where

$$\begin{aligned} \Pi_1 &= \tfrac{1}{4}E\left[(\varrho-1)^2(BS_{t-1} \cdot BS_{t-2}) + (\varrho-1)BS_t \cdot BS_{t-2} + (\varrho-1) + (BS_{t-1} \cdot BS_{t-2})\right] \\ \Pi_3 &= \tfrac{1}{4}E[1 + 2(\varrho-1)(BS_{t-1} \cdot BS_{t-2}) + BS_t \cdot BS_{t-2}] \end{aligned} \tag{19}$$

Equation (18) suggests that when there are IC&AS components, the covariance of the conjectural mid-price returns ($Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1})$) is no longer a function of conjectural errors (Ω) only but also a function of the true spread (SP).

Unlike the simple version of the estimator, the right hand side of Equation (43) is quadratic in $(SP - \Omega)$ and Ω instead of Ω only. The covariance of adjacent conjectural mid-price returns will not be maximised when $\Omega = 0$ In other words, the estimator will be biased.

We now discuss possible errors if we still let $(SP - \Omega)$ which maximises the covariance between two adjacent conjectures of mid-price returns $Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1})$ be the estimate of the true spread. Thus, the estimate is given by:

$$\widehat{SP} = -\left[\frac{-2\Pi_1 SP + \Pi_3 SP\varrho}{2\left(\Pi_1 + \frac{1}{4}E(BS_{t-1} \cdot BS_{t-2})\varrho^2 - \Pi_3\varrho\right)}\right] \tag{20}$$

where $\widehat{SP}$ is the value of $(SP - \Omega)$ which maximises $Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1})$, and is the estimate of the true spread.

When there are IC&AS components of the spread. The estimated spread should be (the detail of the deduction is presented in appendix 10.2):

$$\widehat{SP} = \left(1 - \frac{\varrho}{2}\right) \cdot SP \tag{21}$$

Equation (21) suggests that when the transaction cost is not the only component of the spread, the estimator will underestimate the true spread. In the simulation section, an adjustment will be introduced to overcome this issue.

When there are no IC&AS components of the spread ($\varrho = 0$) and no autocorrelated order flows ($E(BS_{t-1} \cdot BS_{t-2}) = 0$ and $\Pi_1 = -1$ and $\Pi_3 = 1$), the equation above becomes:

$$\widehat{SP} = SP \tag{22}$$

and thus,

$$\Omega = 0$$

The equations above suggest that under the ideal conditions, Equation (20) reduces to the simple version of the estimator and in this circumstance, the estimator is unbiased.

It can be shown that the estimator will not be influenced by the autocorrelation of order flows.

5 Simulation Experiments

In this section, simulated data are used to examine the performance of the new estimator. The aim of this section is to assess the effects of the following factors on the performance of the estimators: mid-price changes caused by order flows and lagged feedback trading. The performance of the Huang and Stoll model is also presented for comparison. Bleaney and Li (2013) evaluate the performance of the Roll, HS, Corwin and Schultz and Hasbrouck estimators under various conditions. The paper suggests that the HS model outperforms others in higher frequencies and, although the Corwin and Schultz estimator has the lowest standard deviation, it significantly under-/overestimates the spread in higher/lower frequencies and does not exhibit consistency across the sampling frequencies. Goyenko et al. (2009) and Corwin and Schultz (2012) show that the Hasbrouck and Corwin and Schultz estimators are better than the LOT estimator. When order flow information is available, the HS model is the best choice. The Holden (2009) estimator cannot be evaluate by simulation experiments. Therefore, in this paper, we only compare the performances of the new estimator and the HS estimator.

There are 500 replications simulated for each case. There are 432000 periods in a replication. Let one period represent one minute, and there is one trade per minute. Thus there are 300 trading days (1440 minutes and 1440 trades per day). For each replication, data are considered in various sampling periods: tick-by-tick, five minutes, fifteen minutes, one hour, four hours, twelve hours and 24 hours. Thus, there are eight subgroups for each replication. For five-minute intervals, only every fifth trade is used and the intervening trades discarded, and similarly for longer intervals.

Each replication includes data on order flows, bid-ask spreads, mid-prices, and translation prices. Data are generated according to the following system. An order has two possible values 1 and -1. Order flows are either random or positively correlated with current and (possibly) lagged mid-price returns (the feedback-trading case). Formally, the order flow series is given by

$$\begin{aligned} BS_t &= \psi F(\Delta M_t + \eta \Delta M_{t-1}) + (1-\psi)\varpi_t \\ \psi &= 0 \; or \; 1 \end{aligned} \tag{23}$$

where BS_t is the order flow, which is either random ($\psi = 0$) or a function $F(\Delta M_{t-1} + \eta \Delta M_t)$ is a function of the past mid-price returns ($\psi = 1$), which suggests the existence of feedback trading. η describes the existence of lagged feedback trading. For example, $\eta = 0.5$ suggests that lagged feedback trading is 50% weaker than current feedback trading. ϖ_t is a binomial random variable, which follows a binomial distribution with one trial and 50% probability i.e. $B(1, 0.5)$. It suggests that order flows are drawn from a binomial distribution randomly and both the buy and sell orders carry the same weight in the series. The function $F(\cdot)$ reflects the following relationship between order flows and past mid-price returns.

$$BS_t \sim \begin{cases} B(1,\kappa) & if \quad \Delta M_t > 0 \\ B(1, 1-\kappa) & if \quad \Delta M_t < 0 \end{cases} \tag{24}$$

where $B(1,\kappa)$ is a binomial distribution with one trial and κ probability. When $\kappa = 0.5$, there is no feedback trading, and when $\kappa > 0.5$, there is positive feedback trading and vice versa.

Mid-price returns are generated using the following equation,

$$\Delta M_t = \varrho \, BS_{t-1} \cdot \tfrac{SP_t}{2} + \epsilon_t \tag{25}$$

where ϵ_t follows a normal distribution with zero mean and standard deviation σ; SP_t is the bid-ask spread which is assumed fixed. ϱ is the fraction of the spread that is caused by inventory control and asymmetric information, and thus $(1-\varrho)$ represents the order-processing part of the spread. When $\varrho = 0$, the order-processing part is the only component of the spread, and mid-price follow a random walk process.

Transaction prices are generated by

$$s_t = M_t + \frac{SP_t}{2} \cdot BS_t \tag{26}$$

where s_t is the transaction price.

5.1 Ideal Conditions

In this section, the ideal case for the estimators is considered, where order flows are random; mid-prices follow a random walk process and the spread is fixed. Under these circumstances, both the basic and adjusted estimators are unbiased. Formally, the standard deviations of mid-price returns is $\sigma = 0.0002$, which is similar to that observed for major currencies in foreign exchange markets. Let $\psi = 0$ in Equation (23), which suggests that order flows are random. Let $\varrho = 0$ in Equation (25), which suggests that the mid-price follows a random walk process and the spread is fixed at 0.0003. The system is given by,

$$\begin{aligned}
&BS_t = \varpi_t \\
&\varpi_t \sim B(1, 0.5) \\
&\Delta M_t = \epsilon_t \\
&\epsilon_t \sim N(0, 4 \times 10^{-8}) \\
&SP_t = 0.0003 \\
&s_t = M_t + \tfrac{SP_t}{2} \cdot BS_t
\end{aligned} \tag{27}$$

Five hundred replications, each of which has 432000 periods, are generated according to the system above. Each replication has eight subgroups according to various sampling periods.

Transaction returns and order flows are used for estimations. The standard deviation of mid-price returns is also calculated. Thus, for every subgroup, there are 500 estimated spreads for each estimator and 500 standard deviations of mid-price returns.

The results are presented in Table (1). The first column shows the results when every transaction is used (tick-by-tick data). The other columns show the results when the transactions are sampled at increasingly long intervals, from five minutes to 24 hours. There are four panels which report the summary statistics and the results of the estimators respectively. The rows in each panel are as follows. *Midstd* reports the average of the standard deviations of mid- price returns over the relevant interval. *Estimates* indicates the average of estimated spreads, and *Relative Estimates* shows the ratio of this to the true spread. *Est-Std* reports the standard deviations of the estimated spreads. *RMSE* is the root mean square error, or the standard deviation of the estimates about the true spread, so it incorporates the effect of bias as well as the standard deviation of the estimates about their own mean. It is the best indicator of the likely error in an estimate of the spread from an individual series.

The row of *Midstd* shows the time interval and the standard deviation of mid-price returns have a positive relationship, as a result of the random walk in returns. In the tick-by-tick case, the average standard deviation of mid-price returns is 2×10^{-4} which is the same as the setting of the system. In the 24-hour case, the standard deviation is 7.58×10^{-3}. Thus the ratio of the spread to the standard deviation varies from 1.5 in the tick-by-tick case to 0.0396 at 24 hours. Spreads are harder to estimate when this ratio is smaller, and hence the standard deviation and *RMSE* increase with the time interval.

It can be seen from Table 1 that both the new estimator and the HS estimator are highly accurate in tick-by-tick data, with an RMSE of less than 0.25% of the true spread of 0.0003. As the sampling frequency falls, the RMSE rises quickly, exceeding 12% of the true spread in one-hour samples. The performance of the two estimators is extremely similar.

