Measuring Transaction Costs in OTC markets

Filip Zikes*

February 1, 2018

Abstract

This paper develops measures of transaction costs in the absence of transaction timestamps and information about who initiates transactions, which are data limitations that often arise in studies of over-the-counter markets. I propose new measures of the effective spread and study the performance of all estimators analytically, in simulations, and present an empirical illustration with small-cap stocks for the 2005–2014 period. My theoretical, simulation, and empirical results provide new insights into measuring transaction costs and may help guide future empirical work.

JEL Classification: C14, C15, G20
Keywords: OTC markets, transaction costs, effective spread, simulated method of moments

*Board of Governors of the Federal Reserve System, Division of Financial Stability, 20th Street and Constitution Avenue N.W., Washington, D.C. 20551, United States. Phone: +1-202-475-6617. Email: filip.zikes@frb.gov. I am grateful to Evangelos Benos, Dobrislav Dobrev, Erik Hjalmarsson, Ivan Ivanov, John Schindler, Michalis Vasios, and Yang-Ho Park for valuable comments, and to Eric Parolin and Margaret Yellen for excellent research assistance. The results reported in the paper were generated using programs written in SAS and the Ox language of Doornik (2007); the programs are available on request. The views expressed in this paper are the sole responsibility of the author and should not be interpreted as representing the views of the Federal Reserve Board or any other person associated with the Federal Reserve System.

1 Introduction

This paper develops measures of transaction costs that do not require observing the intraday transaction times or knowing who initiates trades (buyers vs. sellers). A recent and growing literature on large and previously opaque over-the-counter (OTC) markets employs transaction data that suffer from these limitations. Examples include studies of the credit default swap market (Chen et al., 2011; Benos, Wetherilt, and Zikes, 2013; Biswas, Nikolova, and Stahel, 2014; Du et al., 2017), the interest rate swap market (Chen et al., 2012; Benos, Payne, and Vasios, 2016), the U.K. sovereign bond market (Benos and Zikes, 2016), and the U.S. corporate bond market (Bessembinder et al., 2006). But to the best of my knowledge, the literature has not yet formally tackled the problem of estimating transaction costs when timestamps and trade direction are missing.[1] The contribution of this paper is to fill this gap.

I propose three consistent estimators of the effective spread and study their sampling properties. The first one develops the idea of Benos and Zikes (2016), who suggest inferring the effective spread from the dispersion of the transaction prices from (1) some benchmark or reference price (e.g., end-of-day composite quote) and (2) the average transaction price. The estimator is available in closed form, which allows me to establish its finite-sample properties analytically and compare them to some well-known (infeasible) measures, explicitly quantifying the loss of information due to the missing timestamps, trade direction, or both. I also show that the estimator is robust to stochastic volatility.

The other two measures I propose are also moment-based and combine the ideas of Corwin and Schultz (2012) and Benos and Zikes (2016). The first one is based on the daily range, which is the difference between the daily high and low prices, together with the sample variance of the transaction prices. The advantage of this

[1]Throughout this paper, I use the term "missing timestamps" to mean not only that transaction times are not known, but also that transactions are not chronologically ordered so that intraday price changes cannot be calculated.

measure is that it is based solely on transaction prices and does not require a daily benchmark or reference price, which may be difficult to obtain for some illiquid assets. At the same time, it utilizes all available data (transaction prices), unlike the Corwin and Schultz (2012) measure, which only uses the daily high and low prices. If a benchmark price is available, however, it is, of course, optimal to use all three moment conditions—the two dispersion metrics and the daily range—and this is how I construct my third estimator. Because I do not assume that the number of transactions is large, I have to resort to the simulated method of moments (SMM). Despite having to approximate the expected range by simulation, the estimator turns out to be computationally cheap and easy to implement in practice. I provide simple yet accurate small-sample approximations to the unknown moment that significantly speed up computations.

To summarize my theoretical results, I find that the absence of timestamps or trade direction lead to a reduced convergence rate of the effective spread estimators. While in the case of full information the effective spread can be estimated n-consistently as the number of intraday transactions (n) increases, when timestamps or trade direction are missing, only $\sqrt{n}$-consistency can be achieved, and when both are missing, the effective spread cannot be estimated consistently from intraday data alone–averaging over an increasing number of days (T) is necessary. Thus, accurate estimates can only be obtained from weeks or months worth of transaction data.

To corroborate the theoretical findings and to study how the various estimators perform in practice, both in absolute terms and relative to alternative measures, I conduct two experiments. First, I run Monte Carlo simulations to assess the performance in a controlled environment. Second, I provide an empirical illustration with small-cap stocks listed on the New York Stock Exchange (NYSE) using data from the Trade and Quote Database (TAQ). The second exercise therefore uses real-world rather than simulated data to assess and compare performance, similar to e.g. Goyenko, Holden, and Trzcinka (2009), who compare low-frequency liquidity measures with their high-frequency counterparts for selected U.S. stocks.

To summarize the findings of these exercises, I find that the loss of information due to missing timestamps and/or trade direction is generally large. The measures I propose in this paper deliver accurate estimates of transaction costs in situations where the transactions costs are high relative to the volatility of the efficient price and when the number of intraday transactions is small; in such situations, my estimators often perform on par or even better than their infeasible counterparts. They may be, therefore, suitable for relatively illiquid, infrequently traded assets that exhibit relatively low fundamental volatility, such as some corporate and municipal bonds. They should not be applied, however, to highly liquid assets, such as listed equities, which trade with a tight spread and tend to be quite volatile. Fortunately, for these assets, high-quality time-stamped transaction data are typically available.

In OTC markets, timestamps are often inaccurate or outright missing for various reasons. In the credit default swap and interest rate data mentioned previously, transaction times are simply not reported, and the trade reporting time does not necessarily correspond to the actual trade time, making it impossible to chronologically order transactions. In the U.K. sovereign bond market data used by Benos and Zikes (2016), the timestamps are not accurate in the sense that two parties to the same transaction report widely different transaction times. More generally, though, the trading protocol in OTC markets often involves negotiation that may stretch over a period of time, and so the exact timing of the trade may be ambiguous; consider, for example, the "workup" protocol recently studied in Duffie and Zhu (forthcoming).

Trade direction cannot be easily inferred because trades cannot be aligned with intraday quotes when timestamps are missing and because data on intraday quotes are rarely available in OTC markets, making it impossible to use popular trade-signing algorithms such as that of Lee and Ready (1991). Researchers often assume that clients initiate trades with dealers, motivated by the fact that dealers are the main liquidity providers in these markets (Bessembinder, Maxwell, and Venkataraman, 2006; Edwards, Harris, and Piwowar, 2007). However, as recently shown by Choi and Huh (2017) for the U.S. corporate bond market, dealers often initiate trades with clients as well, implying potentially serious missclassification issues associated with this identification method. Moreover, in some markets, interdealer trades account for more than two-thirds of all transactions (Benos et al., 2013) and so the vast majority of transactions cannot be signed using this approach anyway. Thus, standard methods cannot be used to measure transaction costs in these large and important financial markets given the limitations of the available data.

Apart from the literature on measuring transaction costs (see Harris, 2015, Section 3.1 for a recent overview), my paper is also related to the recent literature on measuring volatility using high-frequency data starting with Andersen and Bollerslev (1998) and Barndorff-Nielsen and Shephard (2002); see Aït-Sahalia and Jacod (2014) for a recent textbook treatment. Some of my estimators employ the range—that is, the difference between intraday high and low prices—and here I draw on the ideas of Christensen and Podolskij (2007) and Christensen, Podolskij, and Vetter (2009). Although the data-generating process I assume is very similar to many papers in this literature, the problem studied in my paper is different in three important ways.

First, my goal is to estimate the effective spread. Thus, what the realized volatility literature (e.g. Aït-Sahalia, Mykland, and Zhang, 2005; Zhang, Mykland, and Aït-Sahalia, 2005, Hansen and Lunde, 2006; Barndorff-Nielsen, Hansen, Lunde, and Shephard, 2008a,b; Kalnina and Linton, 2008; Christensen, Podolskij, and Vetter, 2009) treats as microstructure noise is precisely my object of interest, and what that literature is interested in estimating—the variation of the efficient price—is a source of noise in my framework. Second, in-fill asymptotics do not always apply—that is, increasing the number of intraday observations (n) does not generally improve the precision of the effective spread estimator. I have to rely on an increasing number of days (T) and employ large-T asymptotics or double asymptotics (both $n \to \infty$ and $T \to \infty$). Finally, my estimation framework is model based, unlike the estimation methods in the realized volatility literature that operate in a model-free environment.

The rest of the paper is organized as follows. In Section 2, I set out my theoretical framework. In Section 3, I propose a closed-form measure of the effective spread that does not require either timestamps or trade direction and study its properties analytically, explicitly quantifying the loss of information associated with missing timestamps, trade direction, or both. In Section 4, I introduce range-based estimators of the effective spread and propose simple computational methods to implement the estimator in practice. In Section 5, I discuss robustness to stochastic volatility. Section 6 reports Monte Carlo simulations. In Section 7, I present an empirical application to small-cap equities, and Section 8 concludes. Proofs are collected in the Appendix.

2 Framework

The effective spread is defined as two times the difference between the actual transaction price (P) and the prevailing mid-quote or some proxy for the true value of the asset (efficient price) (M) at the time of the transaction. It can be expressed in absolute terms—that is, $2|P - M|$—or in relative terms—that is, $2|P - M|/M$ or $2|\log(P) - \log(M)|$. Like the bid-offer spread, the effective spread measures round trip transaction costs, but it is based on actual transaction price rather than on quoted prices. The effective spread can also be seen as a measure of the price impact of a trade, and because the price impact and transaction costs tend to vary inversely with liquidity, it is frequently used as a measure of liquidity (Foucault, Pagano, and Roell, 2013).

My theoretical framework is essentially borrowed from Roll (1984) and it can be easily cast in continuous time as in Christensen, Podolskij and Vetter (2009). Suppose we have a sample of T days and divide each day into n subintervals of equal length. I assume that a transaction arrives at the beginning of each of these subintervals and that the associated logarithmic transaction prices—$p_{i,t}$; $i = 1, ..., n$; $t = 1, ..., T$—are related to the logarithmic efficient price, $m_{i,t}$, by

$$p_{i,t} = m_{i,t} + \frac{s}{2} q_{i,t}, \tag{1}$$

where s is the proportional effective spread and $q_{i,t}$ is a binary variable indicating whether the i-th transaction on day t is buyer initiated ($q_{i,t} = 1$) or seller initiated ($q_{i,t} = -1$). I initially assume that the efficient price is observable at the end of the day—that is, at the end of the last subinterval n. I later relax this assumption and propose estimators that do not require observing m at all. Following Roll (1984) and Benos and Zikes (2016), I assume that the logarithmic efficient price m follows a random walk with independently and identically distributed (iid) increments:

$$m_{i+1,t} = m_{i,t} + \epsilon_{i+1,t}, \tag{2}$$

where $\mathrm{E}(\epsilon_{i,t}) = 0$ and $\mathrm{E}(\epsilon_{i,t}^2) = \sigma^2/n$. Thus, the daily integrated variance of the efficient price equals σ^2 for any n and t. Finally, I assume that $q_{i,t}$ is uncorrelated with $m_{j,s}$ for all i, j, t, s and that $q_{i,t}$ is serially uncorrelated with $\mathrm{E}(q_{i,t}) = 0$—that is, there is the same number of buyer- and seller-initiated trades on average. I make no assumptions on the overnight return of the efficient price, $m_{0,t} - m_{n,t-1}$.

3 Baseline estimator

Inspired by Jankowitsch, Nashikkar, and Subrahmanyam (2011), Benos and Zikes (2016) rely on the dispersion of transaction prices around some benchmark price, but they recognize that the dispersion metric is affected by the intraday volatility of the benchmark price in a nontrivial way. They suggest using two dispersion metrics,

$$\hat{d}_t^2 = \frac{1}{n}\sum_{i=1}^{n}(p_{i,t} - m_{0,t})^2, \quad \tilde{d}_t^2 = \frac{1}{n-1}\sum_{i=1}^{n}(p_{i,t} - \bar{p}_t)^2, \tag{3}$$

and show that under the assumptions stated in the previous section, the two metrics satisfy

$$\mathrm{E}(\hat{d}_t^2) = \frac{s^2}{4} + \frac{\sigma^2}{2}\left(\frac{n+1}{n}\right), \quad \mathrm{E}(\tilde{d}_t^2) = \frac{s^2}{4} + \frac{\sigma^2}{6}\left(\frac{n+1}{n}\right). \tag{4}$$

Solving for s^2, censoring at zero, and taking the square root yield the relative effective spread measure:

$$ES_t = \sqrt{\max\{2(3\tilde{d}_t^2 - \hat{d}_t^2), 0\}}. \tag{5}$$

My baseline estimator develops the idea of Benos and Zikes (2016). I define $\hat{s}_t^2 = 2(3\tilde{d}_t^2 - \hat{d}_t^2)$, where $\hat{d}_t^2$ and $\tilde{d}_t^2$ are given in equation (3) and start by deriving the variance of $\hat{s}_t^2$, as it will be invoked repeatedly in the paper.

Proposition 1 *Provided that the fourth moment of $\epsilon_{1,1}$ exists,*

$$\mathrm{Var}(\hat{s}_t^2) = \frac{9s^4}{2n(n-1)} + \frac{2s^2\sigma^2(2n^2+3n+1)}{n^2(n-1)} + \frac{2(2n\sigma^4+\sigma^4+2\kappa)(2n^3+7n^2+7n+2)}{15n^3(n-1)}, \tag{6}$$

where $\kappa = \mathrm{E}(\epsilon_{1,1}^4) - 3\sigma^4$ is the excess kurtosis of $\epsilon_{1,1}$.

