Journal of Financial Economics 161 (2024) 103916

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/finec

Efficient estimation of bid–ask spreads from open, high, low, and close prices☆

David Ardia[a], Emanuele Guidotti[b,*], Tim A. Kroencke[c,d]

[a] GERAD & Department of Decision Sciences, HEC Montréal, Canada
[b] Institute of Finance, Università della Svizzera Italiana, Switzerland
[c] FHNW School of Business, Switzerland
[d] Remaco Asset Management, Switzerland

ARTICLE INFO

Dataset link: Replication Package for "Efficient Estimation of Bid-Ask Spreads from Open, High, Low, and Close Prices"

Keywords:
Bid–ask spread
Trading frictions
Transaction costs

ABSTRACT

Popular bid–ask spread estimators are downward biased when trading is infrequent. Moreover, they consider only a subset of open, high, low, and close prices and neglect potentially useful information to improve the spread estimate. By accounting for discretely observed prices, this paper derives asymptotically unbiased estimators of the effective bid–ask spread. Moreover, we combine them optimally to minimize the estimation variance and obtain an efficient estimator. Through theoretical analyses, numerical simulations, and empirical evaluations, we show that our efficient estimator dominates other estimators from transaction prices, yields novel insights for measuring bid–ask spreads, and has broad applicability in empirical finance.

1. Introduction

The effective bid–ask spread measures the distance of observed transaction prices from the unobserved fundamental price, and it is a predominant measure of transaction costs in financial markets. The literature on measuring bid–ask spreads has proceeded along two complementary paths that focus on either high-frequency or low-frequency data. The high-frequency literature relies on trade and quote data to obtain an explicit proxy of the fundamental price and calculate the distance of transaction prices from it (Holden and Jacobsen, 2014; Stoikov, 2018; Hagströmer, 2021). The low-frequency literature introduces assumptions about the fundamental price to derive estimators from transaction prices only, without requiring any information about quotes (Roll, 1984; Hasbrouck, 2009; Corwin and Schultz, 2012; Abdi and Ranaldo, 2017).

While measures from trades and quotes are typically more accurate, low-frequency estimates are more readily available and are becoming increasingly popular due to the difficulties and costs of obtaining quote data for international markets, historical data samples, and asset classes other than stocks.[1] However, the estimators developed thus far rely on

☆ Dimitris Papanikolaou was the editor for this article. We are grateful to the editor and two anonymous referees for their very helpful comments. We thank Yakov Amihud, Jean-Philippe Bouchaud, Victor DeMiguel, Laurent Frésard, Björn Hagströmer, Angelo Ranaldo, Sebastian Stöckl, Erik Theissen, Dennis Umlandt, and conference and seminar participants at the Theory-based Empirical Asset Pricing Research Workshop (T-BEAR, 2023), the Annual Meeting of the European Finance Association (EFA, 2022), Annual Meeting of the Financial Management Association (FMA, 2022), International Conference on Quantitative Finance and Financial Econometrics (QFFE, 2022), New Zealand Finance Meeting (NZFM, 2021), Annual Financial Market Liquidity Conference (AFML, 2021), Annual Meeting of the German Finance Association (DGF, 2021), International Risk Management Conference (IRMC, 2021), Goethe University Frankfurt, HEC Montréal, and INSEAD business school for helpful suggestions. We also thank Christophe Pérignon and the Certification Agency for Scientific Code and Data (CASCaD) for validating the reproducibility of our work. We acknowledge financial support from IVADO and NSERC (grant RGPIN-2022-03767). This article is based on the first chapter of Emanuele Guidotti's Ph.D. thesis at the University of Neuchâtel.

* Corresponding author.
E-mail address: emanuele.guidotti@usi.ch (E. Guidotti).