[Table 1 near here]

5.2 One-Period Feedback Trading

In this section, most settings are the same as the ones in section (5.1) except that now order flows are assumed to be influenced by the latest mid-price returns. Thus all the differences of the performance of the estimators can be imputed to the existence of feedback trading. Let $\psi = 1$ and $\eta = 0$, which suggests that there is only current feedback trading. Under these circumstances, all the estimators are biased. However, the new estimator should have the least error and the HS estimator should have the greatest error. Formally, let $\psi = 1$ and $\eta = 0$ in Equation (23), which suggests that order flows affected by the past period mid-price returns. Let $\kappa = 0.65$, which implies that there is positive net feedback trading. As in the previous case, the spread is fixed at 0.0003. The system is given by,

$$
\begin{aligned}
&BS_t \sim \begin{cases} B(1, 0.65) & if \;\; \Delta M_t > 0 \\ B(1, 0.35) & if \;\; \Delta M_t < 0 \end{cases} \\
&\Delta M_t = \epsilon_t \\
&\epsilon_t \sim N(0, 9 \times 10^{-12}) \\
&SP_t = 0.0003 \\
&s_t = M_t + \frac{SP_t}{2} \cdot BS_t
\end{aligned}
\tag{28}
$$

The results are shown in Table 2. The bias in the estimators can be seen in the relative estimate for the tick-by-tick case (a bias of +16% for the new estimator, and +32% for the HS estimator). Since the standard deviation of the estimates is very small at short time intervals, the RMSE is dominated by the bias in these cases (up to 30-minute intervals). The bias is slightly larger at longer time intervals for both estimators. Clearly, however, the new estimator outperforms the HS estimator in the presence of feedback trading, as predicted by our earlier analysis.

[Table 2 near here]

5.3 Two-Period Feedback Trading

In this section, most settings are the same as the ones in section (5.2) except that now e add lagged feedback trading. Let $\psi = 1$ and $\eta = 0.5$ which implies that there is both current and lagged feedback trading, with lagged feedback trading in the same direction as but 50% weaker than current trading. Net feedback trading is the summation of them: $\Delta M_t + 0.5\,\Delta M_{t-1}$. In this section, order flows are random; order flows are influenced by the past mid-price returns; and the spread is fixed. Under these circumstances, all the estimators are biased. However, the new estimator should have the least error and the HS estimator should have the greatest error. Formally, let $\psi = 1$ and $\eta = 0.5$ in Equation (23), which suggests that order flows affected by past two periods mid-price returns. Let $\kappa = 0.65$, which suggest there is positive net feedback trading.

The spread is still fixed at 0.0003. The system is given by,

$$\begin{aligned} & BS_t \sim \begin{cases} B(1,0.65) & if \quad \Delta M_t + 0.5\,\Delta M_{t-1} > 0 \\ B(1,0.35) & if \quad \Delta M_t + 0.5\,\Delta M_{t-1} < 0 \end{cases} \\ & \Delta M_t = \epsilon_t \\ & \epsilon_t \sim N(0, 9 \times 10^{-12}) \\ & SP_t = 0.0003 \\ & s_t = M_t + \frac{SP_t}{2} \cdot BS_t \end{aligned} \tag{29}$$

The results are presented in Table 3. The HS estimator performs slightly better than it did in Table 2 (current feedback trading only) in the tick-by-tick case, with a bias of +28.7% compared with +32% in Table 2. However it performs worse than in Table 2 at any longer time interval (for example at five minutes the bias for the HS estimator rises to +42.7%, compared with +32% in Table 2). The new estimator, by contrast, performs even better than in Table 2 in the tick-by-tick case, because of the offsetting effect of current and lagged feedback trading shown in Equation (14). The bias of the new estimator is only +7.3% in the tick-by-tick case, compared with +16% in Table 2. At longer time intervals the new estimator, like the HS estimator, performs worse in Table 3 than in Table 2, but its bias is substantially less at all time intervals than that of the HS estimator.

[Table 3 near here]

5.4 Inventory Control and Asymmetric Information Components

In this section, most settings are the same as the ones in section (5.1) except that the mid-price return is now assumed to be influenced by the past order flow, and thus there are inventory control and the asymmetric information components to the spread. Order flow is assumed to be random, so there is no feedback trading. Let $\varrho = \frac{1}{3}$, which suggests that the inventory control and asymmetric information parts contribute one third of the total spread. Under these circumstances, the new estimator is biased, but the HS estimator is unbiased. Formally, let $\psi = 0$ in Equation (23), which suggests that order flows are random.The spread is 0.0003, as before. The system is given by:

$$\begin{aligned}
BS_t &= \varpi_t \\
\varpi_t &\sim B(1, 0.5) \\
\Delta M_t &= \tfrac{1}{3} BS_{t-1} \cdot \tfrac{SP_t}{2} + \epsilon_t \\
\Delta M_t &= \tfrac{2}{3} BS_{t-1} \cdot \tfrac{SP_t}{2} + \epsilon_t \\
\epsilon_t &\sim N(0, 4 \times 10^{-8}) \\
SP_t &= 0.0003 \\
s_t &= M_t + \tfrac{SP_t}{2} \cdot BS_t
\end{aligned} \tag{30}$$

The results are presented in Table (4), in which the second row (ϱ) reports the coefficient of Equation (25) and represents the proportion of the IC & AS components of the spread.

The standard deviation of mid-price returns is slightly greater than in the previous cases (2.06×10^{-4} for tick-by-tick data, rising to 7.80×10^{-3} for 24-hour intervals). Thus the ratio of the spread to the standard deviation ranges from 1.46 to 0.0385. The estimated ϱ is close to $\frac{1}{3}$, which is the same as the setting, when the time interval is short. When the time interval is longer than one hour, ϱ becomes unstable, because in relatively long runs the microstructure effects are weaker. While the HS estimator remains accurate, the new estimator underestimates the spread by 16.7%, or half of ϱ, as predicted in our earlier theoretical discussion.

[Table 4 near here]

5.5 Both Feedback Trading and Price Impact

In this section, we investigate the performance of the two estimators in the presence of *both* feedback trading (which favours the new estimator) and inventory control and asymmetric information components of the spread (which favours the HS estimator). We assume two-period feedback trading, as in section (5.3), and we investigate two separate settings for ϱ: one-third, as in section (5.4), and a larger one of two-thirds.

Thus $\psi = 1$, $\eta = 0.5$ and $\varrho = \frac{1}{3}$ or $\frac{2}{3}$. The system is given by:

$$\begin{aligned}
BS_t &\sim \begin{cases} B(1, 0.65) & if \quad \Delta M_t + 0.5\,\Delta M_{t-1} > 0 \\ B(1, 0.35) & if \quad \Delta M_t + 0.5\,\Delta M_{t-1} < 0 \end{cases} \\
\Delta M_t &= \tfrac{1}{3}\, BS_{t-1} \cdot \tfrac{SP_t}{2} + \epsilon_t \quad in\ table\ 5 \\
\Delta M_t &= \tfrac{2}{3}\, BS_{t-1} \cdot \tfrac{SP_t}{2} + \epsilon_t \quad in\ table\ 6 \\
\epsilon_t &\sim N(0, 9 \times 10^{-12}) \\
SP_t &= 0.0003 \\
s_t &= M_t + \tfrac{SP_t}{2} \cdot BS_t
\end{aligned} \tag{31}$$

Table (5) shows the results for $\varrho = \frac{1}{3}$. As in the case of two-period feedback trading alone (Table 3), the HS estimator overestimates the spread considerably: by 28.7% in tick-by-tick data and by rather more in time-aggregated data. In fact the numbers for the HS estimator are very similar to those in Table (3); the price impact makes virtually no difference. For the new estimator the picture is very different. The underestimation associated with price impact offsets the overestimation caused by feedback trading. In tick-by-tick data the new estimator underestimates by 9%, but overestimates slightly in time-aggregated data (by about 5% up to four hours, and by quite a bit more at longer intervals).

When $\varrho = \frac{2}{3}$, the simulation results are as show in Table (6). The HS results are very close to those shown in Table (5). For the new estimator, the higher value of ϱ reduces the estimates, as expected. In tick-by-tick data the new estimator now underestimates by 25.7%, and by 12% in five-minute data and by 10% in four-hour data, only overestimating at longer intervals.

[Table 5 near here] [Table 6 near here]

6 Discussion

Our new spread estimator, based on a trial-and-error procedure, was shown to perform almost as well as the HS estimator in ideal conditions of no price impact or feedback trading (Table 1). A little-recognised weakness of the HS estimator is that it is prone to overestimate the spread in the presence of (positive) feedback trading. Our new estimator also overestimates the spread in the presence of feedback trading, but by considerably less than the HS estimator does. With only current feedback trading, the overestimation bias of the new estimator is only half that of the HS estimator in tick-by-tick data (Table 2). With lagged feedback trading as well, the bias in the new estimator is even smaller, both absolutely and relative to the HS estimator (Table 3).

In the presence of inventory control and asymmetric information components of the spread, the HS estimator remains unbiased, because these elements are built into the HS estimation procedure (Table 4). The new estimator, however, underestimates to the tune of half of ϱ, where ϱ is the proportion of the spread attributable to price impact. When both feedback trading and price impact effects are present (Tables 5 and 6), the new estimator benefits from the offsetting effects of the tendency to overestimate in the feedback trading case and to underestimate in the price impact case. It therefore tends to outperform the HS estimator, which performs as poorly in this case as in the pure feedback trading case.