Equation (6) implies that although $\hat{s}^2$ is an unbiased estimator of s^2, it is not consistent as the number of intraday transactions increases because $\mathrm{Var}(\hat{s}_t^2)$ =

$\frac{8}{15}\sigma^4 + O(n^{-1})$ as $n \to \infty$. This result is due to the fact that we are averaging random walks in levels (prices) as opposed to first differences (returns), which cannot be constructed due to missing timestamps.

To derive a consistent estimator of s based on $\hat{s}_t^2$, we need to average $\hat{s}_t^2$ over an increasing number of days before censoring at zero and taking the square root as in equation (5). The resulting estimator, which I denote by $ES_T^{(1)}$, is thus given by

$$ES_T^{(1)} = \sqrt{\max\left\{\frac{1}{T}\sum_{t=1}^{T} 2(3\tilde{d}_t^2 - \hat{d}_t^2), 0\right\}}. \tag{7}$$

Given the nonlinear nature of the estimator, $\mathrm{E}(ES_T^{(1)})$ and $\mathrm{Var}(ES_T^{(1)})$ are not available in closed form. I employ a Taylor series expansion of $ES_T^{(1)}$ around s, $s > 0$, together with equation (6), to establish the leading terms (as $T \to \infty$). The leading term of the bias reads

$$\lim_{T\to\infty} \mathrm{E}[T(ES_T^{(1)} - s)] = -\frac{1}{15}\frac{\sigma^4}{s^3} - \left(\frac{1}{3}\frac{\sigma^4}{s^3} + \frac{1}{15}\frac{\kappa}{s^3} + \frac{1}{2}\frac{\sigma^2}{s}\right)\frac{1}{n} + O\left(\frac{1}{n^2}\right), \tag{8}$$

implying that $ES_T^{(1)}$ tends to underestimate the true effective spread. The limiting variance reads

$$\lim_{T\to\infty} \mathrm{Var}[\sqrt{T}(ES_T^{(1)} - s)] = \frac{2}{15}\frac{\sigma^4}{s^2} + \left(\frac{2}{3}\frac{\sigma^4}{s^2} + \frac{2}{15}\frac{\kappa}{s^2} + \sigma^2\right)\frac{1}{n} + O\left(\frac{1}{n^2}\right). \tag{9}$$

As expected, the absolute bias and variance decrease with the signal-to-noise ratio (SNR) s/σ, so it is more difficult to estimate the effective spread when it is small relative to the volatility of the efficient price. The absolute bias and variance of $ES_T^{(1)}$ also increase with excess kurtosis, but this only matters when the number of transactions is small; the contribution of κ vanishes as $n \to \infty$. The second terms in the expansions also show that for sufficiently large n, the absolute bias and variance of $ES_T^{(1)}$ decrease with the number of transactions, as the coefficients on the n^{-1} terms in equations (8) and (9) are always positive.

It follows from the assumptions stated in Section 2 and standard limit theorems that as $T \to \infty$, $ES_T^{(1)} \to_p s$ and if $s > 0$ and $\kappa < \infty$, $\sqrt{T}(ES_T^{(1)} - s) \to_d \mathrm{N}(0, \omega^2)$, where $\omega^2 = \frac{1}{4s^2}\mathrm{Var}(\hat{s}^2)$. The limiting variance of $ES_T^{(1)}$ has a particularly simple form if we consider the asymptotics where both $T, n \to \infty$, which may be appropriate in situations where the number of daily transactions is large. Then, from equation (9), we obtain $\sqrt{T}(ES_T^{(1)} - s) \to_d \mathrm{N}(0, \frac{2\sigma^4}{15s^2})$. Feasible inference can be obtained by replacing the unknown s and σ^2 in the limiting variance with their sample counterparts, $ES_T^{(1)}$ and $\hat{\sigma}_T^2$, respectively, where

$$\hat{\sigma}_T^2 = \max\left\{\frac{1}{T}\sum_{t=1}^{T} 3(\hat{d}_t^2 - \tilde{d}_t^2), 0\right\} \tag{10}$$

is a consistent estimator of σ^2.

3.1 Comparison with estimators that use timestamps or trade direction

In this section, I compare the baseline estimator with several well-known measures of the effective spread that require timestamps or trade direction, or both. The goal is to assess how serious the loss of information associated with these data limitations is.[2]

3.1.1 Observable timestamps

Should timestamps be available, one would typically use the Roll (1984) estimator, which is equal to minus 4 times the sample first-order autocovariance of intraday returns:

$$\hat{\gamma}_t^2 = -\frac{4}{n-2}\sum_{i=3}^{n}(p_{i,t} - p_{i-1,t})(p_{i-1,t} - p_{i-2,t}). \tag{11}$$

[2]In the rest of this section, all analytical results are presented without proof to save space. The derivations follow similar steps as the derivation of equation (6) and can be obtained upon request.

It is easy to show that in my setup the estimator is unbiased for s^2, and its variance reads

$$\mathrm{Var}(\hat{\gamma}_t^2) = \left(\frac{16\sigma^4}{n^2} + \frac{16 s^2 \sigma^2}{n} + \frac{2 s^4 (n-3)}{(n-2)} + 3 s^4 \right) \frac{1}{n-2}. \tag{12}$$

Clearly, $\mathrm{Var}(\hat{\gamma}_t^2) = \frac{5s^4}{n} + O(n^{-2})$, so $\hat{\gamma}_t^2$ is a consistent estimator of s^2 as $n \to \infty$. Similar to $ES_T^{(1)}$, $\hat{\gamma}_t^2$ can be averaged over T days and censored at zero to obtain a nonnegative estimator of s:

$$Roll_T = \sqrt{\max \left\{ \frac{1}{T} \sum_{t=1}^{T} \hat{\gamma}_t^2, 0 \right\}}. \tag{13}$$

Unlike $ES_T^{(1)}$, $Roll_T$ is consistent for s as $n \to \infty$ for any T.

But the Roll estimator is not the only $\sqrt{n}$-consistent measure of s. Christensen, Podolskij, and Vetter (2009) propose an estimator based on realized volatility, which can be tailored to my framework as follows:

$$\hat{\omega}_t^2 = \frac{2}{n-1} \sum_{i=1}^{n-1} (p_{i+1,t} - p_{i,t})^2. \tag{14}$$

It is straightforward to show that

$$\mathrm{E}(\hat{\omega}_t^2) = s^2 + \frac{2\sigma^2}{n} \quad \text{and} \quad \mathrm{Var}(\hat{\omega}_t^2) = \frac{s^4}{n-1} + \frac{8 s^2 \sigma^2}{n(n-1)} + \frac{4(\kappa - \sigma^4)}{n^2(n-1)} \tag{15}$$

where $\kappa = \mathrm{E}(\epsilon_{1,1}^4)$, which implies that $\hat{\omega}^2$ is asymptotically unbiased and its variance satisfies $\mathrm{Var}(\hat{\omega}_t^2) = \frac{s^4}{n} + O(n^{-2})$. The limiting variance is five times smaller than that of the Roll estimator $\hat{\gamma}_t^2$, but the finite-sample bias can be large when σ^2 is large. Unlike $\hat{s}^2$ or $\hat{\gamma}^2$, the estimator $\hat{\omega}_t^2$ is non negative by construction, and hence there is no need for censoring when constructing a consistent estimator of s based on T days worth of data:

$$RV_T^{all} = \sqrt{\frac{1}{T} \sum_{t=1}^{T} \hat{\omega}_t^2}. \tag{16}$$

Similar to $Roll_T$, and unlike $ES_T^{(1)}$, RV_T^{all} is consistent for s as $n \to \infty$ regardless of T.

3.1.2 Observable trade direction

As suggested by Warga (1991) and Schultz (2001), when the trade direction is observable one can simply regress the difference between the transaction price and some benchmark price on the trade indicator. Here I continue assuming that the benchmark price equals the end-of-day mid-quote and suggest running the following OLS regression:

$$p_{i,t} - m_{0,t} = \beta_t^{(1)} q_{i,t} + u_{i,t}^{(1)}. \tag{17}$$

If the daily benchmark prices are not available, one can use the average transaction price instead and run the regression

$$p_{i,t} - \bar{p}_t = \beta_t^{(2)} q_{i,t} + u_{i,t}^{(2)}. \tag{18}$$

If model (1) is the data-generating process, the regression innovations are given by $u_{i,t}^{(1)} = m_{i,t} - m_{0,t}$ and $u_{i,t}^{(2)} = m_{i,t} - \bar{m}_t - (s/2)\bar{q}_t$, respectively. It is easy to show that the ordinary least squares (OLS) estimators of $\beta_t^{(1)}$ and $\beta_t^{(2)}$in regressions (17) and (18) satisfy, under my assumptions:

$$\mathrm{E}(2\hat{\beta}_t^{(1)}) = s, \quad \mathrm{Var}(2\hat{\beta}_t^{(1)}) = \frac{2\sigma^2(n+1)}{n^2}, \tag{19}$$

$$\mathrm{E}(2\hat{\beta}_t^{(2)}) = s - \frac{s}{n}, \quad \mathrm{Var}(2\hat{\beta}_t^{(2)}) = \frac{2\sigma^2(n+1)(n-1)}{3n^3} + 2s^2\left(\frac{1}{n^2} - \frac{1}{n^3}\right) \tag{20}$$

Replacing $m_{i,t}$ with $m_{0,t}$ or $\bar{p}_t$ therefore does not render the regression-based estimator inconsistent as it did for $\hat{d}_t^2$ and $\hat{d}_t^2$ in Section 3. The OLS estimators $2\hat{\beta}_t^{(i)}$, $i = 1, 2$ will converge in probability to s at rate $\sqrt{n}$ as $n \to \infty$ as did the Roll and RV-based measures in Section 3.1.

Similar to the effective spread measures in the previous section, none of these OLS estimators are guaranteed to be non-negative. Thus, I censor them at zero and denote the regression-based estimators of the effective spread s for day t by $RS_t^{(i)} = \max\{2\hat{\beta}_t^{(i)}, 0\}$, $i = 1, 2$. When estimating the effective spread using the full sample of T days with n transactions each, one simply runs the regressions (17) and (18) using all nT observations. I denote these estimators by $RS_T^{(i)}$, $i = 1, 2$. Note that in practice, n does not have to be the same for all days in the sample; I only make this assumption here to simplify derivations.

Hong and Warga (2000) propose an estimator that is closely related to the regression-based estimators discussed in the previous paragraph. They suggest to compare the same-day same-bond purchase prices with sales prices. Formally, their estimator for day t is given by

$$\hat{\delta}_t = \begin{cases} \frac{\sum_{i=1}^n p_{i,t} 1\{q_{i,t}=1\}}{\sum_{i=1}^n 1\{q_{i,t}=1\}} - \frac{\sum_{i=1}^n p_{i,t} 1\{q_{i,t}=-1\}}{\sum_{i=1}^n 1\{q_{i,t}=-1\}} & \text{if } \exists i, j \text{ s.t. } q_{i,t} = 1 \text{ and } q_{j,t} = -1, \\ 0 \quad \text{otherwise}, \end{cases} \tag{21}$$

where I explicitly account for the fact that there may be no buys or sales on any given day. In a sample of n transactions, this case occurs, under my assumptions, with probability $(1/2)^{n-1}$ and has a negligible effect on the properties of the estimator. Due to the nonlinearity of the estimator, it is difficult to derive the exact variance of $\hat{\delta}_t$ in closed form, but it can be shown that

$$\mathrm{E}(\hat{\delta}_t) = s - s\left(\frac{1}{2}\right)^{n-1}, \quad \lim_{n\to\infty} n\mathrm{Var}(\hat{\delta}_t) = \frac{2\sigma^2}{3}. \tag{22}$$

Thus, the Hong and Warga (2000) estimator has the same limiting variance as the estimator based on regression (18). In fact, it can be shown that the two estimators are asymptotically equivalent—that is, $\sqrt{n}(\hat{\delta}_t - 2\hat{\beta}_t^{(2)}) \xrightarrow{p} 0$ as $n \to \infty$. They can differ in small samples, however, so I will consider both estimators in the simulation and empirical application later in the paper. $\hat{\delta}_t$ is not guaranteed to be non negative and so I censor it at zero. Finally, to obtain an estimator that utilizes the full sample of T days, I simply average the daily estimates $\hat{\delta}_t$ before censoring. Denoting the resulting estimator by $RS_T^{(3)}$, it is given by

$$RS_T^{(3)} = \max\left\{\frac{1}{T}\sum_{t=1}^{T}\hat{\delta}_t, 0\right\}. \tag{23}$$

3.1.3 Observable timestamps and trade direction

Finally, I investigate the loss of efficiency associated with both missing timestamps and trade direction. If timestamps were available, one could run the regression of p on q in first differences:

$$\Delta p_{i,t} = \beta_t^{(3)}\Delta q_{i,t} + u_{i,t}^{(3)}. \tag{24}$$

The gain in efficiency compared with the regressions in levels is simply due to the fact that $u_{i,t}^{(3)} = \epsilon_{i,t}$, which has much smaller variance than either $u_{i,t}^{(1)}$ or $u_{i,t}^{(2)}$ and is serially uncorrelated. A complication with the standard OLS estimator $\hat{\beta}^{(3)}$ in regression (24) is that it is not always well defined. We have $\hat{\beta}^{(3)} = \sum_{i=2}^{n}\Delta p_{i,t}\Delta q_{i,t}/\sum_{i=2}^{n}(\Delta q_{i,t})^2$ and it is not difficult to show that $\mathbb{P}(\sum_{i=2}^{n}(\Delta q_{i,t})^2 = 0) = (1/2)^{n-1}$. But because $\frac{1}{n-1}\sum_{i=2}^{n}(\Delta q_{i,t})^2$ converges in probability to 2 as $n \to \infty$, an asymptotically equivalent, well-defined estimator can be obtained by simply setting $\sum_{i=2}^{n}(\Delta q_{i,t})^2$ equal to $2(n-1)$ in $\hat{\beta}_t^{(3)}$ whenever $\sum_{i=2}^{n}(\Delta q_{i,t})^2 = 0$—that is, I define

$$\tilde{\beta}_t^{(3)} = \frac{\sum_{i=2}^{n}\Delta p_{i,t}\Delta q_{i,t}}{1\{\sum_{i=2}^{n}(\Delta q_{i,t})^2 = 0\}2(n-1) + \sum_{i=2}^{n}(\Delta q_{i,t})^2}. \tag{25}$$

Clearly, $\mathrm{E}(2\tilde{\beta}_t^{(3)}) = s$, so $2\tilde{\beta}_t^{(3)}$ is an unbiased estimator of s. It is difficult to derive the exact variance of $\tilde{\beta}_t^{(3)}$, but it is easy to show that the limiting variance satisfies $\lim_{n\to\infty} n^2\mathrm{Var}(2\hat{\beta}_t^{(3)}) = 2\sigma^2$ and that $2\tilde{\beta}_t^{(3)}$ is a n-consistent estimator of s.