[1] For instance, recent use cases of low-frequency estimators include: stock return predictability and asset pricing anomalies (McLean and Pontiff, 2016; Hou et al., 2018; Chen et al., 2018; Birru, 2018; Hua et al., 2019; Jacobs and Müller, 2020; Patton and Weller, 2020; Amihud and Noh, 2020; Chaieb et al., 2020); municipal and corporate bonds (Schwert, 2017; Bongaerts et al., 2017; Cai et al., 2019; Kaviani et al., 2020; Bali et al., 2021; Ding et al., 2022); bond funds (Goldstein et al., 2017; Choi et al., 2020); currency markets (Michaelides et al., 2019; Ranaldo and de Magistris, 2022); OTC derivatives (Loon and Zhong, 2016); interest rates (Ranaldo et al., 2021); monetary policy (Grosse-Rueschkamp et al., 2019); institutional trading costs (Eaton et al., 2021); investor behavior (Li et al., 2018); information and dark pools (Brogaard and Pan, 2021); and machine learning (Easley et al., 2020). See Table I.1 in the Internet Appendix for a survey.

https://doi.org/10.1016/j.jfineco.2024.103916
Received 31 January 2023; Received in revised form 3 August 2024; Accepted 5 August 2024
Available online 14 August 2024
0304-405X/

the assumption that prices are observed continuously. In contrast, the number of trades within any time interval is finite in real markets, and prices unfold in discrete time. We show that this assumption causes a downward bias when the number of trades per observation period is small. Moreover, these estimators consider only a subset of open, high, low, and close prices and thus neglect potentially useful information to improve the spread estimate and reduce the estimation variance. Jahan-Parvar and Zikes (2023) show that a larger estimation variance causes a larger upward bias when the spread is small compared to volatility due to the methods employed to guarantee non-negativity of the spread estimates in small samples. In summary, current estimators understate bid–ask spreads when they are expected to be the largest and overstate bid–ask spreads when they are expected to be the smallest.

In this paper, we develop an asymptotically unbiased estimator with minimum variance by accounting for discretely observed prices and optimally considering the complete information set of open, high, low, and close prices. First, we derive multiple bid–ask spread estimators from several combinations of prices. Our methodology yields estimators with an analytical term that depends on the probability that opening or closing prices coincide with the highest or lowest prices. Such probability would be zero if prices were observed continuously, and it can be regarded as an analytical correction term accounting for discretely observed prices. To give a sense of the importance of correcting by this term, Fig. 1 displays the probability that daily opening or closing prices coincide with the highest or lowest prices for U.S. common stocks from 1926 to 2021. The probability ranges between 25% for large stocks and 75% for small stocks in the last century and has decreased in the last two decades. Thus, while correcting by this term is less significant for more recent periods, it becomes essential when analyzing historical samples, and it is increasingly important for smaller stocks. Moreover, this term also depends on the sampling frequency used for estimation. Indeed, if intraday – instead of daily – prices are used, then the number of trades observed per time interval decreases, and the probability that opening or closing prices coincide with the highest or lowest prices increases. Thus, this term corrects a bias that varies in the time series and the cross-section and depends on the sampling frequency of open, high, low, and close prices.

Next, we combine our estimators to construct an efficient estimator. All estimators are asymptotically unbiased, so their efficient combination is obtained by minimizing the estimation variance. We proceed as follows. First, we identify two estimators that achieve minimum variance when the spread is small compared to volatility. Second, we identify two other estimators that exhibit the opposite behavior and achieve minimum variance when the spread is large compared to volatility. Third, we show that these estimators can be written as moment conditions and apply the generalized method of moments (Hansen, 1982) to construct our efficient estimator that achieves minimum variance across small and large spreads. By minimizing the estimation variance, our efficient estimation also minimizes the upward bias that arises in small samples due to the methods employed to guarantee non-negativity of the spread estimates (Jahan-Parvar and Zikes, 2023).

We compare our efficient estimator with the seminal Roll (1984) estimator and with those by Corwin and Schultz (2012) and Abdi and Ranaldo (2017) as they have been shown to deliver more accurate estimates than previous approaches.