7 Conclusions

We have proposed a new bid-ask spread estimator based on the principle that the covariance of successive mid-price returns tends to be maximised at the true value of the spread. A grid search or trial-and-error procedure for maximising this covariance over alternative conjectures about the spread may therefore be used to estimate the true spread. The information requirements are the same as for Huang and Stoll's (1997) estimator: transaction prices and trade direction. Theoretically it was shown that the new estimator overestimates the spread in the presence of positive feedback trading (a rise in price making a buy order more likely), but by considerably less than the Huang-Stoll estimator. Price impact causes the new estimator to underestimate the spread, with a bias equal to half the proportion of the spread represented by price impact. Simulation results confirm the theoretical findings. Simulation results for the combination of feedback trading and price impact show that the bias effects identified in the separate cases are approximately additive. This means that the Huang-Stoll estimator performs as poorly in the combined case as in the pure feedback trading case, whereas the new estimator tends to perform better in the combined case than in the pure price impact case, because the two biases offset one another (assuming that feedback trading is positive).

For practical purposes, a sensible approach would be to use the Huang-Stoll estimator initially, and then to estimate the degree of feedback trading using the estimated mid-price returns implied by the HS-estimated spread. If feedback trading is significant, use our new spread estimator in preference to the HS estimator.

Appendix

8 Proof of Proposition (3.1)

Definition Let A be a set of all conjectures of the true spread $A = \{\widetilde{SP}_1, \widetilde{SP}_2, \cdots, \widetilde{SP}_n\}$

Definition Let B be a set of covariances of two adjacent conjectural mid-price returns obtained according to the conjecture of the true spread $B = \{Cov_1, Cov_2, \cdots, Cov_n\}$, where $Cov_i = Cov[\widetilde{M}(\widetilde{SP}_i)_t, \widetilde{M}(\widetilde{SP}_i)_{t-1}]$.

One can find that sets A and B are one to one mapping.

Proposition (3.1): If there is no feedback trading, and no inventory control or asymmetric information components of the spread, then the spread and its conjecture, and thus the conjectural error, are serially independent or are fixed. If a conjecture of the spread $\widetilde{SP}_i \in A$ corresponds to $Cov_i = max(B)$, it equals the true spread i.e. $\widetilde{SP}_i = SP$.

Proof The covariance of two adjacent conjectures of mid-price returns is:

$$\begin{aligned} & Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) \\ = & E\{[\Delta\widetilde{M}_t - E(\Delta\widetilde{M}_t)][\Delta\widetilde{M}_{t-1} - E(\Delta\widetilde{M}_{t-1})]\} \end{aligned} \tag{32}$$

Assume the conjectural errors are fixed, expectations of errors are given by:

$$E(\Omega_t) = E(\Omega_{t-1}) = E(\Omega_{t-2}) = \Omega$$

and the exceptions of the multiplication of the conjectural errors are given by:

$$E(\Omega_t\Omega_{t-1}) = E(\Omega_t\Omega_{t-2}) = E(\Omega_{t-1}\Omega_{t-2}) = \Omega^2$$

Furthermore, assume the expectation of the conjectural mid-prices is zero. Thus, the covariance can be written as:

$$\begin{aligned}
&Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) \\
=& E[\Delta\widetilde{M}_t \cdot \Delta\widetilde{M}_{t-1}] \\
=& E[(\Delta M_t + \tfrac{1}{2}\Omega BS_t - \tfrac{1}{2}\Omega BS_{t-1})(\Delta M_{t-1} + \tfrac{1}{2}\Omega BS_{t-1} - \tfrac{1}{2}\Omega BS_{t-2})] \\
=& E[(\Delta M_t + \tfrac{1}{2}\Omega BS_t - \tfrac{1}{2}\Omega BS_{t-1})\Delta M_{t-1} \\
&+\tfrac{1}{2}\Omega BS_{t-1}(\Delta M_t + \tfrac{1}{2}\Omega BS_t - \tfrac{1}{2}\Omega BS_{t-1}) \\
&-\tfrac{1}{2}\Omega BS_{t-2}(\Delta M_t + \tfrac{1}{2}\Omega BS_t - \tfrac{1}{2}\Omega BS_{t-1})] \\
=& E[(\Delta M_t\Delta M_{t-1} + \tfrac{1}{2}\Omega BS_t\Delta M_{t-1} - \tfrac{1}{2}\Omega BS_{t-1}\Delta M_{t-1}) \\
&+(\Delta M_t\tfrac{1}{2}\Omega BS_{t-1} + \tfrac{1}{2}\Omega BS_t\tfrac{1}{2}\Omega BS_{t-1} - \tfrac{1}{2}\Omega BS_{t-1}\tfrac{1}{2}\Omega BS_{t-1}) \\
&-(\Delta M_t\tfrac{1}{2}\Omega BS_{t-2} + \tfrac{1}{2}\Omega BS_t\tfrac{1}{2}\Omega BS_{t-2} - \tfrac{1}{2}\Omega BS_{t-1}\tfrac{1}{2}\Omega BS_{t-2})]
\end{aligned} \tag{33}$$

Re-arrange the equation further, we have:

$$\begin{aligned}
=& E[(\Delta M_t\Delta M_{t-1} + \tfrac{1}{2}\Omega BS_t\Delta M_{t-1} - \tfrac{1}{2}\Omega BS_{t-1}\Delta M_{t-1}) \\
&+(\tfrac{1}{2}\Delta M_t\Omega BS_{t-1} + \tfrac{1}{4}\Omega^2 BS_t BS_{t-1} - \tfrac{1}{4}\Omega^2 {BS_{t-1}}^2) \\
&-(\Delta M_t\tfrac{1}{2}\Omega BS_{t-2} + \tfrac{1}{4}\Omega^2 BS_t BS_{t-2} - \tfrac{1}{4}\Omega^2 BS_{t-1}BS_{t-2})] \\
=& E(\Delta M_t\Delta M_{t-1}) \\
&+E(\tfrac{1}{2}\Omega BS_t\Delta M_{t-1} - \tfrac{1}{2}\Omega BS_{t-1}\Delta M_{t-1} + \tfrac{1}{2}\Delta M_t\Omega BS_{t-1} - \Delta M_t\tfrac{1}{2}\Omega BS_{t-2} \\
&+\tfrac{1}{4}\Omega^2 BS_t BS_{t-1} - \tfrac{1}{4}\Omega^2 {BS_{t-1}}^2 - \tfrac{1}{4}\Omega^2 BS_t BS_{t-2} + \tfrac{1}{4}\Omega^2 BS_{t-1}BS_{t-2})
\end{aligned} \tag{34}$$

Because the variable BS is a binary variable (1 or -1), then:

$$E(BS_{t-1}^2) = 1$$

Furthermore, because we assume there is no feedback trading, then:

$$E(BS_t\Delta M_{t-1}) = 0$$
$$E(BS_{t-1}\Delta M_{t-1}) = 0$$

Because we assume there is no IC&AS components, then:

$$E(BS_{t-1}\Delta M_t) = 0$$
$$E(BS_{t-2} \cdot \Delta M_t) = 0$$

Equation (34) can be written as:

$$\begin{aligned}
&Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) \\
=& E(\Delta M_t\Delta M_{t-1}) \\
&+E(\tfrac{1}{4}\Omega^2 BS_t BS_{t-1} - \tfrac{1}{4}\Omega^2 - \tfrac{1}{4}\Omega^2 BS_t BS_{t-2} + \tfrac{1}{4}\Omega^2 BS_{t-1}BS_{t-2}) \\
=& Cov(\Delta M_t \cdot \Delta M_{t-1}) + \tfrac{1}{4}\Omega^2 \cdot [2E(BS_t \cdot BS_{t-1}) - E(BS_t \cdot BS_{t-2}) - 1]
\end{aligned} \tag{35}$$

The right hand side of the equation is a quadratic polynomial of the expectation of the error of the conjecture. For a given series, the first term on the right hand side is a constant. It is straightforward that when the expectation of the error is zero (i.e. $E(\Omega) = 0$), the second term is zero. Furthermore, when $E(\Omega) = 0$, there is a global extreme for the right hand side polynomial, symmetrically, the left hand side of the equation $Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1})$ is also at the extreme value:

$$\underset{\Omega}{arg\ max}\, Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) = 0 \tag{36}$$

When the conjectural error is zero, the conjectural spread is the true spread:

$$\Omega = SP - \widetilde{SP}_i = 0 \tag{37}$$

Therefore the conjectural spread which maximises the covariance equals the true spread.

$$\underset{\widetilde{SP}_i \in A}{arg\ max}\, Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) = SP \tag{38}$$

Q.E.D

9 Simplify equations (11) and (18)

This section shows the detail of the simplification of equations (11) and (18). Feedback trading, inventory holding costs and asymmetric information costs are considered together.