Thus, observing both timestamps and trade direction at the same time improves the convergence rate further: recall that the RMSE of the estimators $Roll_T$, RV_T^{all}, $RS_T^{(i)}$, $i = 1, 2$ only decays at rate $n^{1/2}$. In small samples, $\tilde{\beta}_t^{(3)}$ can be negative with positive probability, so I define $RS_t^{(3)} = \max\{2\tilde{\beta}_t^{(3)}, 0\}$ as an estimator of s for day t and $RS_T^{(3)} = \max\{2\tilde{\beta}_T^{(3)}, 0\}$, where $\tilde{\beta}_T^{(3)}$ is obtained by running regression (24) using all nT observations (T days with n transactions each).

3.1.4 Summary

The analytical comparison reveals that the absence of timestamps and/or trade direction reduces the convergence rates of the effective spread estimators. In the full-information case, one can achieve n-consistency, while in the absence of either timestamps or trade direction, only $\sqrt{n}$-consistency is possible. In the absence of timestamps, the limiting RMSE only depends on s (equations (12) and (15)), while in the case of missing trade direction, it is solely driven by σ^2 (equations (19), (20), and (22)). Finally, when both are missing, consistency cannot be achieved by increasing the number of intraday observations and averaging over an increasing number of days is necessary. The limiting variance of the effective spread estimator depends on the ratio of σ^2 and s, see equation (9).

Now, in practice this means that the relative performance of the various estimators depends on the parameter configuration and the number of intraday observations. Clearly, when the signal-to-noise ratio σ^2/s is high, the absence of time stamps and trade direction lead to significant deterioration in RMSE for any n. But when σ^2/s is small, there may exists a range for n where the infeasible estimators do not really improve much upon the estimator that does not require either timestamps or trade direction. This is a useful result because in practice it is precisely illiquid, infrequently traded asses for which these data limitations occur.

4 Range-based estimators

The baseline estimator is simple to compute, but it requires observing the benchmark price $m_{0,t}$. When these prices or mid-quotes are not available, $\hat{d}_t$ cannot be calculated and we need an alternative moment condition to use together with $\mathrm{E}(\tilde{d}_t^2)$ in equation (4). Inspired by Corwin and Schultz (2012), I use the daily range:

$$\hat{r}_t^2 = (\max_j p_{j,t} - \min_j p_{j,t})^2. \tag{26}$$

The range has a long tradition in financial econometrics, dating back to Parkinson (1980), and has been widely used for estimating volatility from intraday data (Christensen and Podolskij, 2007; Christensen, Podolskij, and Vetter, 2009; Dobrev, 2007). It is clear that $\hat{r}_t^2$ is expected to depend on both s and σ, as do $\tilde{d}_t^2$ and $\tilde{d}_t^2$. Corwin and Schultz (2012) combine equation (26) with a second moment condition based on the squared range over two consecutive days:

$$\hat{r}_{t:t+1}^2 = (\max_j p_{j,t:t+1} - \min_j p_{j,t:t+1})^2, \tag{27}$$

where $p_{j,t:t+1}$ denotes the j-th transaction price in a two-day window starting on day t. They assume that trading takes place continuously ($n \to \infty$), the overnight returns are zero, and the observed high and low prices are related to the high and low efficient prices by $\max_i p_{i,t} = (1 + S/2) \max_i m_{i,t}$ and $\min_i p_{i,t} = (1 - S/2) \min_i m_{i,t}$, where S is the proportional effective spread. These assumptions permit Corwin and Schultz (2012) to derive the expectations of the range in equations (26) and (27) in closed form and solve the two equations for S. In particular, they show that S can be estimated by

$$S_t = \frac{2(e^{\alpha_t} - 1)}{1 + e^{\alpha_t}}, \quad \alpha_t = \frac{\sqrt{2(\hat{r}_t^2 + \hat{r}_{t+1}^2)} - \sqrt{\hat{r}_t^2 + \hat{r}_{t+1}^2}}{3 - 2\sqrt{2}} - \sqrt{\frac{\hat{r}_{t:t+1}^2}{3 - 2\sqrt{2}}} \tag{28}$$

Note that Corwin and Schultz (2012) estimate the proportional effective spread S rather than the log spread s that I focus on in this paper. The two are related by $s = 2\log(1 + S/2)$, and the difference is very small for small values of S and does not materially affect any simulation or empirical results reported later in the paper.

The estimator S_t is not guaranteed to be non-negative. Corwin and Schultz (2012) suggest to either censor S_t at zero before calculating the average over T days:

$$CS_T^{(1)} = \frac{1}{T-1}\sum_{t=0}^{T-1} \max\{S_t, 0\}, \tag{29}$$

or to censor at zero the average S_t:

$$CS_T^{(2)} = \max\left\{\frac{1}{T-1}\sum_{t=0}^{T-1} S_t, 0\right\}. \tag{30}$$

It is not hard to see that $CS_T^{(1)}$ cannot be a consistent estimator of S, as the censoring before averaging produces an asymptotic bias.

Now the Corwin and Schultz (2012) measures, while simple, require two strong assumptions that I do not want to make here: continuous trading and zero overnight returns.[3]; in OTC markets, the daily number of transactions is typically much smaller than in equity markets, where assuming large n is perfectly plausible. Moreover, relying solely on the range means throwing away a lot of data, so I develop measures that use all available transaction data. My range-based estimators therefore use only the daily range in equation (26) and work for any finite n.

[3]Corwin and Schultz (2012) suggest a simple correction for the overnight return, but the correction does not eliminate the overnight return problem completely and the estimator remains generally biased and inconsistent. Abdi and Ranaldo (2016) propose a solution: they suggest using the statistics $(c_t - (\eta_{t+1} + \eta_t)/2)^2$ and $(\eta_t - \eta_{t+1})^2$, where c_t denotes the closing transaction price on day t and $\eta_t = (\max_j p_{j,t} + \min_j p_{j,t})/2$ is the daily midrange. Although this is an ingenious solution, it will not work in my context, as the closing transaction price c_t is not known in the absence of timestamps. Replacing c_t with the closing efficient price $m_{0,t}$ does not solve the problem because it is precisely the half-spread component of c_t that allows Abdi and Ranaldo (2016) to identify s.

I continue with the assumptions stated in Section 2 and additionally assume that the innovations of the efficient price are normally distributed. The expectation of the squared range can then be approximated by simulation for any finite n and the SMM employed to consistently estimate s. I proceed as follows. Let $\boldsymbol{\theta} = (s, \sigma^2)'$ and let $\boldsymbol{p}_s^* = (p_{1s}^*, p_{2s}^*, ..., p_{ns}^*)'$ denote a random draw from model (1) given $\boldsymbol{\theta}$. Taking S independent draws, I approximate the expectation of $\hat{r}^2$ by

$$m_S(\boldsymbol{\theta}, n) = \frac{1}{S} \sum_{s=1}^{S} (\max_j p_{js}^* - \min_j p_{js}^*)^2. \tag{31}$$

The SMM estimator is then obtained by

$$\hat{\boldsymbol{\theta}}_T = \underset{\theta \in \mathbb{R}_{++}}{\arg\min}\, \boldsymbol{g}_T' \boldsymbol{g}_T, \tag{32}$$

where $\boldsymbol{g}_T = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{g}_t$, $\boldsymbol{g}_t = (g_{1t}, g_{2t})'$, $g_{1t}(\boldsymbol{\theta}, n) = \tilde{d}_t^2 - \mathrm{E}(\tilde{d}_t^2)$, and $g_{2t}(\boldsymbol{\theta}, n) = \hat{r}_t^2 - m_S(\boldsymbol{\theta}, n)$. The objective function $\boldsymbol{g}_T' \boldsymbol{g}_T$ must be minimized numerically under the restrictions that s and σ are nonnegative. The range-based estimator of s, which I denote by $ES_T^{(2)}$, is then given by $ES_T^{(2)} = \hat{\theta}_{1,T}$. It follows that if $T/S \to 0$ as $T \to \infty$, $ES_T^{(2)} \xrightarrow{p} s$ and the SMM estimator is asymptotically equivalent to the generalized method of moments (GMM) (see chapter 2 in Gourieroux and Monfort, 1996), and the usual GMM inference applies.

My final estimator follows naturally from the previous two. If the benchmark prices are observable, it is clearly desirable to use all three moment conditions at the same time. Formally, define $g_{3t}(\boldsymbol{\theta}, n) = \hat{d}_t^2 - \mathrm{E}(\hat{d}_t^2)$ and $\boldsymbol{g}_t = (g_{1t}, g_{2t}, g_{3t})'$, where g_{1t} and g_{2t} are previously given. The over-identified SMM estimator of $\boldsymbol{\theta}$ is given by

$$\tilde{\boldsymbol{\theta}}_T = \underset{\theta \in \mathbb{R}_{++}}{\arg\min}\, \boldsymbol{g}_T' \boldsymbol{W}_T \boldsymbol{g}_T \tag{33}$$

for some positive definite matrix $\boldsymbol{W}_T$. I follow the standard two-stage approach, whereby I first use $\boldsymbol{W}_T = \boldsymbol{I}$ to obtain a preliminary estimate $\hat{\boldsymbol{\theta}}_T$ and then use the optimal $\widehat{\boldsymbol{W}_T}$ (sample variance of $\boldsymbol{g}_t$ evaluated at $\hat{\boldsymbol{\theta}}_T$) in the second stage to obtain $\tilde{\boldsymbol{\theta}}_T$. My third estimator of s is then given by $ES_T^{(3)} = \tilde{\theta}_{1,T}$. Again, if $T/S \to 0$ as $T \to \infty$, $ES_T^{(3)} \xrightarrow{p} s$, and we obtain asymptotic equivalence with GMM.

4.1 Computational aspects

When the number of transactions is large, the previously described simulation-based estimation may be slow. The minimization must be done numerically and the evaluation of the objective function can be costly. Fortunately, simulation-based or analytical approximations for $\mathrm{E}(r_t^2)$ can be devised that significantly speed up computations.

Observe that the price process in equation (2) can be approximated by a process $\sigma W(t) + \frac{s}{2}q(t)$, where $W(t)$ is standard Brownian motion and $q(t)$ is a continuous-time process such that for any t, $q(t) = 1$ with a probability of $1/2$, and $q(t) = -1$ otherwise. Now due to continuity of Brownian motion, the range of $\sigma W(t) + \frac{s}{2}q(t)$ equals the range of $\sigma W(u)$ plus s. Thus, for large n, we can approximate the range of $p_{i,t}$ by the range of $m_{i,t}$ plus s:

$$\mathrm{E}[(\max_j p_{j,t} - \min_j p_{j,t})^2] \approx \mathrm{E}[(\sigma(\max_j z_j - \min_j z_j) + s)^2], \tag{34}$$

where z_j, $j = 0, ..., n$, is a discretized Brownian motion on $[0, 1]$. All that has to be simulated, then, is the expectation of the range and squared range of a discretized Brownian motion. This simulation needs to be done only once, before the SMM estimation begins, and not every time the objective function is evaluated. When n is large, this approximation leads to significant gains in computational speed.

The approximation can be further improved by using the decomposition of Christensen, Podolskij, and Vetter (2009), Lemma A.1, where the maximum of the efficient price is only taken over buyer-initiated transactions and the minimum over seller-initiated transactions when calculating the range of z in (34). Formally, let b_i, $i = 0, ..., n$ be an iid binary process independent of z, where $b_i = 1$ with a probability of $1/2$ and $b_i = -1$ otherwise. Given a sample path of b and z, the range of z is calculated over the set $I = \{(i,j)|b_i = 1, b_j = -1]\}$. The approximation then becomes

$$\mathrm{E}[(\max_j p_{j,t} - \min_j p_{j,t})^2] \approx \mathrm{E}[(\sigma \max_{(i,j)\in I}(z_i - z_j) + s)^2]. \tag{35}$$

As before, the range on the right-hand side of equation (35) needs to be simulated only once and not every time the objective function is evaluated.