In our simulation experiments, we study the bias and variance of the estimators. In agreement with our theoretical analysis, we find that other estimators understate the spread in simulations that use few trades per period, and the estimate shrinks to zero as the number of trades declines. Instead, our estimator remains unbiased even for simulations where we expect, on average, only a single trade per period. For simulations that use many trades per period, we find that all estimators are asymptotically unbiased, and they correctly estimate the spread used in the simulation. In this case, the best estimator has the lowest variance because it delivers unbiased estimates with higher precision. We find that the estimator by Corwin and Schultz (2012) has a lower variance than the estimator by Abdi and Ranaldo (2017) for small spreads, while it has a higher variance for large spreads. Our efficient estimator provides the most precise estimates with a variance lower than the other approaches across small and large spreads. In summary, our estimator dominates other approaches by yielding unbiased estimates when other estimators are biased and achieving minimum variance when all estimators are unbiased.

Our empirical analysis uses the Center for Research in Security Prices (CRSP) U.S. stock database to compute bid–ask spread estimates from daily prices. We compare them with the effective spread computed by matching high-frequency trades with quotes via the NYSE Trades and Quotes (TAQ) database in the sample period 1993–2021. The simulation-based results carry over to the empirical data. Our efficient estimator dominates all other estimators, and it is more correlated and considerably closer to the high-frequency benchmark in each subperiod, in each market venue, for small and large stocks, both in time series and cross-sectional studies, for each sample size and evaluation metric.

We illustrate the broad applicability of our estimator in low- and high-frequency both within and outside the U.S. stock market. First, we revisit historical spread estimates from daily prices in the U.S. stock market since 1926. For small stocks, our estimator closely overlaps with the high-frequency benchmark. In contrast, other estimators understate the spread, and their bias increases for older sample periods, mirroring that these estimators are more biased when trading becomes less frequent. Indeed, their bias reduces for larger stocks, which are presumably traded more frequently. For all stocks, we find that the end-of-day quoted spread is higher than our effective spread estimates by a factor of two. Thus, our estimator reproduces previous findings that the quoted spread overstates the effective spread finally paid by traders by up to 100% (Huang and Stoll, 1994; Petersen and Fialkowski, 1994; Bessembinder and Kaufman, 1997; Bacidorea et al., 2003), due to dealers offering a better price than the quotes, also known as trading inside the spread (Lee, 1993). In summary, our estimator makes available the most realistic effective spread estimates for the U.S. stock market from 1926 to the advent of high-frequency data.

Second, we show that our estimator can exploit intraday prices to improve the spread estimate significantly and that this approach is more effective than increasing the estimation sample with more daily data. Using minute prices – instead of daily – increases the correlation of the estimates with the benchmark from 56.17% to 88.79% in the challenging sample from October 2003 to December 2021, where the spread is small compared to volatility. The fraction of non-positive estimates reduces from 34.15% to 0.02%, and the upward bias induced by resetting negative estimates to zero essentially vanishes (Jahan-Parvar and Zikes, 2023). These results show that our estimator can be applied at any frequency, and, in this sense, it reconciles the high-frequency and low-frequency literature. Moreover, by relying on transaction prices only, our estimator is insensitive to the quality of quote data, which causes issues in measuring effective spreads by matching trades with quotes in fast and competitive markets (Holden and Jacobsen, 2014).

Third, we apply the estimator outside the stock market and analyze low- and high-frequency estimates for cryptocurrencies. We find that other estimators are dominated by their downward bias in high frequency and produce a tenfold difference between estimates that use daily or intraday prices. Instead, our estimator produces estimates from daily prices that closely overlap with those from hourly and minute prices. We conclude that our efficient estimator can potentially reduce a significant source of non-standard errors (Menkveld et al., 2024) in the measurement of transaction costs.

This paper is structured as follows. Section 2 reviews high- and low-frequency estimators of the effective bid–ask spread. Section 3 introduces our methodology and develops our estimators. Sections 4 and 5 present our simulation and empirical results, respectively. Section 6 illustrates the advantages and wide applicability of our efficient