Considering the inventory control and asymmetric information components of the spread, the true mid-price returns are given by:

$$\Delta M_t = \frac{1}{2}\varrho SPBS_{t-1} + \epsilon_t$$

The covariance of the two adjacent conjectural mid-price returns is given by,

$$\begin{aligned} &Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) \\ &= E\{[\Delta\widetilde{M}_t - E(\Delta\widetilde{M}_t)][\Delta\widetilde{M}_{t-1} - E(\Delta\widetilde{M}_{t-1})]\} \end{aligned} \tag{39}$$

Assume the expectation of conjectural mid-price returns to be zero. Thus the equation above can be written as:

$$
\begin{aligned}
& Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) \\
= {} & E[\Delta\widetilde{M}_t \cdot \Delta\widetilde{M}_{t-1}] \\
= {} & E\Big\{\Big[\tfrac{1}{2}(\varrho-1)\Omega BS_{t-1} + \tfrac{1}{2}\Omega BS_t + \tfrac{1}{2}\varrho\widetilde{SP}_{t-1}BS_{t-1} + \epsilon_t\Big] \\
& \cdot \Big[\tfrac{1}{2}(\varrho-1)\Omega BS_{t-2} + \tfrac{1}{2}\Omega BS_{t-1} + \tfrac{1}{2}\varrho\widetilde{SP}_{t-2}BS_{t-2} + \epsilon_{t-1}\Big]\Big\} \\
= {} & E\Big\{\tfrac{1}{2}(-1+\varrho)\Omega BS_{t-2}\Big[\tfrac{1}{2}(-1+\varrho)\Omega BS_{t-1} + \epsilon_t + \tfrac{1}{2}\Omega BS_t + \tfrac{1}{2}\varrho\widetilde{SP}_{t-1}BS_{t-1}\Big] \\
& +\epsilon_{t-1}\Big[\tfrac{1}{2}(-1+\varrho)\Omega BS_{t-1} + \epsilon_t + \tfrac{1}{2}\Omega BS_t + \tfrac{1}{2}\varrho\widetilde{SP}_{t-1}BS_{t-1}\Big] \\
& +\tfrac{1}{2}\Omega BS_{t-1}\Big[\tfrac{1}{2}(-1+\varrho)\Omega BS_{t-1} + \epsilon_t + \tfrac{1}{2}\Omega BS_t + \tfrac{1}{2}\varrho\widetilde{SP}_{t-1}BS_{t-1}\Big] \\
& +\tfrac{1}{2}\varrho\widetilde{SP}_{t-2}BS_{t-2}\Big[\tfrac{1}{2}(-1+\varrho)\Omega BS_{t-1} + \epsilon_t + \tfrac{1}{2}\Omega BS_t + \tfrac{1}{2}\varrho\widetilde{SP}_{t-1}BS_{t-1}\Big]\Big\} \\
= {} & E\Big\{\Big[\tfrac{1}{2}\cdot\tfrac{1}{2}(-1+\varrho)\Omega BS_{t-2}(-1+\varrho)\Omega BS_{t-1} + \tfrac{1}{2}(-1+\varrho)\Omega BS_{t-2}\epsilon_t \\
& +\tfrac{1}{2}\cdot\tfrac{1}{2}(-1+\varrho)\Omega BS_{t-2}\Omega BS_t + \tfrac{1}{2}\cdot\tfrac{1}{2}(-1+\varrho)\Omega BS_{t-2}\varrho\widetilde{SP}_{t-1}BS_{t-1}\Big] \\
& + \Big[\tfrac{1}{2}(-1+\varrho)\Omega BS_{t-1}\epsilon_{t-1} + \epsilon_t\epsilon_{t-1} + \tfrac{1}{2}\Omega BS_t\epsilon_{t-1} + \tfrac{1}{2}\varrho\widetilde{SP}_{t-1}BS_{t-1}\epsilon_{t-1}\Big] \\
& + \Big[\tfrac{1}{2}\cdot\tfrac{1}{2}\Omega BS_{t-1}(-1+\varrho)\Omega BS_{t-1} + \tfrac{1}{2}\Omega BS_{t-1}\epsilon_t + \tfrac{1}{2}\cdot\tfrac{1}{2}\Omega BS_{t-1}\Omega BS_t \\
& +\tfrac{1}{2}\cdot\tfrac{1}{2}\Omega BS_{t-1}\varrho\widetilde{SP}_{t-1}BS_{t-1}\Big] + \Big[\tfrac{1}{2}\cdot\tfrac{1}{2}\varrho\widetilde{SP}_{t-2}BS_{t-2}(-1+\varrho)\Omega BS_{t-1} \\
& +\tfrac{1}{2}\varrho\widetilde{SP}_{t-2}BS_{t-2}\epsilon_t + \tfrac{1}{2}\cdot\tfrac{1}{2}\varrho\widetilde{SP}_{t-2}BS_{t-2}\Omega BS_t + \tfrac{1}{2}\cdot\tfrac{1}{2}\varrho\widetilde{SP}_{t-2}BS_{t-2}\varrho\widetilde{SP}_{t-1}BS_{t-1}\Big]\Big\}
\end{aligned} \tag{40}
$$

Re-arrange the equation further, we have:

$$
\begin{aligned}
& Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) \\
= & E\Big\{\Big[\tfrac{1}{4}(-1+\varrho)^2\Omega^2 BS_{t-1}BS_{t-2} + \tfrac{1}{2}(-1+\varrho)\Omega BS_{t-2}\epsilon_t \\
& +\tfrac{1}{4}(-1+\varrho)\Omega^2 BS_t BS_{t-2} + \tfrac{1}{4}\varrho(-1+\varrho)\widetilde{SP}BS_{t-1}BS_{t-2}\Omega\Big] \\
& + \Big[\tfrac{1}{2}(-1+\varrho)\Omega BS_{t-1}\epsilon_{t-1} + \epsilon_t\epsilon_{t-1} + \tfrac{1}{2}\Omega BS_t\epsilon_{t-1} + \tfrac{1}{2}\varrho\widetilde{SP}BS_{t-1}\epsilon_{t-1}\Big] \\
& + \Big[\tfrac{1}{4}(-1+\varrho)\Omega^2 {BS_{t-1}}^2 + \tfrac{1}{2}\Omega BS_{t-1}\epsilon_t + \tfrac{1}{4}\Omega^2 BS_t BS_{t-1} \\
& +\tfrac{1}{4}\varrho\widetilde{SP}{BS_{t-1}}^2\Omega\Big] + \Big[\tfrac{1}{4}\varrho(-1+\varrho)\Omega BS_{t-1}BS_{t-2}\widetilde{SP} + \tfrac{1}{2}\varrho\widetilde{SP}BS_{t-2}\epsilon_t\Big] \\
& +\tfrac{1}{4}\varrho\widetilde{SP}\Omega BS_t BS_{t-2} + \tfrac{1}{4}\varrho^2\widetilde{SP}^2 BS_{t-2}BS_{t-1}\Big]\Big\} \\
= & E(\epsilon_t\epsilon_{t-1}) + E\Big[\tfrac{1}{4}(-1+\varrho)^2\Omega^2 BS_{t-1}BS_{t-2} + \tfrac{1}{4}(-1+\varrho)\Omega^2 BS_t BS_{t-2} \\
& +\tfrac{1}{4}(-1+\varrho)\Omega^2 {BS_{t-1}}^2 + \tfrac{1}{4}\Omega^2 BS_t BS_{t-1} + \tfrac{1}{2}(-1+\varrho)\Omega BS_{t-2}\epsilon_t \\
& +\tfrac{1}{2}(-1+\varrho)\Omega BS_{t-1}\epsilon_{t-1} + \tfrac{1}{2}\Omega BS_t\epsilon_{t-1} + \tfrac{1}{2}\Omega BS_{t-1}\epsilon_t \\
& +\tfrac{1}{4}\varrho(-1+\varrho)\widetilde{SP}BS_{t-1}BS_{t-2}\Omega + \tfrac{1}{4}\varrho\widetilde{SP}{BS_{t-1}}^2\Omega \\
& +\tfrac{1}{4}\varrho(-1+\varrho)\Omega BS_{t-1}BS_{t-2}\widetilde{SP} + \tfrac{1}{4}\varrho\widetilde{SP}\Omega BS_t BS_{t-2} \\
& +\tfrac{1}{2}\varrho\widetilde{SP}BS_{t-1}\epsilon_{t-1} + \tfrac{1}{2}\varrho\widetilde{SP}BS_{t-2}\epsilon_t + \tfrac{1}{4}\varrho^2\widetilde{SP}^2 BS_{t-2}BS_{t-1}\Big] \\
= & E(\epsilon_t\epsilon_{t-1}) + E\Big\{\Big[\tfrac{1}{4}(-1+\varrho)^2 BS_{t-1}BS_{t-2} + \tfrac{1}{4}(-1+\varrho)BS_t BS_{t-2} \\
& +\tfrac{1}{4}(-1+\varrho){BS_{t-1}}^2 + \tfrac{1}{4}BS_t BS_{t-1}\Big]\Omega^2 + \Big[\tfrac{1}{2}(-1+\varrho)BS_{t-2}\epsilon_t \\
& +\tfrac{1}{2}(-1+\varrho)BS_{t-1}\epsilon_{t-1} + \tfrac{1}{2}BS_t\epsilon_{t-1} + \tfrac{1}{2}BS_{t-1}\epsilon_t\Big]\Omega \\
& + \Big[\tfrac{1}{4}(-1+\varrho)BS_{t-1}BS_{t-2} + \tfrac{1}{4}{BS_{t-1}}^2 \\
& +\tfrac{1}{4}(-1+\varrho)BS_{t-1}BS_{t-2} + \tfrac{1}{4}BS_t BS_{t-2}\Big]\varrho\widetilde{SP}\Omega \\
& + \Big[\tfrac{1}{2}BS_{t-1}\epsilon_{t-1} + \tfrac{1}{2}BS_{t-2}\epsilon_t\Big]\varrho\widetilde{SP} \\
& +\tfrac{1}{4}\varrho^2\widetilde{SP}^2 BS_{t-2}BS_{t-1}\Big\}
\end{aligned}
\tag{41}
$$

Let following symbols to represent some parts of the equation above, because we assume order flows do not influence the mid-price shocks following parts are zeros.