But the simulation can be avoided altogether because accurate analytical approximations for the range of discretized Brownian motion in equation (34) exist. Using Lemma A.8 in Andersen, Dobrev, and Schaumburg (2013) together with equation (34) leads to the approximation

$$\mathrm{E}[(\max_j p_{j,t} - \min_j p_{j,t})^2] \approx (4\log 2)\sigma^2 + 2\sqrt{\frac{8}{\pi}}\sigma s + s^2 + \frac{\zeta(1/2)}{\sqrt{2\pi}}\left(\sqrt{\frac{8}{\pi}}\sigma^2 + \sigma s\right)\frac{4}{\sqrt{n}}, \tag{36}$$

where $\zeta(1/2)/\sqrt{2\pi} \approx -0.5826$.

To see how these approximations work, I plot expressions (34),(35), and (36) together with the true value $\mathrm{E}[(\max_j p_{j,t} - \min_j p_{j,t})^2]$ in Figure 1 for different values of n. I find that all approximations are generally quite close to the true value for $n > 1000$. Interestingly, there is virtually no difference between the analytical approximations in equations (36) and (34); clearly, the first-order correction in Andersen, Dobrev, and Schaumburg (2013) works very well, even for small n. But both of these approximations are significantly upward biased when n is small. Fortunately, the approximation in equation (35), based on the idea of Christensen, Podolskij, and Vetter (2009), is significantly more accurate for all values of n and is very close to the true value when $n > 100$. Thus, it seems that in practice one should simulate $\mathrm{E}[(\max_j p_{j,t} - \min_j p_{j,t})^2]$ when n is small—say, less than 100—and use the approximation in equation (35) to speed up computations when $n > 100$. For very large values of n, one can avoid simulations altogether and use equation (36).

Another issue that obviously arises in practice is that the number of transactions is not the same every day. This issue poses no problems for my estimators. All one needs to do is to replace n in equations (3) and (36) with n_t, where n_t denotes the number of transactions on day t. Similarly, in the SMM estimation, one would simply simulate the squared range with the appropriate n_t for each t and then take the average.

5 Stochastic volatility

The assumption of constant volatility I have maintained so far is a strong one, but it is easy to show that, under certain conditions, the estimator $ES_T^{(1)}$ is robust to stochastic volatility. It follows from the proof of Proposition 1 and the Law of Iterated Expectations that if $\mathrm{E}(\epsilon_{i,t}^2) = const$, $\hat{s}_t^2$ remains an unbiased estimator of s^2; see Section B of the Appendix. If, in addition, the long-run variance of u_s^2, where $u_{n(t-1)+i} = \epsilon_{i,t}$, $i = 1, ..., n$, $t = 1, ..., T$, goes to zero as $T \to \infty$, $\mathrm{Var}(1/T \sum_{t=1}^{T} \hat{s}_t^2) \to 0$ and hence $ES_T^{(1)} \xrightarrow{p} s$. These conditions allow for long-memory in volatility and the so-called leverage effect, i.e. the correlation between volatility and the efficient price innovations. Now, when the variance is nonstationary or deterministic (and time-varying), $\mathrm{E}(\hat{s}_t^2) \neq s_t$ in general, and $ES_T^{(1)}$ can be asymtotically biased and inconsistent. The bias is a function of the entire path of σ, and it cannot be evaluated analytically in the absence of timestamps.

Unlike $ES_T^{(1)}$, the range-based estimators are not generally consistent in the presence of time-varying volatility. The expectation of the range cannot be easily expressed in terms of moments of the volatility of the efficient price, and so the previously used argument does not apply.

6 Performance assessment with simulated data

To assess the performance of my estimators $ES_T^{(i)}$, $i = 1, 2, 3$, in a controlled environment—both in absolute terms and relative to the other measures discussed in the paper—I run a Monte Carlo experiment. I set the daily integrated volatility of the efficient price (σ) to 35 basis points, which is approximately equal to the daily volatility of the 10-year Treasury futures price, and let the efficient price innovations follow a normal distribution. I vary the true effective spread (s) between 5 and 50 basis points, the number of daily transactions (n) between 10 and 250, and the number of days (T) in the sample between 25 and 250. Recall that the absolute values of s and σ are not that important for the RMSE of my measures—what matters are their relative values. Each simulation is based on 10,000 Monte Carlo replications.

Table 2 reports the average effective spread obtained in the simulation together with the associated RMSE. Starting with the results for the baseline estimator $ES_T^{(1)}$, which are reported in the top two rows of each panel, I find that the bias of the estimator can be either positive or negative in small samples depending on the true effective spread. But as predicted by theory (see equation (8)), the bias does become negative for sufficiently large T before eventually converging to zero as $T \to \infty$. The RMSE of the estimator approaches zero at a rate that is broadly in line with $\sqrt{T}$ consistency.

Turning to the just-identified range-based estimator, $ES_T^{(2)}$, reported in rows 3 and 4 of each panel in Table 2, I find that the estimator exhibits a bias that can be either positive or negative depending on n, T, and s, but both the bias and the RMSE decline as $T \to \infty$, as expected. Comparing the performance of $ES_T^{(2)}$ with the baseline estimator $ES_T^{(1)}$, I find that the two estimators can perform quite differently. On the one hand, $ES_T^{(1)}$ does well when s is large and T is small; for example, when $s = 50$, $n = 50$, and $T = 50$, the RMSE of $ES_T^{(1)}$ is around three times smaller than that of $ES_T^{(2)}$. On the other hand, $ES_T^{(2)}$ works relatively well when s is small and T is large; for example, when $s = 5$, $n = 250$, and $T = 250$, the RMSE of $ES_T^{(1)}$ is more than three times larger than that of $ES_T^{(2)}$.

It is therefore not surprising that the over-identified estimator, $ES_T^{(3)}$, which combines the moment conditions underlying $ES_T^{(1)}$ and $ES_T^{(2)}$ using the optimal weighting matrix, generally performs the best. The results reported in rows 5 and 6 of Table 2 show that the estimator is typically the most precise in terms of RMSE, except when T is very small.

The Corwin and Schultz (2012) estimators $CS_T^{(1)}$ and $CS_T^{(2)}$ perform markedly different from my estimators and from each other. The $CS_T^{(1)}$ estimator is significantly upward biased, especially for small values of s, and the bias does not disappear as $n \to \infty$, $T \to \infty$, or both. The bias remains simply due to the fact that S_t is censored at zero before averaging (see equation (29)). The alternative estimator, $CS_T^{(2)}$, which is obtained by censoring the average S_t, performs better when n is large, and its RMSE declines with T as expected. For small values of n, the estimator does not work very well, but recall that the Corwin-Schultz estimators are derived under the assumption of $n \to \infty$, so this is hardly surprising; in contrast, my range-based estimators work for any finite n as they rely on simulated moments. I should also stress that the Corwin-Schultz estimators only use intraday high and low prices, while my estimators use all available transaction prices and therefore more information.

Turning to the measures that require transactions to be ordered in time but do not need the trade direction—$Roll_T$ and RV_T^{all}—I find that in line with my analytical results, they generally perform better than my estimators, except when the number of transactions (n) is small (Panel A) and the effective spread is large; see, for example, the case of $n = 10$ and $s = 50$, where the $ES_T^{(3)}$ estimator outperforms both measures in terms of RMSE. The bias associated with RV_T^{all} tends to be quite large for small n, which compromises its RMSE, and it makes the estimator competitive with that of Roll only for large n.

Turning to the measures that require observing the trade direction but no ordering of transaction in time—$RS_T^{(1)}$, $RS_T^{(2)}$, and $RS_T^{(3)}$—I find that the Hong and Warga (2000) measure delivers the most accurate estimates across different sample sizes and values of s. Centering by the average transaction price ($RS_T^{(2)}$) produces biased estimates for small n as expected. For large n, the estimator outperforms its competitor $RS_T^{(1)}$, especially for small values of s, where it reaches accuracy similar to that of Hong and Warga (2000). Note that in line with the theoretical results, the bias—and, hence, RMSE—of $RS_T^{(2)}$ do not vanish as T increases, a property not shared by the other estimators. Finally, observing both the time stamps and trade direction—$RS_T^{(4)}$—produces the most accurate estimates uniformly across n, T, and s, as expected. Relative to my estimators that require neither time stamps nor trade direction, the improvement in efficiency is large: the RMSE is an order of magnitude smaller.

7 Performance assessment with small-cap equity data

Having explored the behavior of the various effective spread estimators in a controlled environment, I now repeat the exercise with empirical data. Ideally, I would like to employ data from an OTC market, since that is where I expect my measures would naturally find applications, but to the best of my knowledge, no time-stamped OTC trade and quote data are publicly available that would allow me to do this exercise. I therefore employ the widely-used TAQ data for selected NYSE-listed stocks; the TAQ data are time stamped to the second and contain information about m and q, so that the true effective spread can be readily computed and used as a benchmark.

7.1 Data and descriptive statistics

The universe of NYSE-listed stocks is too broad to consider here in full. Because my measures of effective spread would typically be applied to OTC-traded contracts, which tend to be less liquid and trade less frequently than exchange-traded instruments, I focus on stocks with small market capitalization. In particular, I take all stocks in the TAQ database that satisfy two criteria. The first is that the stock was included in the S&P Small-Cap 600 Index for the entire period between January 2, 2005, and December 31, 2014. The second is that there are trade and quote data available for this stock in the TAQ database for every trading day in this period. These criteria select 147 stocks; the list of TAQ tickers of these stocks is provided in Appendix C.

For each stock and day in my sample, I download from the Wharton Research Data Services (WRDS) the WRDS-derived trades files (WCT data sets), which contain trades matched with the prevailing National Best Bid and Offer quotes. I then filter the data and retain only those trades with trade times between 9:35 a.m. and 4:00 p.m., positive transaction price, positive prevailing mid-quote, and positive quoted spread. In addition, I drop all quotes where the prevailing quoted spread is greater than 50 times the median quoted spread for the same day, and all trades where the implied proportional effective spread is greater than 50 times the median proportional effective spread for the same day; these rules are similar to those proposed by Barndorff-Nielsen et al. (2008a).

Table 3 reports some descriptive statistics for the data, separately for five two-year periods that span my sample. The average daily number of trades varies between 1,000 and 2,000 for a typical stock-day. The average effective spread varies between 12 and 16 basis points, while the average daily realized volatility varies between 150 and 250 basis points. The average SNR, which I define here as the ratio of effective spread and realized volatility (s/σ), fluctuated between 6 and 9 percent. Thus, despite being small cap, the typical stock in my sample traded relatively frequently and with a fairly tight spread during my sample period. At the same time, there are stocks and trading days with relatively little trading and fairly large effective spreads as indicated by the 5th and 95th percentiles reported in the table.

My main empirical results are summarized in Panel A of Table 4. I run the effective spread estimations separately for each stock-month, stock-quarter, stock-half-year, and stock-year in the sample period and compare the estimates with the actual effective spreads observed for a given stock in a given time period. Specifically, I calculate the bias and RMSE associated with each estimator and the correlation of the estimated spread with the actual effective spreads calculated from TAQ.

I find that estimating the effective spread without timestamps and trade direction is very challenging. The infeasible estimators are significantly less biased and an order of magnitude more accurate in terms of RMSE than my estimators or the Corwin and Schultz (2012) ones; they are also much more closely correlated with the actual effective spreads. In line with theory, the over-identified estimator $ES_T^{(3)}$ is generally most accurate in terms of RMSE out of my three estimators, although $ES_T^{(2)}$ tends to be more closely correlated with the actual spread. The results do not improve as the number of observations used for estimation increases. As expected, all infeasible estimators deliver RMSE that is an order of magnitude lower than that of my feasible estimators. The estimator based on realized volatility performs remarkably well, exhibiting almost no bias and having significantly lower RMSE than the Roll measure.

The relatively poor performance of the feasible estimators should not come as a surprise: The average SNR for the 147 stocks in my sample is very small, and the analytical and simulation results presented previously clearly indicate that in such circumstances all feasible estimators struggle. To shed more light on the relationship between the SNR and RMSE, I perform the following experiment.

Rather than using the original transaction prices when computing effective spread estimates, I construct a new set of transaction prices $\tilde{p}$ where I set $\tilde{p}_{i,t} = m_{i,t} + 10(p_{i,t} - m_{i,t})$ for all i and t—that is, I artificially inflate the actual effective spread by a factor of 10. This procedure leaves the intraday volatility and time-series dynamics of the mid-quote unchanged, but it increases the SNR tenfold. I then reproduce the results reported in Panel A of Table 4 using the artificial transaction prices $\tilde{p}$ in place of the actual transaction prices p.

The results are reported in Panel B. I find that the relative performance of my estimators improves significantly. Although they are still upward biased, their RMSE is now much smaller relative to the actual spread. Also, while the two estimators that utilize timestamps (Roll and RV^{all}) still outperform my estimators, the differences in terms of RMSE have become smaller. The regression-based estimators that require trade direction perform better that either Roll or RV^{all}. Finally, the correlation between my estimates and the true spreads has increased significantly.

In addition to the experiment with the SNR, I study how the intraday number of transactions (n) affects performance. I do this by sampling sparsely from the set of transaction prices, retaining only every 10th observation on a given day for a given stock, and re-run all estimations on the sparsely sampled data. The results are reported in Panels C and D of Table 4; the former reports results based on the original data, while the latter shows results based on the artificial transaction prices previously described (inflated true effective spreads).