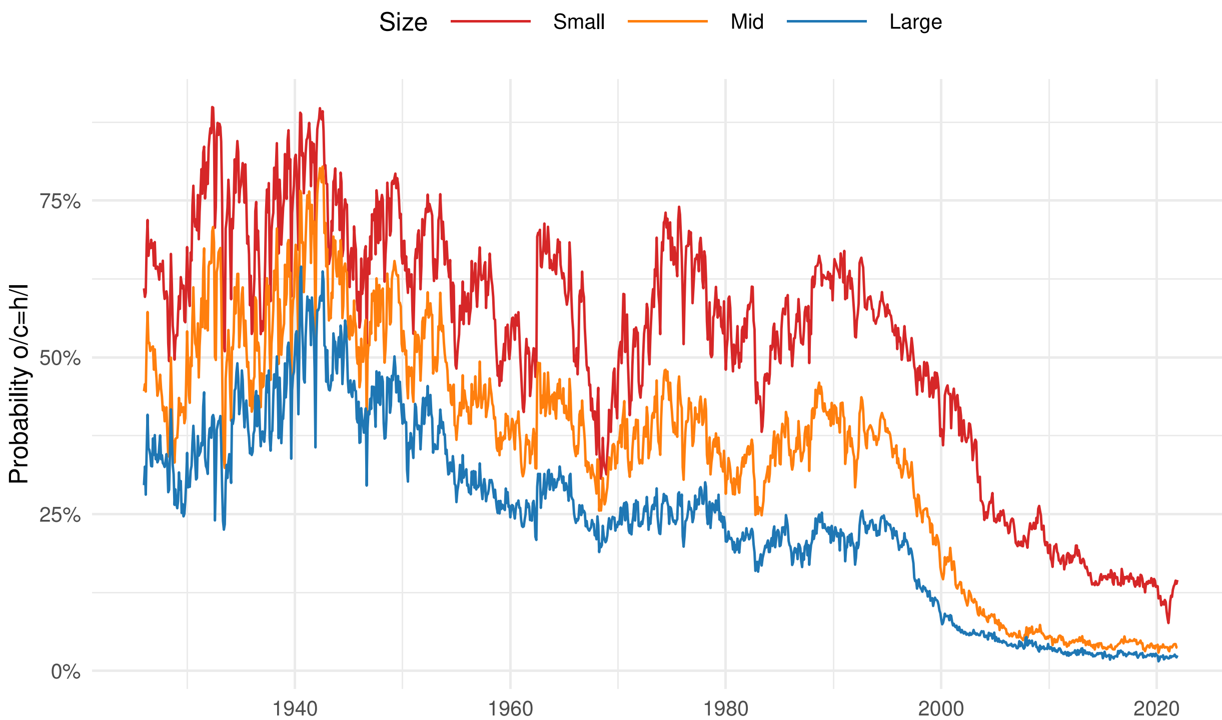

Fig. 1. Probability that Open/Close prices are High/Low prices.
The probability is computed for each stock-month as the average across: (i) the fraction of days such that the open matches the high, (ii) the fraction of days such that the open matches the low, (iii) the fraction of days such that the close matches the high, (iv) the fraction of days such that the close matches the low. Stocks are sorted into small-caps, mid-caps, and large-caps based on their market capitalization at the end of each month and using the 50th and 80th percentiles as breakpoints. The figure reports the average probability across stocks for each month and size group from 1926 to 2021. Open prices are missing from July 1962 through June 1992.

estimator. Finally, Section 7 concludes. To guarantee reproducibility, we make available software for the R statistical environment (R Core Team, 2020) that implements all the results in this paper. To facilitate adoption, we make our efficient estimator available in various programming languages. We also release open data containing all our spread estimates.[2]

2. The effective Bid–Ask spread

For a given trade, the relative effective bid–ask spread S is defined as:

$$S = \frac{2D(P - \tilde{P})}{\tilde{P}}, \tag{1}$$

where P is the observed transaction price, $\tilde{P}$ is the unobserved fundamental price, and D is a direction of trade indicator taking the value $+1$ for buyer-initiated trades, and -1 for seller-initiated trades. As the fundamental price $\tilde{P}$ is unobserved, different ways of estimating the spread exist, which depend on different proxies for $\tilde{P}$. Here, we review popular measures of the effective bid–ask spread that arise from different proxies. We classify these measures into two groups. First, we discuss measures that require trade and quote data and are typically used in the high-frequency literature. Then, we discuss measures that only require transaction prices and are typically used in the low-frequency literature.