$$E(BS_{t-2}\epsilon_t) = E(BS_{t-1}\epsilon_t) = 0$$

Because the variable BS is a binary variable and with a mean of zero, then:

$$E(BS_{t-1}^2) = 1$$

$$
\begin{aligned}
\Pi_0 &= E(BS_{t-1}BS_{t-2}) \\
\Pi_1 &= E\left[\tfrac{1}{4}(-1+\varrho)^2 BS_{t-1}BS_{t-2} + \tfrac{1}{4}(-1+\varrho)BS_tBS_{t-2} + \tfrac{1}{4}(-1+\varrho) + \tfrac{1}{4}BS_tBS_{t-1}\right] \\
&= \tfrac{1}{4}E\left[(-1+\varrho)^2\Pi_0 + (-1+\varrho)BS_tBS_{t-2} + (-1+\varrho) + \Pi_0\right] \\
\Pi_2 &= E\left[\tfrac{1}{2}(-1+\varrho)BS_{t-2}\epsilon_t + \tfrac{1}{2}(-1+\varrho)BS_{t-1}\epsilon_{t-1} + \tfrac{1}{2}BS_t\epsilon_{t-1} + \tfrac{1}{2}BS_{t-1}\epsilon_t\right] \\
&= \tfrac{1}{2}E\left[(-1+\varrho)BS_{t-1}\epsilon_{t-1} + BS_t\epsilon_{t-1}\right] \\
\Pi_3 &= E\left[\tfrac{1}{4}(-1+\varrho)BS_{t-1}BS_{t-2} + \tfrac{1}{4}{BS_{t-1}}^2 + \tfrac{1}{4}(-1+\varrho)BS_{t-1}BS_{t-2} + \tfrac{1}{4}BS_tBS_{t-2}\right] \\
&= \tfrac{1}{4}E[2(-1+\varrho)\Pi_0 + 1 + BS_tBS_{t-2}] \\
\Pi_4 &= E(\tfrac{1}{2}BS_{t-1}\epsilon_{t-1} + \tfrac{1}{2}BS_{t-2}\epsilon_t) \\
&= \tfrac{1}{2}E(BS_{t-1}\epsilon_{t-1})
\end{aligned} \tag{42}
$$

Substitute above equations into Equation (41), we have:

$$
\begin{aligned}
& Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) \\
= & E(\epsilon_t\epsilon_{t-1}) + \Pi_1\Omega^2 + \Pi_2\Omega + \Pi_3\varrho\widetilde{SP}\Omega \\
& + \Pi_4\varrho\widetilde{SP} + \tfrac{1}{4}\varrho^2\widetilde{SP}^2\Pi_0
\end{aligned} \tag{43}
$$

Equation (43) suggests that when there are IC&AS components and feedback trading, the covariance of the conjectural mid-price returns ($Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1})$) is no longer a function of conjectural errors (Ω) only but also a function of the conjecture of the spread ($\widetilde{SP}$). Furthermore, because the true spread (SP) is certain for a given series, the conjectural errors (Ω) is a function of the conjecture of the spread. To investigate the relationship between the true spread and the conjecture of it, we re-arrange the equation to make $\widetilde{SP}$ be the only variable of the equation. Replace the conjectural error (Ω) by the true spread (SP) and the conjectural spread ($\widetilde{SP}$), we have:

$$
\begin{aligned}
& Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) \\
= & E(\epsilon_t\epsilon_{t-1}) + \Pi_1(SP-\widetilde{SP})^2 + \Pi_2\left(SP-\widetilde{SP}\right) \\
& + \Pi_3\left[\varrho\widetilde{SP}(SP-\widetilde{SP})\right] + \Pi_4\varrho\widetilde{SP} + \tfrac{1}{4}\Pi_0\varrho^2\widetilde{SP}^2 \\
= & E(\epsilon_t\epsilon_{t-1}) + \Pi_1SP^2 - 2\Pi_1\widetilde{SP}SP + \Pi_1\widetilde{SP}^2 + \Pi_2SP \\
& - \Pi_2\widetilde{SP} + \Pi_3\varrho\widetilde{SP}SP - \Pi_3\varrho\widetilde{SP}^2 + \Pi_4\varrho\widetilde{SP} + \tfrac{1}{4}\varrho^2\widetilde{SP}^2\Pi_0 \\
= & E(\epsilon_t\epsilon_{t-1}) + \Pi_1SP^2 + \Pi_2SP - 2\Pi_1\widetilde{SP}SP - \Pi_2\widetilde{SP} \\
& + \Pi_3\varrho\widetilde{SP}SP + \Pi_4\varrho\widetilde{SP} + \Pi_1\widetilde{SP}^2 + \tfrac{1}{4}\varrho^2\widetilde{SP}^2\Pi_0 - \Pi_3\varrho\widetilde{SP}^2 \\
= & (\Pi_1 + \tfrac{1}{4}\varrho^2\Pi_0 - \varrho\Pi_3)\cdot\widetilde{SP}^2 + [\varrho\Pi_3\cdot SP + \varrho\Pi_4 - 2\Pi_1\cdot SP - \Pi_2] \\
& \cdot\widetilde{SP} + E(\epsilon_t\cdot\epsilon_{t-1}) + \Pi_1\cdot SP^2 + \Pi_2\cdot SP
\end{aligned} \tag{44}
$$

We now discuss possible errors that if we still let $(SP - \Omega)$ which maximises the covariance between two adjacent conjectures of mid-price returns $Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1})$ to be the estimate of the true spread. Thus, the estimate is given by:

$$\widehat{SP} = -\left[\frac{-2\Pi_1 SP - \Pi_2 + \Pi_3 SP\varrho + \Pi_4\varrho}{2\left(\Pi_1 + \frac{1}{4}\Pi_0\varrho^2 - \Pi_3\varrho\right)}\right] \tag{45}$$

where $\widehat{SP}$ is the value of $(SP - \Omega)$ which maximises $Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1})$, and is the estimate of the true spread. When there are no IC&AS components of the spread ($\varrho = 0$), no feedback trading ($\Pi_2 = \Pi_4 = 0$), and no autocorrelated order flows ($\Pi_o = 0$ and $\Pi_1 = -1$ and $\Pi_3 = 1$), the equation above becomes:

$$\widehat{SP} = SP \tag{46}$$

and thus,

$$\Omega = 0$$

The equations above suggest that under the ideal conditions, Equation (45) reduce to the simple version of the estimator and in this circumstance, the estimator is unbiased.

10 Errors of the estimator

10.1 Feedback trading

Assume there is feedback trading and there are no inventory control and asymmetric information components of the spread, thus $\varrho = 0$, then Equation (43) becomes,

$$\begin{aligned} & Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) \\ &= E(\Delta\widetilde{M}_t \cdot \Delta\widetilde{M}_{t-1}) \\ &= E(\epsilon_t \cdot \epsilon_{t-1}) + \Pi_1 \cdot \Omega^2 + \Pi_2 \cdot \Omega \end{aligned} \tag{47}$$

where

$$\begin{aligned} \Pi_0 &= E(BS_{t-1} \cdot BS_{t-2}) \\ \Pi_1 &= \tfrac{1}{4} E\left[\Pi_0 - BS_t \cdot BS_{t-2} - 1 + \Pi_0\right] \\ \Pi_2 &= \tfrac{1}{2} E\left[-BS_{t-1} \cdot \epsilon_{t-1} + BS_t \cdot \epsilon_{t-1}\right] \end{aligned} \tag{48}$$

Because the covariance of order flows are usually very small compare to 1, it is safe to take approximation that let $\Pi_1 = -\frac{1}{4}$. From (48) into Equation (45), it becomes:

$$\begin{aligned} &\widehat{SP} \\ &= -\frac{-2\Pi_1 SP - \Pi_2}{2(\Pi_1)} \\ &= SP + \frac{\Pi_2}{2\Pi_1} \\ &= SP + \frac{\frac{1}{2}E(BS_t\epsilon_{t-1} - BS_{t-1}\epsilon_{t-1})}{-2\cdot\frac{1}{4}} \\ &= SP + E(BS_{t-1}\epsilon_{t-1}) - E(BS_t\epsilon_{t-1}) \end{aligned} \tag{49}$$

10.2 Inventory Control and Asymmetric Information Components

Assume there are inventory control and asymmetric information components of the spread and there is no feedback trading, then Equation (43) becomes,

$$\begin{aligned} &Cov(\Delta\widetilde{M}_t, \Delta\widetilde{M}_{t-1}) \\ &= E(\Delta\widetilde{M}_t \cdot \Delta\widetilde{M}_{t-1}) \\ &= E(\epsilon_t \cdot \epsilon_{t-1}) + \Pi_1 \cdot \Omega^2 + \varrho\Pi_3 \cdot (SP - \Omega) \cdot \Omega \\ &\quad + \frac{1}{4}\varrho^2\Pi_0 \cdot [(SP - \Omega)^2] \end{aligned} \tag{50}$$

where

$$\begin{aligned} \Pi_0 &= E(BS_{t-1} \cdot BS_{t-2}) \\ \Pi_1 &= \frac{1}{4}E\left[(\varrho - 1)^2\Pi_0 + (\varrho - 1)BS_t \cdot BS_{t-2} + (\varrho - 1) + \Pi_0\right] \\ \Pi_3 &= \frac{1}{4}E[1 + 2(\varrho - 1)\Pi_0 + BS_t \cdot BS_{t-2}] \end{aligned} \tag{51}$$