Starting with Panel C, I find that the performance of my estimators is largely unaffected by sparse sampling. This finding is in line with the theoretical result that n has only a second-order effect on the RMSE of these estimators. In contrast, the infeasible Roll and RV-based estimators exhibit a significant deterioration in precision as n decreases. Notably, the RV-based estimator now exhibits a significant upward bias and a substantial increase in RMSE. When I artificially inflate the actual effective spreads by a factor of 10 (Panel D), the differences between the RMSE of my estimators and the infeasible ones become even smaller.

In summary, the empirical results based on the 147 stock in my sample are largely consistent with the theoretical and simulation results reported previously. The key driver of the performance of my estimators is the SNR. To illustrate this finding graphically, Figure 2 plots the RMSE expressed as a fraction of the true spread separately for stock-months sorted into deciles by their SNR. The figure is based on the same data as Panel D in Table 4. Clearly, as the SNR increases, the performance improves, and gradually approaches the performance of the infeasible estimators. This is very much in line with the analytical results in Section 3.

8 Conclusion

In this paper, I have studied the problem of estimating transaction costs in the absence of timestamps and trade direction, which are data limitations that often occur in OTC datasets. Building on insights from the previous literature, I proposed several measures of the effective spread, studied their sampling properties, and compared their performance with some well-known, infeasible measures within the simple framework of Roll (1984); table 1 lists the various measures studied in this paper together with the data required to implement them. I corroborated my theoretical findings using a Monte Carlo simulation and additionally assessed the performance of my estimators in an empirical application to selected NYSE-listed small-cap stocks.

The theoretical, simulation-based, and empirical results show that the loss of information due to missing timestamps and trade direction is large. My estimators are suitable for measuring transaction costs in illiquid OTC markets, where effective spreads tend to be wide relative to the fundamental volatility and where trading is infrequent, but not necessarily in highly liquid exchange-based markets such as those for equities and futures contracts, where the opposite is generally true. But in those cases, accurate transaction timestamps are typically available and trade direction can be reliably inferred, so my estimators would not be necessary.

Throughout the paper, I have worked in the widely used framework of Roll (1984). Whereas the simplicity of this framework allows for straightforward analytical derivations, future work may consider more elaborate microstructure models, such as those by Huang and Stoll (1997); Madhavan, Richardson, and Roomans (1997); Bessembinder, Maxwell, and Venkataraman (2006); and Edwards, Harris, and Piwowar (2007). These models allow for a richer relationship between order flow and returns, and they relax some of the arguably restrictive assumptions of the Roll (1984) model. It would be interesting to explore whether these models can be reliably estimated when timestamps and trade direction are missing.

References

Abdi, F., and A. Ranaldo,2017, A simple estimation of bid-ask spreads from daily close, high, and low prices, *Review of Financial Studies*, **30** (12), 4437–4480.

Aït-Sahalia, Y., P. Mykland and L. Zhang, 2005, How often to sample a continuous-time process in the presence of market microstructure noise, *Review of Financial Studies*, **18**, 351-416.

Aït-Sahalia, Y., and J. Jacod, 2014, *High-frequency financial econometrics*, Princeton University Press.

Andersen, T. G., and T. Bollerslev, 1998, Answering the sceptics: Yes, standard volatility models do provide accurate volatility forecasts, *International Economic Review*, **39**(4), 885–905.

Andersen, T. G., D. Dobrev, and E. Schaumburg, 2013, Duration-based volatility estimation, working paper, Northwestern University.

Barndorff-Nielsen, O. E., and N. Shephard, 2002, Econometric analysis of realized volatility and its use in estimating stochastic volatility models, *Journal of the Royal Statistical Society, Series B*, **64**(2), 253–280.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard, 2008, Realized kernels in practice: Trades and quotes, *Econometrics Journal*, **4**(1), 1–32.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard, 2008, Designing realized kernels to measure the ex post variation of equity prices in the presence of noise, *Econometrica*, **76**(6), 1481–1536.

Benos, E., A. Wetherilt, and F. Zikes, 2013, The structure and dynamics of the UK credit default swap market, Financial Stability Paper No. 25, Bank of England.

Benos, E., R. Payne, and M. Vasios, 2016, Centralized trading, transparency and interest rate swap market liquidity: Evidence from the implementation of the Dodd-Frank Act, Staff Working Paper, No. 580, Bank of England.

Benos, E., and F. Zikes, 2016, Liquidity determinants in the UK gilt market, Working Paper No. 600, Bank of England.

Bessembinder, H., W. Maxwell, and K. Venkataraman, 2006, Market transparency, liquidity externalities, and institutional trading costs in corporate bonds, *Journal of Financial Economics*, **82**, 251–288.

Biswas, G., S. Nikolova, and C. W. Stahel, 2014, The transaction costs of trading corporate credit, working paper, Securities and Exchange Commission.

Brandt, M. W. and F. X. Diebold, 2006, A no-arbitrage approach to range-based estimation of return covariances and correlations, *Journal of Business*, **79**, 61-74.

Chen, K., M. Fleming, J. Jackson, A. Li, and A. Sarkar, 2011, An analysis of CDS transactions: Implications for public reporting, Staff Report No. 517, Federal Reserve Bank of New York.

Chen, K., M. Fleming, J. Jackson, A. Li, and A. Sarkar, 2012, An analysis of OTC interest rate derivatives transactions: Implications for public reporting, Staff Report No. 557, Federal Reserve Bank of New York.

Choi, J. and Y. Huh, 2017, Customer liquidity provision: implications for corporate bond transaction costs, Finance and Economics Discussion Series 2017-116, Board of Governors of the Federal Reserve System.

Christensen, K., and M. Podolskij, 2007, Realized range-based estimation of integrated variance, *Journal of Econometrics*, **141**, 323–349.

Christensen, K., M. Podolskij, and M. Vetter, 2009, Bias-correcting the realized range-based variance in the presence of microstructure noise, *Finance and Stochastics*, **13**, 239–268.

Corwin, S., and P. Schultz, 2012, A simple way to estimate bid-ask spreads from daily high and low prices, *Journal of Finance*, **67**(2), 719–759.

Dobrev, D., 2007, Capturing volatility from large price moves: Generalized range theory and applications, working paper, Northwestern University.

Doornik, J. A., 2007, *Object-oriented matrix programming using Ox*, 3rd ed. London: Timberlake Consultants Press and Oxford: www.doornik.com.

Du, W., S. Gadgil, M. Gordy, and C. Vega, 2017, Counterparty risk and counterparty choice in the credit default swap market, working paper.

Duffie, D. and H. Zhu, forthcoming, Size discovery, *Review of Financial Studies*.

Edwards, A. K., L. E. Harris, and M. S. Piwowar, 2007, Corporate bond market transaction costs and transparency, *Journal of Finance*, **62**(3), 1421–1451.

Foucault, T., M. Pagano, and A. Roell, 2013, *Market liquidity: Theory, evidence and policy*, Oxford University Press.

Gourieroux, C., and A. Monfort, 1996, *Simulation-based econometric methods*, Oxford University Press.

Goyenko, R. Y., C. W. Holden, and C. A. Trzcinka, 2009, Do liquidity measure measure liquidity?, *Journal of Financial Economics*, **92**, 153–181.

Hansen, P. R. and A. Lunde, 2006, Realized variance and market microstructure noise (with discussion), *Journal of Business and Economic Statistics*, **24**, 127218.

Harris, L., 2015, Transaction costs, trade throughs, and riskless principle trading in corporate bond markets, working paper, USC Marshall School of Business.

Hong, G., and A. Warga, 2000, An empirical study of bond market transactions, *Financial Analyst Journal*, **56**, 32–46.

Huang, R., and H. Stoll, 1997, The components of the bid-ask spread: A general approach, *Review of Financial Studies*, **10**, 995-1034.

Jankowitsch, R., A. Nashikkar, and M. G. Subrahmanyam, 2011, Price dispersion in OTC markets: A new measure of liquidity, *Journal of Banking and Finance*, **35**, 343–357.

Kalnina, I. and O. Linton, 2008, Estimating quadratic variation in the presence of endogenous and diurnal microstructure noise," *Journal of Econometrics*, **147**, 47–59.

Lee, C. M. C., and M. J. Ready, 1991, Inferring trade direction from intraday data, *Journal of Finance*, **46**(2), 733–746.

Madhavan, A., M. Richardson, and M. Roomans, 1997, Why do securities prices change: A transactions level analysis of NYSE-listed stocks, *Review of Financial Studies*, **10**, 1035–1064.

Parkinson, M., 1980, The extreme value method for estimating the variance of the rate of return, *Journal of Business*, **53**, 61-65.

Roll, R., 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, *Journal of Finance*, **39**, 1127–1139.

Schultz, P., 2001, Corporate bond trading costs: A peak behind the curtain, *Journal of Finance*, **56**(2), 677–698.

Warga, A., 1991, Corporate bond price discrepancies in the dealer and exchange markets, *Journal of Fixed Income*, **1**, 7–16.

Zhang, L., P. Mykland, and Y. Aït-Sahalia, 2005, Tale of two time scales: Determining integrated volatility with noisy high-frequency data, *Journal of the American Statistical Association*, **100**, 1394-1411.

A Proof of Proposition 1

Dropping the subscript t to simplify notation, we have

$$\begin{aligned}\hat{s}^2 - s^2 &= 2(3\tilde{d} - \hat{d}) - s^2, && (37)\\ &= \frac{3s^2}{2}\left[\left(\frac{1}{n-1}\sum_{i=1}^{n}(q_i - \bar{q})^2\right) - 1\right] \\ &\quad + 2s\left[\frac{3}{n-1}\sum_{i=1}^{n}(m_i - \bar{m})(q_i - \bar{q}) - \frac{1}{n}\sum_{i=1}^{n}(m_i - m_0)q_i\right] \\ &\quad + 2\left[\frac{3}{n-1}\sum_{i=1}^{n}(m_i - \bar{m})^2 - \frac{1}{n}\sum_{i=1}^{n}(m_i - m_0)^2\right] && (38)\\ &=: A_n + B_n + C_n. && (39)\end{aligned}$$

By construction, $\mathrm{E}(A_n) = \mathrm{E}(B_n) = \mathrm{E}(C_n) = 0$, and it is easy to show that $\mathrm{E}(A_nB_n) = \mathrm{E}(A_nC_n) = \mathrm{E}(B_nC_n) = 0$ because $\mathrm{E}(m_iq_j) = 0$ for all i and j. Thus, $\mathrm{Var}(\hat{s}^2) = \mathrm{E}(A_n^2) + \mathrm{E}(B_n^2) + \mathrm{E}(C_n^2)$. It is clear from the equation above that $\hat{s}^2$ does not depend on m_0, so we will set it to zero to simplify notation.

Starting with $\mathrm{E}(A_n^2)$, write

$$A_n^2 = \frac{9s^4}{4}\left[\frac{1}{(n-1)^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(q_i - \bar{q})^2(q_j - \bar{q})^2 - \frac{2}{n-1}\sum_{i=1}^{n}(q_i - \bar{q})^2 + 1\right]. \qquad (40)$$

Because $\mathrm{E}(q_iq_j) = 0$ if $i \neq j$ and $q_i^2 \equiv 1$, we have

$$\begin{aligned}\mathrm{E}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}(q_i - \bar{q})^2(q_j - \bar{q})^2\right) &= n^2 - 2\mathrm{E}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}q_iq_j\right) + \frac{1}{n^2}\mathrm{E}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}q_iq_jq_kq_l\right), && (41)\\ &= n^2 - 2n + 3 - \frac{2}{n}. && (42)\end{aligned}$$

This, together with $\mathrm{E}\left(\sum_{i=1}^{n}(q_i-\bar{q})^2\right)=n-1$, gives after some algebra

$$\mathrm{E}(A_n^2)=\frac{9s^4}{2n(n-1)}. \tag{43}$$

Turning to $\mathrm{E}(B_n^2)$, write

$$\begin{aligned} B_n^2 &= 4s^2\left[\frac{9}{(n-1)^2}-\frac{6}{n(n-1)}+\frac{1}{n^2}\right]\sum_{i=1}^{n}\sum_{j=1}^{n} m_i m_j q_i q_j \\ &+ 4s^2\left[\frac{6}{n^2(n-1)}-\frac{18}{n(n-1)^2}\right]\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} m_i m_j q_j q_k \\ &+ \frac{36s^2}{n^2(n-1)^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} m_i m_j q_k q_l. \end{aligned} \tag{44}$$

Because $\mathrm{E}(\epsilon_i\epsilon_j)=0$ if $i\neq j$,

$$\mathrm{E}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} m_i m_j q_i q_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{E}(m_i m_j)\mathrm{E}(q_i q_j)=\sum_{i=1}^{n}\mathrm{E}(m_i^2)q_i^2=\sum_{i=1}^{n}\sum_{p=1}^{i}\mathrm{E}(\epsilon_p^2)=\frac{\sigma^2}{2}(n+1). \tag{45}$$