2.1. High-frequency measures of effective spreads

One way to measure effective spreads is obtaining a proxy of the fundamental price from trade and quote data to plug in Eq. (1). This class of estimators measures the distance of transaction prices from the given proxy. Popular proxies are the quoted midpoint, the weighted midpoint, and the microprice.

[2] All code and data are available at https://github.com/eguidotti/bidask.

2.1.1. Quoted midpoint

A simple proxy of the fundamental price is the average of the bid and ask prices. The quoted midpoint P_M is defined as:

$$P_M = \frac{P_A + P_B}{2}, \tag{2}$$

where P_A and P_B are the ask and bid prices, respectively. Using $\tilde{P} = P_M$ in Eq. (1) we obtain the so-called midpoint effective spread (Hagströmer, 2021). This midpoint-based measure is required in U.S. regulations (SEC current Rule 605, Rule 11ac1-5 before 2007) and is often referred to as the effective spread. Here, we use the more precise terminology of Hagströmer (2021) to highlight that effective spreads are not observable and depend on the choice of the fundamental price. The midpoint effective spread is one possible measure of effective spreads.

2.1.2. Weighted midpoint

Hagströmer (2021) challenges using the quoted midpoint as a proxy of the fundamental price and shows that it leads to overstating effective spreads in markets with discrete prices and elastic liquidity demand. To overcome this problem, he proposes to use the weighted midpoint:

$$P_W = \frac{P_A Q_B + P_B Q_A}{Q_A + Q_B}, \tag{3}$$

where Q_A and Q_B are the depths quoted at the ask and bid prices, respectively.

2.1.3. Microprice

Stoikov (2018) criticizes the midpoint and weighted midpoint as proxies of the fundamental price for generating autocorrelated returns and proposes an alternative proxy – the microprice – that is a martingale by construction. We refer the reader to Stoikov (2018) for the construction of the microprice and to Hagströmer (2021) for a comparison of effective spreads obtained with the midpoint, weighted midpoint, and microprice.

2.2. Low-frequency measures of effective spreads

Another way to measure effective spreads is by introducing desirable assumptions about the data-generating process to develop an estimator that does not require an explicit proxy of the fundamental price. This class of estimators measures the distance of transaction prices from a fundamental price implicitly defined by the model's assumptions.

Several contributions (Roll, 1984; Hasbrouck, 2009; Corwin and Schultz, 2012; Abdi and Ranaldo, 2017) have proposed to derive an estimator of the effective spread by writing Eq. (1) in logarithmic prices $p = \log(P)$ such that:

$$p = \tilde{p} + Z\,, \tag{4}$$

where $Z = S/2D$ is the bid–ask bounce and the basic assumptions are that:

Assumption 1. Fundamental returns are not serially correlated.

Assumption 2. Fundamental returns are uncorrelated with bid–ask bounces.

Assumption 3. Bid–ask bounces are uncorrelated and have zero mean.

Assumptions 1–3 are the representative set of assumptions underlying previous contributions. However, they are more general than those required by each of them. For instance, the Roll (1984) model further assumes that buys and sells are equally likely. The Bayesian approach by Hasbrouck (2009) requires that fundamental returns are i.i.d. with normal distribution. Corwin and Schultz (2012) rely on the idea that high prices are buyer-initiated and low prices are seller-initiated and they model the fundamental price with a geometric Brownian motion with zero-mean returns, which is also used by Abdi and Ranaldo (2017). They further assume that spread and volatility are constant, ruling out stylized facts such as heteroscedasticity and jumps. To mitigate these restrictions, they advocate in favor of measuring the spread over two-day rolling periods and averaging these estimates. However, Jahan-Parvar and Zikes (2023) show this approach produces inconsistent estimators. Finally, one important limitation of all the previous contributions is that they do not account for the discrete nature of trades. Specifically, they require (explicitly or implicitly) the restrictive assumption that there is always at least one trade between two time instants such that prices are observed continuously.