And Equation (45) becomes:

$$\widehat{SP} = -\frac{-2\Pi_1 SP + \Pi_3 SP\varrho}{2(\Pi_1 + \frac{1}{4}\Pi_0\varrho^2 - \Pi_3\varrho)} = \frac{(2\Pi_1 - \Pi_3\varrho)}{2(\Pi_1 + \frac{1}{4}\Pi_0\varrho^2 - \Pi_3\varrho)} \cdot SP \tag{52}$$

The numerator of Equation (52) can be simplified:

$$
\begin{aligned}
&(2\Pi_1 - \Pi_3\varrho) \\
&= 2(-1+\varrho)^2\Pi_0 + 2(-1+\varrho)BS_tBS_{t-2} \\
&\quad +2(-1+\varrho) + 2\Pi_0 - 2(-1+\varrho)\Pi_0\varrho - \varrho - BS_tBS_{t-2}\varrho \\
&= 2(-1+\varrho)^2\Pi_0 + 2\Pi_0 - 2(-1+\varrho)\Pi_0\varrho \\
&\quad +2(-1+\varrho)BS_tBS_{t-2} - BS_tBS_{t-2}\varrho \\
&\quad +2(-1+\varrho) - \varrho \\
&= 2\Pi_0\varrho^2 - 4\Pi_0\varrho + 4\Pi_0 - 2\Pi_0\varrho^2 + 2\Pi_0\varrho \\
&\quad +2\varrho BS_tBS_{t-2} - 2BS_tBS_{t-2} - BS_tBS_{t-2}\varrho \\
&\quad -2+\varrho \\
&= 4\Pi_0 - 2\Pi_0\varrho + \varrho BS_tBS_{t-2} - 2BS_tBS_{t-2} - 2 + \varrho \\
&= 4BS_{t-1}BS_{t-2} - 2BS_tBS_{t-2} - 2 + (BS_tBS_{t-2} + 1 - 2BS_{t-1}BS_{t-2})\varrho
\end{aligned} \tag{53}
$$

The dominator of Equation (52) can be simplified:

$$
\begin{aligned}
&2(\Pi_1 + \Pi_0\varrho^2 - \Pi_3\varrho) \\
&= 2([(-1+\varrho)^2\Pi_0 + (-1+\varrho)BS_tBS_{t-2} + (-1+\varrho) + \Pi_0] \\
&\quad +\Pi_0\varrho^2 - [2(-1+\varrho)\Pi_0 + 1 + BS_tBS_{t-2}]\varrho) \\
&= 2(\Pi_0\varrho^2 - 2\Pi_0\varrho + \Pi_0 + (-1+\varrho)BS_tBS_{t-2} - 1 + \varrho + \Pi_0 \\
&\quad +\Pi_0\varrho^2 + 2(1-\varrho)\Pi_0\varrho - \varrho - BS_tBS_{t-2}\varrho) \\
&= 2(2BS_{t-1}BS_{t-2} - BS_tBS_{t-2} - 1) \\
&\quad +2(\Pi_1 + \Pi_0\varrho^2 - \Pi_3\varrho) \\
&= 2([(-1+\varrho)^2\Pi_0 + (-1+\varrho)BS_tBS_{t-2} + (-1+\varrho) + \Pi_0] \\
&\quad +\Pi_0\varrho^2 - [2(-1+\varrho)\Pi_0 + 1 + BS_tBS_{t-2}]\varrho) \\
&= 2(\Pi_0\varrho^2 - 2\Pi_0\varrho + \Pi_0 + (-1+\varrho)BS_tBS_{t-2} - 1 + \varrho + \Pi_0 \\
&\quad +\Pi_0\varrho^2 + 2(1-\varrho)\Pi_0\varrho - \varrho - BS_tBS_{t-2}\varrho) \\
&= 2(2BS_{t-1}BS_{t-2} - BS_tBS_{t-2} - 1)
\end{aligned} \tag{54}
$$

Substitute the results of the simplification back to Equation (52), the equation becomes,

$$
\begin{aligned}
&\widehat{SP} \\
&= \tfrac{(2\Pi_1 - \Pi_3\varrho)}{2(\Pi_1 + \frac{1}{4}\Pi_0\varrho^2 - \Pi_3\varrho)} \cdot SP \\
&= \tfrac{3BS_{t-1}BS_{t-2} - 2BS_tBS_{t-2} - 2 + (BS_tBS_{t-2} + 1 - 2BS_{t-1}BS_{t-2})\varrho}{(4BS_{t-1}BS_{t-2} - 2BS_tBS_{t-2} - 2)} \cdot SP \\
&= \left[1 + \tfrac{(BS_tBS_{t-2} + 1 - 2BS_{t-1}BS_{t-2})\varrho}{(4BS_{t-1}BS_{t-2} - 2BS_tBS_{t-2} - 2)}\right] \cdot SP \\
&= \left(1 - \tfrac{1}{2}\varrho\right) \cdot SP
\end{aligned}
\tag{55}
$$

References

Bleaney, M. and Z. Li (2013). The performance of bid-ask spread estimators under less than ideal conditions. *University of Nottingham School of Economics Discussion Papers 13/05*. 1, 12

Choi, J. Y., D. Salandro, and K. Shastri (1988). On the estimation of bid-ask spreads: Theory and evidence. *The Journal of Financial and Quantitative Analysis 23*(2), 219–230. 1, 2

Corwin, S. A. and P. Schultz, P. (2012). A simple way to estimate bid-ask spreads from daily high and low prices. *The Journal of Finance 67*(2), 719–760. 1, 3, 12

Daníelsson, J. and R. Love (2006). Feedback trading. *International Journal of Finance & Economics 11*(1), 35–53. 1

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann (1990). Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance 45*(2), 379–395. 1

Evans, M. D. D. and R. K. Lyons (2002). Order flow and exchange rate dynamics. *The Journal of Political Economy 110*(1), 170–180. 10

George, T. J., G. Kaul, and M. Nimalendran (1991). Estimation of the bid-ask spread and its components: A new approach. *The Review of Financial Studies 4*(4), 623–656. 1, 2

Glosten, L. and P. Milgrom (1985). Bid ask and transaction prices in a specialist market with heterogeneously informed trades. *Journal of Financial Economics 14*(1), 71–100. 3

Glosten, L. R. and L. E. Harris (1988). Estimating the components of the bid/ask spread. *Journal of Financial Economics 21*(1), 123–142. 1, 2

Goyenko, R. Y., C. W. Holden, and C. A. Trzcinka (2009). Do liquidity measures measure liquidity? *Journal of Financial Economics 92*(2), 153 – 181. 1, 3, 12

Hasbrouck, J. (1991). Measuring the information content of stock trades. *The Journal of Finance 46*(1), 179–207. 1

Hasbrouck, J. (2004). Liquidity in the futures pits: Inferring market dynamics from incomplete data. *The Journal of Financial and Quantitative Analysis 39*(2), 305–326. 1, 2

Hasbrouck, J. (2009). Trading costs and returns for u.s. equities: Estimating effective costs from daily data. *The Journal of Finance 64*(3), 1445–1477. 1, 2

Holden, C. W. (2009). New low-frequency spread measures. *Journal of Financial Markets 12*(4), 778 – 813. 1, 3, 12

Huang, R. D. and H. R. Stoll (1997). The components of the bid-ask spread: A general approach. *The Review of Financial Studies 10*(4), 995–1034. 1, 2, 3, 20

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica 53*(6), 1315–1335. 2

Lesmond, D. A., J. P. Ogden, and C. A. Trzcinka (1999). A new estimate of transaction costs. *The Review of Financial Studies 12*(5), pp. 1113–1141. 1, 2

Nofsinger, J. R. and R. W. Sias (1999). Herding and feedback trading by institutional and individual investors. *The Journal of Finance 54*(6), 2263–2295. 1

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance 39*(4), 1127–1139. 1, 2, 7

Stoll, H. R. (1989). Inferring the components of the bid-ask spread: Theory and empirical tests. *The Journal of Finance 44*(1), 115–134. 1, 2

Table 1: Ideal Conditions

	Tick	5-Min	15-Min	30-Min	1-Hour	4-Hour	12-Hour	24-Hour
Midstd $\times 10^{-2}$	0.0200	0.0447	0.0775	0.110	0.155	0.310	0.536	0.758
The New Estimator								
Estimates $\times 10^{-3}$	0.300	0.300	0.300	0.300	0.298	0.310	0.289	0.278
Relative Estimates	1	1	1	1	0.993	1.033	0.963	0.927
Est-Std $\times 10^{-3}$	0.000658	0.00302	0.00884	0.0177	0.0375	0.123	0.423	0.850
RMSE $\times 10^{-3}$	0.000658	0.00302	0.00884	0.0177	0.0376	0.123	0.423	0.850
Huang and Stoll 1997								
Estimates $\times 10^{-3}$	0.300	0.300	0.301	0.301	0.301	0.301	0.282	0.280
Relative Estimates	1	1	1.003	1.003	1.003	1.003	0.94	0.933
Est-Std$\times 10^{-3}$	0.000622	0.00289	0.00898	0.0169	0.0371	0.146	0.414	0.865
RMSE$\times 10^{-3}$	0.000622	0.00289	0.00904	0.0169	0.0371	0.146	0.414	0.865

There are 500 replications. There are 432000 periods, each of which represents one minute, in each replication. Data of each replication are generated according to the following system. The order flow is drawn from a binomial distribution, i.e. $BS_t \sim B(1, 0.5)$. The mid-price return is drawn from a normal distribution of which the mean is zero and the variance is 4×10^{-8}, i.e. $\Delta M_t \sim N(0, 4 \times 10^{-8})$. The spread is fixed and equals to 0.0003, i.e. $SP_t = 0.0003$. The transaction price is the mid-price plus or minus a half spread, i.e. $s_t = M_t + \frac{SP_t}{2} \cdot BS_t$. Each replication is also sampled into longer time intervals: five-minute, fifteen-minute, thirty-minute, one-hour, four-hour, twelve-hour and twenty-four- hour, and only the close observations are kept. Thus, there are eight subgroups for each replication. For each subgroup, the standard deviation of mid-price returns, and the estimated spread are collected.