Similarly, $\mathrm{E}\left(\sum_i\sum_j\sum_k m_i m_j q_j q_k\right)=\sum_i\sum_j \mathrm{E}(m_i m_j)$ and $\mathrm{E}\left(\sum_i\sum_j\sum_k\sum_l m_i m_j q_k q_l\right)=n\sum_i\sum_j \mathrm{E}(m_i m_j)$. Thus, it remains to derive $\sum_i\sum_j \mathrm{E}(m_i m_j)$. The case $i=j$ follows from above, so we focus on the case when $i\neq j$:

$$\begin{aligned} \sum_{\substack{i=1\\ i\neq j}}^{n}\sum_{j=1}^{n}\mathrm{E}(m_i m_j) &= 2\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathrm{E}(m_i m_j), && (46)\\ &= 2\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathrm{E}\left(\sum_{p=1}^{i}\left(\sum_{r=1}^{i}+\sum_{r=i+1}^{j}\right)\epsilon_p\epsilon_r\right), && (47)\\ &= 2\sum_{i=1}^{n}\sum_{j=i+1}^{n}\sum_{p=1}^{i}\mathrm{E}(\epsilon_p^2), && (48)\\ &= \frac{2\sigma^2}{n}\sum_{i=1}^{n}\sum_{j=i+1}^{n} i, && (49)\\ &= \sigma^2 n(n+1)-\frac{1}{3}\sigma^2(n+1)(2n+1). && (50) \end{aligned}$$

Plugging (44) and (49) into the expectation of (43) and simplifying gives

$$\mathrm{E}(B_n^2)=\frac{2s^2\sigma^2(2n^2+3n+1)}{n^2(n-1)}. \tag{51}$$

Finally, we derive $\mathrm{E}(C_n^2)$. Write

$$C_n^2 = 4\left[\frac{9}{(n-1)^2} - \frac{6}{n(n-1)} + \frac{1}{n^2}\right]\sum_{i=1}^{n}\sum_{j=1}^{n} m_i^2 m_j^2$$
$$+ 4\left[\frac{6}{n^2(n-1)} - \frac{18}{n(n-1)^2}\right]\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} m_i m_j m_k^2$$
$$+ \frac{36}{n^2(n-1)^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} m_i m_j m_k m_l. \tag{52}$$

We focus on the last term because the other two terms follow from the derivation of this term. Observe that

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} m_i m_j m_k m_l = {} & \sum_{i=1}^{n} m_i^2 + 3\sum_{\substack{i=1 \\ i\neq j}}^{n}\sum_{j=1}^{n} m_i^2 m_j^2 + 4\sum_{\substack{i=1 \\ i\neq j}}^{n}\sum_{j=1}^{n} m_i^3 m_j \\ & + 12\sum_{\substack{i=1 \\ i<j<k}}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} m_i^2 m_j m_k + 12\sum_{\substack{i=1 \\ i<j<k}}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} m_i m_j^2 m_k \\ & + 12\sum_{\substack{i=1 \\ i<j<k}}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} m_i m_j m_k^2 + 24\sum_{\substack{i=1 \\ i<j<k<l}}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} m_i m_j m_k m_l.\end{aligned} \tag{53}$$

To save space, we derive here only the expectation of the last term; the other terms follow using the same approach:

$$\begin{aligned} & \mathrm{E}\left(\sum_{\substack{i=1 \\ i<j<k<l}}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} m_i m_j m_k m_l\right) \\ & = \mathrm{E}\left(\sum_{i=1}^{n}\sum_{j=i+1}^{n}\sum_{k=j+1}^{n}\sum_{l=k+1}^{n}\sum_{p=1}^{i}\sum_{r=1}^{j}\sum_{s=1}^{k}\sum_{t=1}^{l}\epsilon_p\epsilon_r\epsilon_s\epsilon_t\right), \qquad (54) \\ & = \mathrm{E}\left(\sum_{i=1}^{n}\sum_{j=i+1}^{n}\sum_{k=j+1}^{n}\sum_{l=k+1}^{n}\sum_{p=1}^{i}\left(\sum_{r=1}^{i} + \sum_{r=i+1}^{j}\right)\left(\sum_{s=1}^{i} + \sum_{s=i+1}^{j} + \sum_{s=j+1}^{k}\right)\right. \\ & \qquad \left. \times \left(\sum_{t=1}^{i} + \sum_{t=i+1}^{j} + \sum_{t=j+1}^{k} + \sum_{t=k+1}^{l}\right)\epsilon_p\epsilon_r\epsilon_s\epsilon_t\right), \qquad (55) \\ & = \sum_{i=1}^{n}\sum_{j=i+1}^{n}\sum_{k=j+1}^{n}\sum_{l=k+1}^{n}\left[\mathrm{E}\left(\sum_{p=1}^{i}\sum_{r=1}^{i}\sum_{s=1}^{i}\sum_{t=1}^{i}\epsilon_p\epsilon_r\epsilon_s\epsilon_t\right) + 3\mathrm{E}\left(\sum_{p=1}^{i}\sum_{r=1}^{i}\sum_{s=i+1}^{j}\sum_{t=i+1}^{j}\epsilon_p\epsilon_r\epsilon_s\epsilon_t\right)\right.\end{aligned}$$

$$+\mathrm{E}\left(\sum_{p=1}^{i}\sum_{r=1}^{i}\sum_{s=j+1}^{k}\sum_{t=j+1}^{k}\epsilon_p\epsilon_r\epsilon_s\epsilon_t\right)\Bigg], \tag{56}$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\sum_{k=j+1}^{n}\sum_{l=k+1}^{n} 3\sigma^4 i^2 + \kappa i + 3\sigma^4 i(j-i) + \sigma^4 i(k-j), \tag{57}$$

$$= \frac{\sigma^4}{3}n^4 + \left(\sigma^4 + \frac{\kappa}{5}\right)n^3 + \left(\frac{13\sigma^4}{12} + \frac{\kappa}{2}\right)n^2 + \left(\frac{\sigma^4}{2} + \frac{\kappa}{3}\right)n + \frac{\sigma^4}{12} - \frac{\kappa}{30}\frac{1}{n}, \tag{58}$$

where $\kappa = \mathrm{E}(\epsilon_1^4) - 3\sigma^4$. Above, we use the fact that

$$\mathrm{E}\left(\sum_{p=1}^{i}\sum_{r=1}^{i}\sum_{s=1}^{i}\sum_{t=1}^{i}\epsilon_p\epsilon_r\epsilon_s\epsilon_t\right) = \sum_{p=1}^{i}\sum_{r=1}^{i}\sum_{s=1}^{i}\sum_{t=1}^{i}\mathrm{E}(\epsilon_p\epsilon_r\epsilon_s\epsilon_t), \tag{59}$$

$$= 3\sum_{\substack{p=1 \\ p\neq r}}^{i}\sum_{r=1}^{i}\mathrm{E}(\epsilon_p^2\epsilon_r^2) + \sum_{p=1}^{i}\mathrm{E}(e_p^4), \tag{60}$$

$$= \frac{1}{n^2}(3\sigma^4 i^2 + \kappa i), \tag{61}$$

and

$$\mathrm{E}\left(\sum_{p=1}^{i}\sum_{r=1}^{i}\sum_{s=i+1}^{j}\sum_{t=i+1}^{j}\epsilon_p\epsilon_r\epsilon_s\epsilon_t\right) = \mathrm{E}\left(\sum_{p=1}^{i}\sum_{r=1}^{i}\epsilon_p\epsilon_r\right)\mathrm{E}\left(\sum_{s=i+1}^{j}\sum_{t=i+1}^{j}\epsilon_s\epsilon_t\right), \tag{62}$$

$$= \left(\sum_{p=1}^{i}\mathrm{E}(\epsilon_p^2)\right)\left(\sum_{r=i+1}^{j}\mathrm{E}(\epsilon_r^2)\right), \tag{63}$$

$$= \frac{1}{n^2}\sigma^4 i(j-i). \tag{64}$$

The expectation of the other terms in C_n^2 can be obtained analogously. We obtain

$$\mathrm{E}(C_n^2) = \frac{2(2\sigma^4 n + \sigma^4 + 2\kappa)(2n^3 + 7n^2 + 7n + 2)}{15n^3(n-1)}. \tag{65}$$

The variance of $\hat{s}^2$ then follows after some algebra. ■

B Stochastic volatility

Clearly, $\mathrm{E}(A_n) = \mathrm{E}(B_n) = 0$ even under time-varying volatility because q and m are independent. Now for C_n, we have after some algebra

$$\mathrm{E}(C_n|\{\sigma_i\}_{i=1}^n) = \left[\frac{1}{n} + \frac{1}{n-1} - \frac{1}{n(n-1)}\right] \sum_{i=1}^{n} \sum_{p=1}^{i} \mathrm{E}(\epsilon_p^2|\sigma_p^2) - \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=i+1}^{n} \sum_{p=1}^{j} \mathrm{E}(\epsilon_p^2|\sigma_p^2) \tag{66}$$

Now, because $\epsilon_i = \frac{\sigma_i}{\sqrt{n}} z_i$, where z_i is *iid* with zero mean and unit variance, $\mathrm{E}[\mathrm{E}(\epsilon_p^2|\sigma_p^2)] = \frac{1}{n}\mathrm{E}(\sigma_i^2)$, and if $\mathrm{E}(\sigma_i^2) = const$, by the LIE, $\mathrm{E}[\mathrm{E}(C_n|\{\sigma_i\}_{i=1}^n)] = 0$. Thus, $\mathrm{E}(\hat{s}^2) = s^2$.

C List of TAQ tickers used in the empirical illustration

AIR, ABM, CHE, PNK, GFF, IVC, LZB, MCS, MYE, NJR, NWN, OXM, PNY, AGYS, KWR, AWR, SWX, SMP, RGR, WDFC, WGO, WWW, SKYW, JJSF, HTLD, CSH, LNN, MLI, TG, DGII, FBP, SIGI, VICR, ACAT, BHE, ETH, SONC, IO, PRGS, SMRT, UIL, CDI, LDL, APOG, UFPI, SXI, CBM, PKE, AXE, CBR, UTEK, CKP, GNCMA, SWM, KLIC, MW, BRC, ESIO, B, CTS, FMBI, UBSI, NSIT, SGY, CKH, SSD, BBOX, GPI, WTS, CNMD, MRCY, RLI, COHU, DEL, LG, POOL, ASGN, HSII, VSAT, AVA, FWRD, PSEM, HLIT, CRY, KOPN, BELFB, ROG, GCO, BRKS, HAE, BGG, MSCC, SRDX, CENX, DCOM, EXAR, STC, EE, MINI, EME, BPFH, RTEC, VECO, LNCE, FRED, CW, PLCE, DSPG, HVT, ICUI, KNX, WTFC, CUB, IART, FUL, SHLM, AIN, PZZA, JCOM, CCRN, BRKL, EPIQ, LXP, GB, PJC, MTH, DAKT, WPP, ITG, MMSI, LFUS, HIBB, SCSC, PRA, WRLD, EPR, SAFM, VVI, PVTB, CBU, ALE, AMED, PKY, SAH, NP, MTSC, SUP.

D Figures and Tables

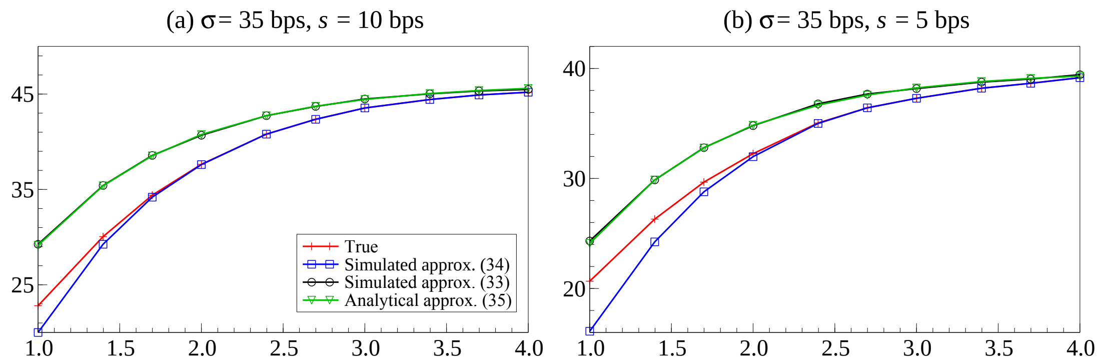

Figure 1: Expected squared range of p and its approximations as a function of $\log_{10} n$. The line labeled "True" shows the true expectation $\mathrm{E}[(\max_j p_{j,t} - \min_j p_{j,t})^2]$, "Simulated approx. (34)" shows the right-hand side of (35), "Simulated approx. (33)" shows the right-hand side of (34), and "Analytical approx. (35)" shows the right-hand side of (36). The various expectations are approximated by simulation with 100,000 replications. The efficient price innovations are normally distributed with a volatility of 35 bps.

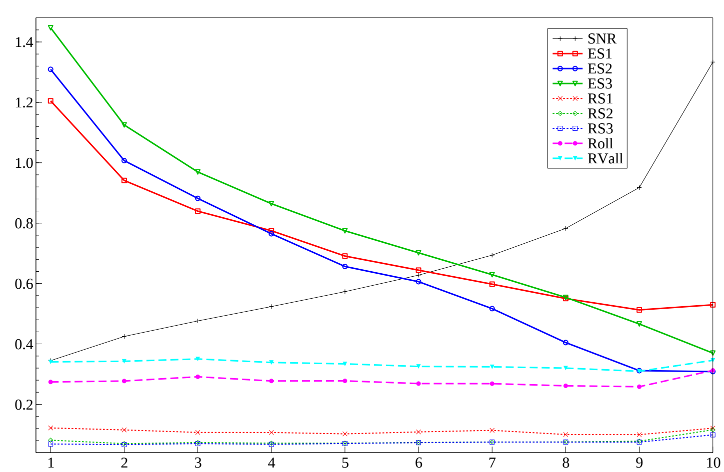

Figure 2: Average RMSE expressed as a fraction of the true effective spread for deciles based on the signal-to-noise (SNR) ratio. Every month, the stocks in the sample are sorted into deciles by their SNR. The RMSE for each stock-month decile is then calculated by averaging across all stock-month observations in the decile.