Overall, the class of estimators based on Assumptions 1–3 aims at measuring the distance of transaction prices from a fundamental price with serially uncorrelated returns that are not correlated with bid–ask bounces. Such a class of estimators is the central focus of this paper, and we review the most popular approaches below. Other works alter the definition of the fundamental price by adding a dependence between the fundamental returns and the bid–ask bounces to accord an informational role to the trade directions, and they are outside the scope of this paper (see *e.g.*, Chen et al., 2017).

2.2.1. Close prices

The seminal work by Roll (1984) computes the serial covariance of observed returns to estimate the effective spread from closing prices. He shows that:

$$S^2 = -4\mathbb{C}\mathrm{ov}[\Delta c_t, \Delta c_{t-1}]\,, \tag{5}$$

where S^2 is the mean squared spread in the estimation sample and $\Delta c_t = c_t - c_{t-1}$ where c_t is the closing log-price of period t. The main limitation of this approach is that it has a large estimation variance, and the squared spread turns out to be negative in 50% of the cases using a yearly sample of daily closing prices (Roll, 1984). To improve the estimation accuracy, Hasbrouck (2009) proposes a Gibbs estimation of the Roll model. However, the method requires an iterative procedure, is computationally expensive, and needs many observations to converge.[3]

2.2.2. High and low prices

Corwin and Schultz (2012) propose an alternative estimator from high and low prices with smaller variance than the Roll (1984) estimator. Their methodology is based on the idea that high (low) prices are almost always buy (sell) trades. Hence, the high-low ratio incorporates both the volatility of the fundamental price and the bid–ask spread. As volatility increases with the return interval, while the spread does not, it is possible to derive a spread estimator from the high-low ratios over different time intervals. To link the high-low ratios with volatility, they assume that the fundamental price follows a geometric Brownian motion and use the equations by Parkinson (1980) and Garman and Klass (1980). However, these equations hold only if the price is observed continuously and are biased in practice as the number of trades within any time interval is finite.

2.2.3. Close, high, and low prices

Abdi and Ranaldo (2017) propose an estimator that jointly uses closing and high-low prices to achieve smaller variance than the Roll (1984) estimator and smaller bias than the Corwin and Schultz (2012) estimator. They show that:

$$S^2 = 4\mathbb{E}[(c_{t-1} - \eta_{t-1})(c_{t-1} - \eta_t)]\,, \tag{6}$$

where $\eta_t = (h_t + l_t)/2$ is the average of the high and low log-prices. However, their methodology also requires that the fundamental price follows a geometric Brownian motion with continuously observed prices. As a consequence, the estimator is still biased. Moreover, it does not exploit the full information set of open, high, low, and close prices to further improve the spread estimate.

3. Methodology

This paper relaxes the assumption that prices are observed continuously – and several other assumptions that were required by previous contributions – by deriving bid–ask spread estimators using Eq. (4) under only Assumptions 1–3. By accounting for the discrete nature of trades, we drastically reduce the estimation bias. By exploiting the full information set of open, high, low, and close prices, we minimize the estimation variance.

We start by introducing the indicator variable:

$$\tau_t = \begin{cases} 0 & \text{if} \quad h_t = l_t = c_{t-1} \\ 1 & \text{otherwise} \end{cases} \tag{7}$$

that equals 0 if the highest price matches the lowest price and the previous close, and it equals 1 otherwise. The value $\tau_t = 0$ indicates that either (i) all trades in period t are executed at the previous closing price, which is increasingly likely when the number of trades per period is smaller, or (ii) there is no trading and the open, high, low, and close prices of period t are filled with the previous close. The value $\tau_t = 1$ is the complementary case and ensures that prices are not forward-filled.