Midstd is the average of the standard deviations of mid-price returns.

Estimates is the average of the estimated spreads.

Relative Estimate represents the average of estimated spreads divided by the true spread. It is one if the estimate equals the true spread.

Est-Std is the standard deviation of the estimated spreads.

RMSE is the Root Mean Square Error.

Table 2: One-period Feedback Trading

	Tick	5-Min	15-Min	30-Min	1-Hour	4-Hour	12-Hour	24-Hour
Mid-price returns SD $\times 10^{-3}$	0.200	0.447	0.774	1.09	1.55	3.09	5.36	7.58
Spread/(returns SD)	1.5	0.671	0.387	0.273	0.194	0.0968	0.0560	0.0396
$Cov(\Delta M_t, BS_t)$ $\times 10^{-3}$	0.0479	0.0479	0.0479	0.0480	0.0477	0.0447	0.0584	0.0544
The New Estimator								
Estimation $\times 10^{-3}$	0.348	0.348	0.348	0.346	0.350	0.349	0.359	0.388
Relative Estimate	1.16	1.16	1.16	1.15	1.17	1.16	1.20	1.29
Est-Std$\times 10^{-3}$	0.000683	0.00305	0.00923	0.0180	0.0373	0.139	0.439	0.879
RMSE$\times 10^{-3}$	0.0480	0.0481	0.0489	0.0494	0.0624	0.147	0.443	0.883
Huang and Stoll 1997								
Estimates $\times 10^{-3}$	0.396	0.396	0.395	0.395	0.398	0.400	0.408	0.422
Relative Estimate	1.32	1.32	1.32	1.32	1.33	1.33	1.36	1.41
Est-Std$\times 10^{-3}$	0.000579	0.00304	0.00889	0.0175	0.0369	0.145	0.446	0.876
RMSE$\times 10^{-3}$	0.0960	0.0960	0.0954	0.0966	0.105	0.176	0.459	0.884

There are 500 replications. There are 432000 periods, each of which represents one minute, in each replication. Data of each replication are generated according to the following system. The order flow is positively autocorrelated. The mid-price return is drawn from a normal distribution of which the mean is zero and the variance is 4×10^{-8}, i.e. $\Delta M_t \sim N(0, 4 \times 10^{-8})$. Order flows is positively correlated to mid-price returns. The probability of a buy (sell) order being after a positive (negative) return is 65%. i.e. The spread is fixed and equals to 0.0003, i.e. $BS_t \sim B(1, 0.65)$ if $\Delta M_t > 0$ and $BS_t \sim B(1, 0.35)$ if $\Delta M_t < 0$. $SP_t = 0.0003$. The transaction price is the mid-price plus or minus a half spread, i.e. $s_t = M_t + \frac{SP_t}{2} \cdot BS_t$. Each replication is also sampled into longer time intervals: five-minute, fifteen-minute, thirty-minute, one-hour, four-hour, twelve-hour and twenty-four- hour, and only the close observations are kept. Thus, there are eight subgroups for each replication. For each subgroup, the standard deviation of mid-price returns, and the estimated spread are collected.

$Cov(\Delta M_t, BS_t)$ is the covariance of mid-price returns and order flows, which reflects the existence of feedback trading.

The other settings are the same as Table (1)

Table 3: Two-Period Feedback Trading

	Tick	5-Min	15-Min	30-Min	1-Hour	4-Hour	12-Hour	24-Hour
Midstd $\times 10^{-3}$	0.200	0.447	0.775	1.10	1.55	3.10	5.36	7.59
$Cov(\Delta M_t, BS_t)$ $\times 10^{-4}$	0.428	0.428	0.428	0.429	0.430	0.427	0.427	0.430
$Cov(\Delta M_{t-1}, BS_t)$ $\times 10^{-5}$	2.14	0.00421	-0.00239	-0.00424	-0.00943	0.0472	-0.0322	-0.0317
The New Estimator								
Estimates $\times 10^{-3}$	0.322	0.364	0.365	0.365	0.362	0.364	0.372	0.358
Relative Estimates	1.073	1.213	1.217	1.217	1.207	1.213	1.24	1.193
Est-Std$\times 10^{-3}$	0.000592	0.00291	0.00948	0.0193	0.0359	0.145	0.460	0.896
RMSE $\times 10^{-3}$	0.0220	0.0641	0.0657	0.0678	0.0716	0.158	0.466	0.898
Huang and Stoll 1997								
Estimates $\times 10^{-3}$	0.386	0.428	0.429	0.429	0.428	0.435	0.455	0.449
Relative Estimates	1.287	1.427	1.43	1.43	1.427	1.45	1.517	1.497
Est-Std$\times 10^{-3}$	0.000556	0.00299	0.00868	0.0180	0.0347	0.150	0.461	0.919
RMSE $\times 10^{-3}$	0.0860	0.128	0.129	0.130	0.133	0.202	0.486	0.931

There are 500 replications. There are 432000 periods, each of which represents one minute, in each replication. Data of each replication are generated according to the following system. The order flow is positively autocorrelated. The mid-price return is drawn from a normal distribution of which the mean is zero and the variance is 4×10^{-8}, i.e. $\Delta M_t \sim N(0, 4 \times 10^{-8})$. Order flows is positively correlated to past mid-price returns. The probability of a buy (sell) order with positive (negative) net feedback trading is 65%. i.e. The spread is fixed and equals to 0.0003, i.e. $BS_t \sim B(1, 0.65)$ if $\Delta M_t + 0.5\,\Delta M_{t-1} > 0$ and $BS_t \sim B(1, 0.35)$ if $\Delta M_t + 0.5\,\Delta M_{t-1} < 0$. $SP_t = 0.0003$. The transaction price is the mid-price plus or minus a half spread, i.e. $s_t = M_t + \frac{SP_t}{2} \cdot BS_t$. Each replication is also sampled into longer time intervals: five-minute, fifteen-minute, thirty-minute, one-hour, four-hour, twelve-hour and twenty-four- hour, and only the close observations are kept. Thus, there are eight subgroups for each replication. For each subgroup, the standard deviation of mid-price returns, and the estimated spread are collected.

Midstd is the average of the standard deviations of mid-price returns.

$Cov(\Delta M_t, BS_t)$ is the covariance of mid-price returns and order flows, which reflects the existence of current feedback trading.

$Cov(\Delta M_{t-1}, BS_t)$ is the covariance of mid-price returns and order flows, which reflects the existence of lagged feedback trading.

The other settings are the same as Table (1)

Table 4: Inventory Control and Asymmetric Information Components

	Tick	5-Min	15-Min	30-Min	1-Hour	4-Hour	12-Hour	24-Hour
Midstd $\times 10^{-2}$	0.0206	0.0461	0.0798	0.113	0.160	0.319	0.552	0.780
	0.333	0.333	0.333	0.333	0.328	0.312	0.247	0.200
The New Estimator								
Estimates $\times 10^{-3}$	0.250	0.250	0.250	0.250	0.252	0.254	0.254	0.347
Relative Estimates	0.833	0.833	0.833	0.833	0.84	0.847	0.847	1.157
Est-Std$\times 10^{-3}$	0.000656	0.00305	0.00886	0.0182	0.0374	0.153	0.457	0.917
RMSE$\times 10^{-3}$	0.0500	0.0501	0.0508	0.0532	0.0609	0.160	0.459	0.918
Huang and Stoll 1997								
Estimates $\times 10^{-3}$	0.300	0.300	0.300	0.300	0.303	0.302	0.311	0.350
Relative Estimates	1	1	1	1	1.01	1.007	1.037	1.167
Est-Std$\times 10^{-3}$	0.000626	0.00312	0.00862	0.0173	0.0391	0.155	0.441	0.927
RMSE$\times 10^{-3}$	0.000626	0.00312	0.00862	0.0173	0.0392	0.155	0.441	0.928

There are 500 replications. There are 432000 periods, each of which represents one minute, in each replication. Data of each replication are generated according to the following system. The order flow is drawn from a binomial distribution, i.e. $BS_t \sim B(1, 0.5)$. The mid-price return is influenced by the past order flow and a random shock drawn from a normal distribution of which the mean is zero and the variance is 4×10^{-8}. Thus there are the inventory control and the asymmetric information components in the spread. Formally, the mid-price returns are given by, $\Delta M_t = \frac{1}{3} BS_{t-1} \cdot \frac{SP_t}{2} + \varepsilon_t$ where $\varepsilon_t \sim N(0, 4 \times 10^{-8})$. The spread is fixed and equals to 0.0003, i.e. $SP_t = 0.0003$. The transaction price is the mid-price plus or minus a half spread, i.e. $s_t = M_t + \frac{SP_t}{2} \cdot BS_t$. Each replication is also sampled into longer time intervals: five-minute, fifteen-minute, thirty-minute, one-hour, four-hour, twelve-hour and twenty-four- hour, and only the close observations are kept. Thus, there are eight subgroups for each replication. For each subgroup, the standard deviation of mid-price returns, and the estimated spread are collected.