	Trade Direction	Time-stamps	Benchmark Price	Closed Form
$ES^{(1)}$	no	no	yes	yes
$ES^{(2)}$	no	no	no	no
$ES^{(3)}$	no	no	yes	no
$CS^{(1)}$	no	no	no	yes
$CS^{(2)}$	no	no	no	yes
$Roll$	no	yes	no	yes
RV_{all}	no	yes	no	yes
$RS^{(1)}$	yes	no	yes	yes
$RS^{(2)}$	yes	no	no	yes
$RS^{(3)}$	yes	no	no	yes
$RS^{(4)}$	yes	yes	no	yes

Table 1: Estimators of effective spread studied in this paper. The table shows the data required to implement the estimators and whether they are available in closed form.

	$s = 50$ bps				$s = 20$ bps				$s = 10$ bps				$s = 5$ bps			
T	25	50	100	250	25	50	100	250	25	50	100	250	25	50	100	250
A. $n = 10$																
$ES_T^{(1)}$	49.94	49.90	49.96	50.00	18.88	19.38	19.68	19.93	9.46	9.27	9.01	9.27	7.07	6.26	5.53	4.87
	[3.59]	[2.57]	[1.79]	[1.13]	[7.38]	[5.18]	[3.40]	[2.03]	[7.88]	[6.71]	[5.72]	[4.28]	[7.63]	[6.40]	[5.42]	[4.41]
$ES_T^{(2)}$	49.67	49.89	49.95	49.99	21.33	21.24	21.18	21.11	9.33	9.45	9.66	9.90	6.66	5.96	5.40	4.78
	[3.22]	[2.13]	[1.50]	[0.94]	[5.69]	[4.79]	[3.95]	[3.04]	[7.02]	[6.35]	[5.51]	[3.98]	[6.75]	[5.94]	[5.10]	[3.88]
$ES_T^{(3)}$	50.03	50.04	50.02	50.02	19.33	19.62	19.81	19.91	10.58	10.10	9.96	9.96	7.17	6.01	5.30	4.81
	[2.92]	[2.00]	[1.43]	[0.90]	[6.18]	[4.44]	[3.12]	[1.98]	[6.70]	[5.37]	[4.05]	[2.47]	[6.78]	[5.31]	[4.20]	[3.19]
$CS_T^{(1)}$	31.91	31.93	31.96	31.90	13.76	13.69	13.72	13.71	10.23	10.20	10.19	10.22	9.32	9.34	9.33	9.32
	[18.81]	[18.42]	[18.22]	[18.17]	[7.00]	[6.70]	[6.48]	[6.37]	[2.68]	[1.88]	[1.33]	[0.87]	[5.00]	[4.69]	[4.51]	[4.39]
$CS_T^{(2)}$	24.56	24.64	24.63	24.59	1.96	1.13	0.63	0.21	0.36	0.08	0.01	0.00	0.19	0.03	0.00	0.00
	[26.62]	[25.94]	[25.65]	[25.53]	[18.34]	[18.99]	[19.42]	[19.80]	[9.73]	[9.93]	[9.99]	[10.00]	[4.90]	[4.98]	[5.00]	[5.00]
Roll	49.83	49.95	50.00	50.02	19.86	19.93	19.97	19.98	9.65	9.85	9.92	9.96	4.46	4.50	4.62	4.79
	[4.22]	[2.96]	[2.08]	[1.32]	[2.39]	[1.67]	[1.17]	[0.74]	[2.80]	[1.87]	[1.27]	[0.78]	[3.34]	[2.78]	[2.19]	[1.43]
RV_{all}	52.37	52.39	52.39	52.40	25.36	25.38	25.40	25.39	18.56	18.56	18.57	18.57	16.41	16.43	16.43	16.43
	[3.04]	[2.74]	[2.57]	[2.48]	[5.47]	[5.44]	[5.43]	[5.40]	[8.60]	[8.58]	[8.58]	[8.58]	[11.44]	[11.44]	[11.43]	[11.43]
$RS_T^{(1)}$	50.00	50.00	50.01	50.00	20.03	20.01	19.99	20.00	9.98	9.97	10.01	9.98	5.11	5.00	5.03	4.99
	[3.26]	[2.35]	[1.67]	[1.03]	[3.28]	[2.31]	[1.63]	[1.03]	[3.28]	[2.30]	[1.65]	[1.04]	[3.07]	[2.28]	[1.65]	[1.03]
$RS_T^{(2)}$	45.01	45.00	45.00	45.01	18.02	18.00	18.00	18.00	9.01	9.01	9.00	8.98	4.47	4.48	4.52	4.50
	[5.48]	[5.25]	[5.13]	[5.04]	[2.72]	[2.40]	[2.21]	[2.08]	[2.05]	[1.61]	[1.34]	[1.17]	[1.85]	[1.36]	[1.02]	[0.76]
$RS_T^{(3)}$	49.89	49.91	49.89	49.92	19.98	19.96	19.96	19.97	9.99	10.00	9.98	9.96	4.96	4.97	5.02	4.99
	[2.08]	[1.47]	[1.05]	[0.67]	[2.02]	[1.44]	[1.01]	[0.63]	[2.01]	[1.42]	[1.00]	[0.65]	[1.99]	[1.41]	[1.01]	[0.64]
$RS_T^{(4)}$	49.99	50.00	50.00	50.00	19.99	20.01	20.00	20.00	10.00	9.99	9.99	10.00	5.00	4.99	5.00	5.00
	[1.05]	[0.73]	[0.52]	[0.33]	[1.05]	[0.73]	[0.52]	[0.33]	[1.04]	[0.74]	[0.52]	[0.33]	[1.05]	[0.74]	[0.52]	[0.33]

Cont. on next page.

Cont. from previous page.

	$s = 50$ bps				$s = 20$ bps				$s = 10$ bps				$s = 5$ bps			
T	25	50	100	250	25	50	100	250	25	50	100	250	25	50	100	250
B. $n = 50$																
$ES_T^{(1)}$	49.93	49.98	49.98	49.99	19.21	19.60	19.79	19.93	9.23	9.07	9.12	9.44	6.51	5.90	5.22	4.64
	[2.16]	[1.53]	[1.07]	[0.68]	[5.77]	[3.83]	[2.57]	[1.54]	[7.14]	[6.12]	[5.07]	[3.59]	[6.88]	[5.80]	[4.90]	[4.04]
$ES_T^{(2)}$	48.79	49.65	49.92	49.98	19.67	20.02	19.70	19.49	10.43	10.11	10.46	10.43	5.50	4.64	4.84	5.00
	[7.64]	[4.58]	[2.11]	[0.44]	[3.99]	[3.40]	[2.80]	[2.25]	[4.04]	[3.33]	[2.67]	[1.93]	[4.08]	[3.15]	[2.30]	[1.42]
$ES_T^{(3)}$	49.95	50.04	50.01	49.99	19.76	19.71	19.74	19.85	11.55	10.71	10.57	10.28	6.34	5.11	5.05	5.08
	[1.35]	[0.94]	[0.67]	[0.42]	[4.02]	[3.29]	[2.42]	[1.53]	[4.70]	[3.54]	[2.49]	[1.39]	[4.81]	[3.31]	[2.13]	[1.26]
$CS_T^{(1)}$	44.32	44.38	44.37	44.41	21.48	21.51	21.49	21.51	15.52	15.56	15.55	15.55	13.25	13.22	13.25	13.24
	[8.04]	[6.91]	[6.31]	[5.87]	[4.35]	[3.25]	[2.51]	[1.98]	[6.50]	[6.06]	[5.80]	[5.66]	[8.82]	[8.51]	[8.39]	[8.30]
$CS_T^{(2)}$	41.71	41.77	41.78	41.80	12.81	12.82	12.76	12.80	4.81	4.21	3.79	3.59	2.44	1.68	1.14	0.66
	[10.70]	[9.53]	[8.90]	[8.46]	[9.67]	[8.58]	[7.97]	[7.49]	[7.10]	[6.91]	[6.84]	[6.71]	[4.42]	[4.17]	[4.24]	[4.47]
Roll	49.98	50.01	50.00	50.00	19.99	19.99	19.99	20.00	9.99	10.00	10.00	10.00	4.97	4.99	4.99	5.00
	[1.64]	[1.13]	[0.82]	[0.51]	[0.71]	[0.50]	[0.35]	[0.22]	[0.45]	[0.32]	[0.22]	[0.14]	[0.44]	[0.31]	[0.22]	[0.14]
RV_{all}	50.49	50.50	50.49	50.49	21.19	21.18	21.19	21.19	12.20	12.21	12.20	12.21	8.60	8.60	8.60	8.60
	[0.88]	[0.71]	[0.61]	[0.54]	[1.23]	[1.21]	[1.20]	[1.19]	[2.21]	[2.21]	[2.21]	[2.21]	[3.60]	[3.60]	[3.60]	[3.60]
$RS_T^{(1)}$	50.00	50.00	50.01	50.00	20.00	20.02	19.99	20.00	9.99	10.00	10.01	10.01	5.01	4.99	5.01	5.00
	[1.43]	[0.98]	[0.69]	[0.45]	[1.42]	[1.00]	[0.71]	[0.44]	[1.42]	[0.99]	[0.71]	[0.45]	[1.42]	[0.99]	[0.71]	[0.45]
$RS_T^{(2)}$	48.98	48.99	49.00	49.00	19.59	19.61	19.60	19.60	9.80	9.80	9.80	9.80	4.90	4.90	4.90	4.90
	[1.33]	[1.17]	[1.09]	[1.04]	[0.92]	[0.70]	[0.58]	[0.47]	[0.84]	[0.61]	[0.45]	[0.32]	[0.81]	[0.60]	[0.41]	[0.28]
$RS_T^{(3)}$	49.99	49.99	50.00	50.00	19.99	20.01	20.00	20.00	10.00	10.00	10.00	10.00	5.00	5.00	5.00	5.00
	[0.82]	[0.58]	[0.41]	[0.26]	[0.83]	[0.58]	[0.42]	[0.26]	[0.83]	[0.58]	[0.41]	[0.26]	[0.82]	[0.60]	[0.41]	[0.26]
$RS_T^{(4)}$	50.00	50.00	50.00	50.00	20.00	20.00	20.00	20.00	10.01	10.00	10.00	10.00	5.00	5.00	5.00	5.00
	[0.20]	[0.14]	[0.10]	[0.06]	[0.20]	[0.14]	[0.10]	[0.06]	[0.20]	[0.14]	[0.10]	[0.06]	[0.20]	[0.14]	[0.10]	[0.06]

Cont. on next page.

Cont. from previous page.

	$s = 50$ bps				$s = 20$ bps				$s = 10$ bps				$s = 5$ bps			
T	25	50	100	250	25	50	100	250	25	50	100	250	25	50	100	250
C. $n = 250$																
$ES_T^{(1)}$	49.98	49.96	49.98	49.99	19.35	19.66	19.87	19.94	9.22	9.12	9.09	9.45	6.53	5.74	5.16	4.73
	[1.86]	[1.32]	[0.94]	[0.59]	[5.33]	[3.62]	[2.35]	[1.47]	[7.04]	[5.98]	[4.95]	[3.40]	[6.71]	[5.68]	[4.83]	[3.98]
$ES_T^{(2)}$	49.19	49.73	49.99	50.01	18.94	18.99	18.97	19.12	10.30	10.48	10.68	10.78	5.75	5.40	5.17	5.05
	[6.36]	[3.77]	[1.22]	[0.31]	[4.58]	[3.98]	[3.54]	[3.00]	[3.67]	[3.25]	[2.97]	[2.70]	[3.82]	[2.93]	[1.98]	[1.20]
$ES_T^{(3)}$	49.88	49.92	49.94	49.95	19.47	19.50	19.62	19.74	11.70	11.27	10.89	10.54	6.82	6.01	5.51	5.29
	[0.98]	[0.69]	[0.48]	[0.31]	[3.94]	[3.09]	[2.24]	[1.38]	[4.54]	[3.53]	[2.62]	[1.66]	[4.76]	[3.24]	[1.88]	[0.98]
$CS_T^{(1)}$	49.42	49.54	49.52	49.48	25.37	25.41	25.37	25.39	18.73	18.70	18.74	18.72	15.80	15.80	15.77	15.77
	[5.90]	[4.21]	[2.99]	[1.93]	[7.01]	[6.25]	[5.80]	[5.56]	[9.53]	[9.11]	[8.94]	[8.80]	[11.36]	[11.07]	[10.90]	[10.83]
$CS_T^{(2)}$	47.72	47.85	47.82	47.78	18.65	18.76	18.65	18.70	9.41	9.17	9.13	9.10	5.50	4.92	4.57	4.46
	[7.02]	[5.18]	[3.98]	[3.05]	[6.80]	[4.83]	[3.55]	[2.44]	[6.06]	[4.62]	[3.37]	[2.24]	[5.18]	[3.93]	[3.04]	[2.09]
Roll	50.00	50.00	49.99	50.00	20.00	20.00	20.00	20.00	10.00	10.00	10.00	10.00	5.00	5.00	5.00	5.00
	[0.72]	[0.50]	[0.36]	[0.22]	[0.29]	[0.20]	[0.14]	[0.09]	[0.15]	[0.11]	[0.08]	[0.05]	[0.09]	[0.07]	[0.05]	[0.03]
RV_{all}	50.10	50.10	50.09	50.10	20.24	20.24	20.24	20.24	10.48	10.48	10.48	10.48	5.90	5.90	5.90	5.90
	[0.34]	[0.24]	[0.19]	[0.14]	[0.28]	[0.26]	[0.25]	[0.25]	[0.48]	[0.48]	[0.48]	[0.48]	[0.90]	[0.90]	[0.90]	[0.90]
$RS_T^{(1)}$	49.99	50.00	49.99	50.00	20.00	20.01	20.00	20.00	10.00	10.01	10.00	10.00	5.00	5.01	5.00	5.00
	[0.62]	[0.44]	[0.31]	[0.20]	[0.63]	[0.44]	[0.31]	[0.20]	[0.63]	[0.45]	[0.31]	[0.20]	[0.63]	[0.45]	[0.32]	[0.20]
$RS_T^{(2)}$	49.80	49.79	49.80	49.80	19.92	19.92	19.92	19.92	9.96	9.96	9.96	9.96	4.98	4.98	4.98	4.98
	[0.42]	[0.33]	[0.28]	[0.23]	[0.38]	[0.26]	[0.20]	[0.14]	[0.36]	[0.26]	[0.19]	[0.12]	[0.36]	[0.26]	[0.18]	[0.12]
$RS_T^{(3)}$	50.00	49.99	50.00	50.00	20.00	20.00	20.00	20.00	10.00	10.01	10.00	10.00	5.00	5.00	5.00	5.00
	[0.36]	[0.25]	[0.18]	[0.12]	[0.37]	[0.25]	[0.18]	[0.12]	[0.36]	[0.26]	[0.18]	[0.12]	[0.36]	[0.26]	[0.18]	[0.11]
$RS_T^{(4)}$	50.00	50.00	50.00	50.00	20.00	20.00	20.00	20.00	10.00	10.00	10.00	10.00	5.00	5.00	5.00	5.00
	[0.04]	[0.03]	[0.02]	[0.01]	[0.04]	[0.03]	[0.02]	[0.01]	[0.04]	[0.03]	[0.02]	[0.01]	[0.04]	[0.03]	[0.02]	[0.01]