We now derive an estimator from close-to-open and open-to-mid (de-meaned) returns by considering their serial covariance:

$$\mathbb{C}\mathrm{ov}[\overline{\eta_t - o_t}, o_t - c_{t-1}] = \mathbb{E}[(\overline{\eta_t - o_t})(o_t - c_{t-1})]\,, \tag{8}$$

[3] From Hasbrouck's website (https://pages.stern.nyu.edu/~jhasbrou/Research/GibbsCurrent/gibbsCurrentIndex.html): "I often receive inquiries regarding Gibbs estimates formed at higher frequencies (e.g., monthly or weekly). I do not provide these estimates due to concerns about their reliability. The 2009 paper describes some of the issues that arise. Briefly, the prior distributions used here are diffuse (to ensure that the posteriors are data-dominated). The priors are generally, however, biased. As the sample size drops, the posteriors start resembling the priors, and the bias problem becomes more acute. The only way out of this is to put more structure on the priors. This is not impractical, but it is application-specific".

where $\eta_t = (h_t + l_t)/2$ is the average of the high and low log-prices, o_t is the opening log-price, c_{t-1} is the closing log-price of the previous time interval, and the de-meaned returns are defined as follows:

$$\overline{r_t} = r_t - \tau_t \frac{\mathbb{E}[r_t]}{\mathbb{E}[\tau_t]} . \tag{9}$$

In Appendix A.1, we prove that the covariance in Eq. (8) is equal to:

$$\mathbb{C}\mathrm{ov}[\eta_t - o_t, o_t - c_{t-1} \mid \tau_t = 1]\mathbb{P}[\tau_t = 1] . \tag{10}$$

Next, we replace observed prices with fundamental prices and bid–ask bounces as given in Eq. (4). As fundamental returns are not autocorrelated (Assumption 1), and they are also uncorrelated with bid–ask bounces (Assumption 2), Eq. (10) is equal to:

$$\mathbb{C}\mathrm{ov}[Z_{\eta_t} - Z_{o_t}, Z_{o_t} - Z_{c_{t-1}} \mid \tau_t = 1]\mathbb{P}[\tau_t = 1] , \tag{11}$$

where Z_{o_t} is the bid–ask bounce at the open, $Z_{c_{t-1}}$ is the bid–ask bounce at the previous close, and $Z_{\eta_t} = (Z_{h_t} + Z_{l_t})/2$. Conditional on $\tau_t = 1$, prices are not forward-filled, and thus bid–ask bounces at time t are uncorrelated with bid–ask bounces at time $t-1$ by assumption. Moreover, they have zero mean (Assumption 3). Thus, Eq. (11) is equal to:

$$\mathbb{E}[Z_{\eta_t} Z_{o_t} - Z_{o_t}^2 \mid \tau_t = 1]\mathbb{P}[\tau_t = 1] . \tag{12}$$

We now need to compute the expectation in Eq. (12). To this end, we recall that $Z_{o_t} = S_{o_t}/2D_{o_t}$ and thus $Z_{o_t}^2 = S_{o_t}^2/4$. Hence, we have:

$$\mathbb{E}[Z_{o_t}^2 \mid \tau_t = 1] = \mathbb{E}[S_{o_t}^2]/4 , \tag{13}$$

and the remaining term is calculated in Appendix A.2:

$$\mathbb{E}[Z_{\eta_t} Z_{o_t} \mid \tau_t = 1] = \frac{\mathbb{E}[S_{o_t}^2]}{4} \frac{\mathbb{P}[o_t = h_t \mid \tau_t = 1] + \mathbb{P}[o_t = l_t \mid \tau_t = 1]}{2} . \tag{14}$$

Finally, we substitute Eqs. (13)–(14) into Eq. (12) and solve for the spread. Following the calculations in Appendix A.3, we obtain that the mean squared spread is:

$$S_o^2 = \mathbb{E}[S_{o_t}^2] = \frac{-8\mathbb{E}[(\overline{\eta_t - o_t})(o_t - c_{t-1})]}{\mathbb{P}[o_t \neq h_t, \tau_t = 1] + \mathbb{P}[o_t \neq l_t, \tau_t = 1]} . \tag{15}$$

3.1. Efficient estimation of effective spreads

So far, we have derived an estimator from close-to-open and open-to-mid returns. However, the same methodology can be used to derive estimators from other combinations of prices. This section identifies four estimators that achieve minimum variance under different conditions. Then, we optimally combine the four estimators to minimize the estimation variance under any condition and obtain an efficient estimator.