The other settings are the same as Table (1)

Table 5: Inventory Control and Asymmetric Information Components and Two-Period Feedback Trading

	Tick	5-Min	15-Min	30-Min	1-Hour	4-Hour	12-Hour	24-Hour
Midstd $\times 10^{-2}$	0.0206	0.0535	0.0881	0.123	0.173	0.344	0.594	0.839
ϱ	0.333	0.341	0.342	0.338	0.334	0.329	0.360	0.319
$Cov(\Delta M_t, BS_t)$ $\times 10^{-4}$	0.440	0.440	0.440	0.439	0.440	0.443	0.447	0.444
$Cov(\Delta M_{t-1}, BS_t)$ $\times 10^{-5}$	2.14	-0.00369	-0.00163	-0.00717	-0.0370	-0.0164	-0.0311	-0.00629
The New Estimator								
Estimates $\times 10^{-3}$	0.273	0.314	0.314	0.315	0.317	0.316	0.359	0.443
Relative Estimates	0.91	1.047	1.047	1.05	1.057	1.053	1.197	1.477
Est-Std$\times 10^{-3}$	0.000602	0.00343	0.00993	0.0197	0.04	0.161	0.332	0.426
RMSE$\times 10^{-3}$	0.0270	0.0144	0.0172	0.0248	0.0435	0.162	0.337	0.449
Huang and Stoll 1997								
Estimates $\times 10^{-3}$	0.386	0.431	0.431	0.431	0.432	0.432	0.430	0.413
Relative Estimates	1.287	1.437	1.437	1.437	1.44	1.44	1.433	1.377
Est-Std$\times 10^{-3}$	0.000615	0.00329	0.0105	0.0197	0.0428	0.171	0.485	0.93
RMSE$\times 10^{-3}$	0.0860	0.131	0.131	0.132	0.139	0.216	0.502	0.937

There are 500 replications. There are 432000 periods, each of which represents one minute, in each replication. Data of each replication are generated according to the following system. The order flow is drawn from a binomial distribution, i.e. $BS_t \sim B(1, 0.5)$. The mid-price return is influenced by the past order flow and a random shock drawn from a normal distribution of which the mean is zero and the variance is 4×10^{-8}. Thus there are the inventory control and the asymmetric information components in the spread. Formally, the mid-price returns are given by, $\Delta M_t = \frac{1}{3} BS_{t-1} \cdot \frac{SP_t}{2} + \varepsilon_t$ where $\varepsilon_t \sim N(0, 4 \times 10^{-8})$. The spread is fixed and equals to 0.0003, i.e. $SP_t = 0.0003$. The transaction price is the mid-price plus or minus a half spread, i.e. $s_t = M_t + \frac{SP_t}{2} \cdot BS_t$. Each replication is also sampled into longer time intervals: five-minute, fifteen-minute, thirty-minute, one-hour, four-hour, twelve-hour and twenty-four- hour, and only the close observations are kept. Thus, there are eight subgroups for each replication. For each subgroup, the standard deviation of mid-price returns, and the estimated spread are collected.

The other settings are the same as Table (1)

Table 6: Inventory Control and Asymmetric Information Components and Two-Period Feedback Trading

	Tick	5-Min	15-Min	30-Min	1-Hour	4-Hour	12-Hour	24-Hour
Midstd $\times 10^{-2}$	0.0224	0.0577	0.0983	0.138	0.195	0.390	0.675	0.954
ϱ	0.667	0.682	0.683	0.681	0.666	0.670	0.709	0.702
$Cov(\Delta M_t, BS_t)\ \times 10^{-4}$	0.452	0.452	0.453	0.453	0.453	0.457	0.461	0.46
$Cov(\Delta M_{t-1}, BS_t)\ \times 10^{-5}$	2.14	-0.00296	-0.00992	-0.0132	-0.00441	-0.0295	-0.0374	-0.0908
The New Estimator								
Estimates $\times 10^{-3}$	0.223	0.264	0.264	0.264	0.269	0.270	0.356	0.422
Relative Estimates	0.743	0.880	0.880	0.880	0.897	0.900	1.187	1.407
Est-Std$\times 10^{-3}$	0.00124	0.00425	0.0108	0.0232	0.0461	0.167	0.356	0.425
RMSE$\times 10^{-3}$	0.077	0.036	0.038	0.043	0.056	0.170	0.360	0.442
Huang and Stoll 1997								
Estimates $\times 10^{-3}$	0.386	0.433	0.433	0.433	0.434	0.448	0.465	0.409
Relative Estimates	1.287	1.443	1.443	1.443	1.447	1.493	1.550	1.363
Est-Std$\times 10^{-3}$	0.000615	0.00329	0.0105	0.0197	0.0428	0.171	0.485	0.93
RMSE$\times 10^{-3}$	0.086	0.133	0.133	0.135	0.143	0.242	0.582	1.105

There are 500 replications. There are 432000 periods, each of which represents one minute, in each replication. Data of each replication are generated according to the following system. The order flow is drawn from a binomial distribution, i.e. $BS_t \sim B(1, 0.5)$. The mid-price return is influenced by the past order flow and a random shock drawn from a normal distribution of which the mean is zero and the variance is 4×10^{-8}. Thus there are the inventory control and the asymmetric information components in the spread. Formally, the mid-price returns are given by, $\Delta M_t = \frac{2}{3} BS_{t-1} \cdot \frac{SP_t}{2} + \varepsilon_t$ where $\varepsilon_t \sim N(0, 4 \times 10^{-8})$. The spread is fixed and equals to 0.0003, i.e. $SP_t = 0.0003$. The transaction price is the mid-price plus or minus a half spread, i.e. $s_t = M_t + \frac{SP_t}{2} \cdot BS_t$. Each replication is also sampled into longer time intervals: five-minute, fifteen-minute, thirty-minute, one-hour, four-hour, twelve-hour and twenty-four- hour, and only the close observations are kept. Thus, there are eight subgroups for each replication. For each subgroup, the standard deviation of mid-price returns, and the estimated spread are collected.

The other settings are the same as Table (1)

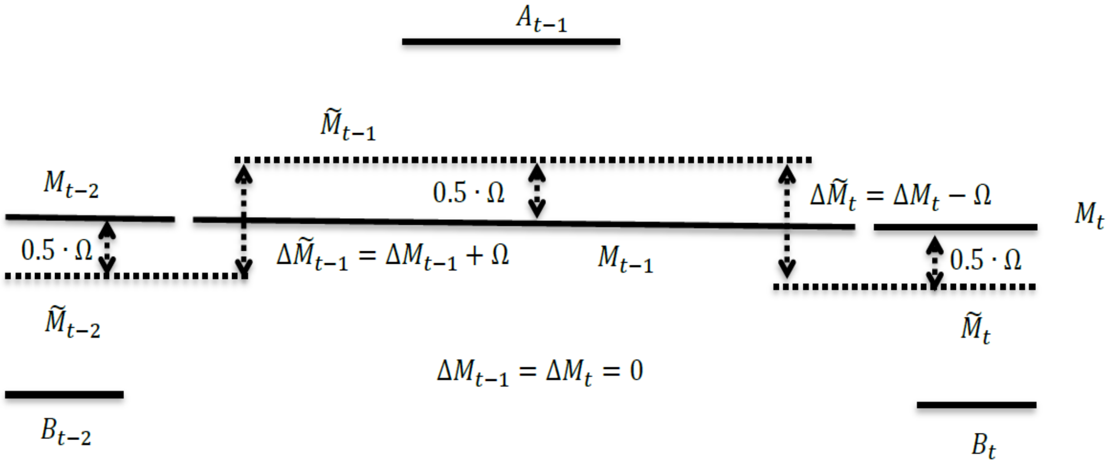

Figure 1: The Conjecture of the Spread

The figure show the general relationships between the returns of true mid-prices and the returns of conjectural mid-prices, when only transaction prices and the directions of transactions are known. The conjectural spread here is less than the true spread. The conjectural mid-prices are obtained from Equation (2) using transaction prices, given a conjectural spread. The dot-lines are conjectural mid-prices and the solid lines are true mid-prices or transaction prices. A and B represent the ask and bid prices respectively; these are the prices that are actually observed. M and $\widetilde{M}$ represent the true mid-price and the conjecture of it respectively. Δ is the first-order difference operator. Ω is the error of the conjecture which is the difference between the true spread and the conjectural spread, or equivalently, between the true mid-price and the conjectural mid-price.

There is a sell order in period $t-2$, and, thus, the bid price is recorded. There is a buy order in period $t-1$, and thus the ask price is recorded. There is a sell order in period t, and thus the bid price is recorded. Because the conjectural spread is less than the true spread, the in periods $t-2$ and t, the conjectural mid-prices are $0.5 \cdot \Omega$ less than the true ones, and in period $t-1$, the conjectural mid-price is $0.5 \cdot \Omega$ greater than the true one.

Between periods $t-2$ and $t-1$, the trade direction changes from selling to buying, so the conjectural error makes the conjectural mid-price return greater than the true mid-price return ($\Delta\widetilde{M}_{t-1} = \Delta M_{t-1} + \Omega = \Omega$).

Between periods $t-1$ and t, the trade direction switches back from buying to selling, so the conjectural error partially cancels the true mid-price return ($\Delta\widetilde{M}_t = \Delta M_t - \Omega = -\Omega$). When the trade direction does not change, the returns of the conjectural mid-prices and of the true mid-prices are the same.