Table 2: Simulation results. The table reports the mean effective spread obtained in the simulation and the associated RMSE in brackets. The efficient price innovations are normally distributed, and the daily integrated volatility of the efficient price is set to 35 bps. The results are based on 10,000 Monte Carlo replications.

	2005–6	2007–8	2009–10	2011–12	2013–14
A. Number of transactions					
Mean	975	2098	1947	1697	1757
Std. dev.	1023	2001	2635	2019	1836
5th percentile	174	409	283	228	268
95th percentile	2740	5831	5867	5248	4874
B. Effective spread (bps)					
Mean	12.6	13.1	15.6	13.2	13.0
Std. dev.	7.3	9.0	12.1	9.0	8.7
5th percentile	5.8	5.3	5.4	4.8	4.8
95th percentile	25.9	27.7	36.2	30.0	31.0
C. Realized volatility (bps)					
Mean	168.8	246.0	252.3	192.0	156.3
Std. dev.	67.7	151.8	133.4	92.3	64.4
5th percentile	86.6	89.7	100.6	84.6	80.9
95th percentile	293.4	559.1	498.0	362.4	272.5
D. Signal-to-noise ratio					
Mean	0.078	0.058	0.063	0.071	0.084
Std. dev.	0.046	0.025	0.030	0.036	0.047
5th percentile	0.039	0.030	0.031	0.033	0.040
95th percentile	0.144	0.103	0.116	0.136	0.169

Table 3: The descriptive statistics are calculated over all stock-days in a given two-year period. The effective spread and realized volatility were winsorized at the 99.5% level, separately for each stock, before pooling and calculating the stock-day descriptive statistics. The sample consists of 147 small-cap stocks over the period from January 2005 to December 2014, spanning 2,517 business days.

	Monthly				Quarterly				Semiannual				Annual			
	Mean	Bias	RMSE	Corr	Mean	Bias	RMSE	Corr	Mean	Bias	RMSE	Corr	Mean	Bias	RMSE	Corr
A. All transactions, original spreads																
True	13.52	-	-	-	13.52	-	-	-	13.52	-	-	-	13.52	-	-	-
$ES_T^{(1)}$	91.34	77.82	95.01	0.342	92.61	79.09	92.39	0.364	92.12	78.60	90.87	0.378	91.51	77.99	89.29	0.410
$ES_T^{(2)}$	79.14	65.62	85.59	0.403	86.31	72.79	97.08	0.472	86.56	73.04	95.88	0.456	88.57	75.05	94.96	0.434
$ES_T^{(3)}$	81.38	67.86	86.61	0.334	79.83	66.30	85.61	0.389	82.58	69.06	85.69	0.390	85.44	71.92	86.09	0.395
$CS_T^{(1)}$	83.82	70.30	84.28	0.515	83.79	70.27	80.89	0.564	83.74	70.22	79.54	0.569	83.69	70.17	77.40	0.572
$CS_T^{(2)}$	28.85	15.33	38.72	0.299	22.56	9.041	25.44	0.351	20.45	6.926	20.97	0.413	19.14	5.621	16.96	0.487
Roll	8.731	-4.789	6.007	0.955	8.813	-4.710	5.785	0.967	8.874	-4.646	5.599	0.969	8.988	-4.533	5.362	0.969
RV_{all}	13.16	-0.357	2.345	0.965	13.25	-0.270	2.154	0.969	13.34	-0.181	2.093	0.968	13.49	-0.027	2.046	0.967
$RS_T^{(1)}$	7.833	-5.687	10.73	0.455	7.106	-6.417	10.07	0.501	6.872	-6.648	9.688	0.542	6.770	-6.752	9.165	0.621
$RS_T^{(2)}$	11.10	-2.426	6.444	0.768	11.00	-2.520	5.317	0.841	10.93	-2.591	4.863	0.872	10.91	-2.617	4.507	0.893
$RS_T^{(3)}$	11.53	-1.993	5.517	0.827	11.52	-2.001	4.378	0.893	11.52	-2.000	3.908	0.916	11.52	-2.000	3.544	0.932
$RS_T^{(4)}$	9.429	-4.092	5.095	0.974	9.407	-4.116	5.082	0.977	9.352	-4.168	5.156	0.975	9.318	-4.204	5.171	0.970
B. All transactions, spreads multiplied by 10																
True	135.3	-	-	-	135.3	-	-	-	135.3	-	-	-	135.3	-	-	-
$ES_T^{(1)}$	207.4	72.13	89.04	0.937	209.5	74.24	87.35	0.957	210.8	75.51	87.60	0.964	212.9	77.63	89.68	0.960
$ES_T^{(2)}$	210.7	75.44	101.2	0.732	213.2	77.87	100.2	0.755	214.9	79.63	100.6	0.763	218.3	82.99	102.0	0.769
$ES_T^{(3)}$	256.4	121.1	136.6	0.775	243.6	108.3	122.1	0.808	238.2	102.9	114.5	0.837	232.0	96.69	106.0	0.866
$CS_T^{(1)}$	402.9	267.7	308.7	0.846	403.5	268.3	301.8	0.878	403.4	268.2	298.1	0.884	403.4	268.1	292.6	0.892
$CS_T^{(2)}$	380.7	245.4	289.0	0.841	381.1	245.8	279.7	0.884	381.0	245.8	275.6	0.893	380.9	245.7	270.3	0.903
Roll	107.6	-27.65	39.87	0.960	108.4	-26.88	37.79	0.965	109.1	-26.15	36.07	0.965	110.4	-24.88	33.84	0.965
RV_{all}	128.9	-6.359	22.04	0.971	129.8	-5.470	20.15	0.975	130.6	-4.612	19.63	0.974	132.2	-3.065	19.30	0.971
$RS_T^{(1)}$	128.9	-6.375	15.07	0.989	128.7	-6.627	13.41	0.994	128.0	-7.220	15.05	0.992	127.5	-7.762	16.45	0.988
$RS_T^{(2)}$	129.5	-5.793	13.82	0.993	129.3	-5.995	13.52	0.994	128.7	-6.570	15.25	0.992	128.2	-7.026	16.76	0.988
$RS_T^{(3)}$	132.7	-2.525	6.320	0.998	132.8	-2.526	5.120	0.999	132.7	-2.525	4.661	0.999	132.7	-2.525	4.272	0.999
$RS_T^{(4)}$	119.9	-15.41	20.62	0.995	119.5	-15.76	21.13	0.995	118.8	-16.41	23.14	0.992	118.5	-16.82	24.48	0.986

Cont. on next page.

Cont. from previous page.

	Monthly				Quarterly				Semiannual				Annual			
	Mean	Bias	RMSE	Corr	Mean	Bias	RMSE	Corr	Mean	Bias	RMSE	Corr	Mean	Bias	RMSE	Corr
C. Every 10th transaction, original spreads																
True	13.52	-	-	-	13.52	-	-	-	13.52	-	-	-	13.52	-	-	-
$ES_T^{(1)}$	89.34	76.44	93.50	0.339	90.63	77.72	90.95	0.365	90.03	77.13	89.39	0.376	89.46	76.55	87.85	0.406
$ES_T^{(2)}$	75.89	62.98	81.96	0.453	85.92	73.01	97.05	0.482	86.90	74.00	97.06	0.477	88.96	76.06	96.90	0.446
$ES_T^{(3)}$	81.84	68.94	85.59	0.379	81.57	68.66	85.21	0.412	84.25	71.35	85.46	0.424	87.23	74.33	85.91	0.444
$CS_T^{(1)}$	71.26	57.89	72.23	0.415	71.29	57.93	68.84	0.456	71.25	57.89	67.47	0.455	71.21	57.85	65.34	0.444
$CS_T^{(2)}$	18.65	5.28	29.92	0.184	12.07	-1.297	18.94	0.172	9.586	-3.775	15.74	0.199	7.744	-5.617	13.57	0.203
Roll	6.901	-6.003	9.607	0.621	6.820	-6.087	9.040	0.655	6.835	-6.068	8.702	0.675	6.911	-5.991	8.366	0.679
RV_{all}	31.28	18.37	22.59	0.870	31.52	18.61	22.63	0.877	31.73	18.82	22.72	0.877	32.08	19.18	22.87	0.874
$RS_T^{(1)}$	8.069	-5.295	11.18	0.422	7.079	-6.285	10.27	0.450	6.728	-6.633	9.868	0.480	6.473	-6.892	9.477	0.539
$RS_T^{(2)}$	11.07	-2.292	7.187	0.715	10.93	-2.440	5.619	0.807	10.84	-2.518	4.975	0.849	10.68	-2.681	4.672	0.873
$RS_T^{(3)}$	11.41	-1.951	6.533	0.759	11.37	-1.996	4.858	0.855	11.36	-1.999	4.167	0.893	11.36	-2.004	3.707	0.916
$RS_T^{(4)}$	8.597	-4.767	6.067	0.925	8.562	-4.802	5.905	0.952	8.505	-4.856	5.922	0.960	8.470	-4.895	5.928	0.957
D. Every 10th transaction, spreads multiplied by 10																
True	135.3	-	-	-	135.3	-	-	-	135.3	-	-	-	135.3	-	-	-
$ES_T^{(1)}$	206.3	72.66	91.13	0.933	208.6	74.88	88.97	0.956	209.8	76.12	89.04	0.964	211.9	78.21	90.94	0.962
$ES_T^{(2)}$	183.3	49.64	76.69	0.762	182.9	49.24	73.47	0.782	183.6	49.97	71.87	0.792	186.7	53.03	74.19	0.777
$ES_T^{(3)}$	209.7	76.05	90.98	0.855	202.7	69.01	82.66	0.874	199.7	66.08	77.21	0.899	197.2	63.53	71.75	0.926
$CS_T^{(1)}$	236.9	103.2	134.6	0.754	237.1	103.4	127.2	0.793	237.0	103.3	124.6	0.797	236.9	103.2	120.8	0.803
$CS_T^{(2)}$	198.7	65.01	109.9	0.712	198.3	64.58	94.80	0.782	198.2	64.53	90.49	0.797	198.1	64.40	85.80	0.810
Roll	161.5	27.84	41.81	0.968	162.9	29.23	40.13	0.978	164.0	30.33	41.24	0.978	166.0	32.36	43.53	0.975
RV_{all}	170.5	36.81	48.75	0.978	171.8	38.14	48.52	0.984	172.9	39.24	50.00	0.983	175.0	41.33	52.45	0.979
$RS_T^{(1)}$	127.7	-6.016	16.35	0.985	127.5	-6.225	13.40	0.992	126.8	-6.818	14.73	0.991	124.7	-9.049	18.19	0.987
$RS_T^{(2)}$	128.2	-5.454	13.54	0.992	128.1	-5.598	12.59	0.994	127.5	-6.155	14.26	0.992	125.5	-8.222	18.21	0.986
$RS_T^{(3)}$	131.3	-2.432	7.645	0.997	131.3	-2.433	5.646	0.998	131.2	-2.437	4.829	0.999	131.2	-2.453	4.396	0.999
$RS_T^{(4)}$	126.4	-7.264	12.29	0.995	126.2	-7.523	11.67	0.996	125.5	-8.174	13.96	0.994	125.1	-8.558	16.28	0.988

Table 4: Empirical results for the various effective spread estimators applied to 147 small-cap stocks between January 2005 and December 2014. The table reports the mean effective spread estimates (Mean), the bias (Bias) with respect to the actual effective spreads observed in the TAQ data (True), the RMSE, and the correlation between estimated and actual values of the effective spread (Corr). The results are obtained by pooling across stocks.