For illustration, we consider the case where high and low prices always differ from open or close prices, and returns have zero mean such that $\overline{r_t} = r_t$. From Eq. (15) we obtain that the spread is proportional to $S_o^2 = \mathbb{E}[(\eta_t - o_t)(o_t - c_{t-1})]$ and thus the estimation variance is proportional to:

$$\mathbb{V}\mathrm{ar}[\hat{S}_o^2] = \mathbb{V}\mathrm{ar}[(\eta_t - o_t)(o_t - c_{t-1})] . \tag{16}$$

Eq. (16) shows that the estimation variance depends on the volatility of observed returns. Thus, it depends on the volatility of the fundamental price and the size of the bid–ask spread. We now consider two complementary cases where the spread is either small or large compared to the volatility of the fundamental price.

In the first case, $S \to 0$ and observed prices p coincide with fundamental prices $\tilde{p}$. In this case, the estimation variance is proportional to:

$$\mathbb{V}\mathrm{ar}[\hat{S}_o^2] = \mathbb{V}\mathrm{ar}[(\tilde{\eta}_t - \tilde{o}_t)(\tilde{o}_t - \tilde{c}_{t-1})] = \mathbb{V}\mathrm{ar}[\tilde{\eta}_t - \tilde{o}_t]\mathbb{V}\mathrm{ar}[\tilde{o}_t - \tilde{c}_{t-1}] . \tag{17}$$

Eq. (17) shows that the estimation variance decreases with the sampling frequency because the volatility of the fundamental price reduces at higher frequencies and makes the estimation variance smaller. In other words, we obtain that the bid–ask spread should be estimated with the highest frequency data possible and that estimators considering higher time lags are dominated by estimators considering the smallest possible lag. The estimator in Eq. (15) is optimal because it considers subsequent close-to-open and open-to-mid returns. An equivalent estimator is obtained by considering subsequent mid-to-close and close-to-open returns, as we expect open-to-mid returns to be distributed similarly to mid-to-close returns. All other estimators have a larger variance and are dominated by these two because they require higher time lags.

In the second case, $S \to \infty$ and observed returns are driven by bid–ask bounces. Moreover, as the spread is large, high prices are buys, and low prices are sells. In this case, $Z_{\eta_t} = (Z_{h_t} + Z_{l_t})/2 = S/4 - S/4 = 0$ and the estimation variance is proportional to:

$$\mathbb{V}\mathrm{ar}[\hat{S}_o^2] = \mathbb{V}\mathrm{ar}[(Z_{\eta_t} - Z_{o_t})(Z_{o_t} - Z_{c_{t-1}})] = \mathbb{V}\mathrm{ar}[Z_{o_t}]^2 + \mathbb{V}\mathrm{ar}[Z_{o_t}]\mathbb{V}\mathrm{ar}[Z_{c_{t-1}}] . \tag{18}$$

Eq. (18) shows that the estimation variance can be reduced by using the mid price η_{t-1} ($\mathbb{V}\mathrm{ar}[Z_{\eta_{t-1}}] = 0$) instead of the closing price c_{t-1} ($\mathbb{V}\mathrm{ar}[Z_{c_{t-1}}] \to \infty$). In this case, the estimation variance becomes $\mathbb{V}\mathrm{ar}[Z_{o_t}]^2$, which is strictly smaller than that in Eq. (18). In other words, it is convenient to consider subsequent mid-to-open and open-to-mid returns when the spread is large compared to volatility. An equivalent estimator is obtained by considering subsequent mid-to-close and close-to-mid returns, as we expect (i) mid-to-close returns to be distributed similarly to open-to-mid returns and (ii) close-to-mid returns to be distributed similarly to mid-to-open returns.

Table 1
Discrete generalized estimators.

Estimators			
OHL	$S_o^2 = \pi_o\ \mathbb{E}[(\overline{\eta_t - o_t})(o_t - \eta_{t-1})]$	$S_c^2 = \pi_c\ \mathbb{E}[(\overline{\eta_t - c_{t-1}})(c_{t-1} - \eta_{t-1})]$	CHL
OHLC	$S_o^2 = \pi_o\ \mathbb{E}[(\overline{\eta_t - o_t})(o_t - c_{t-1})]$	$S_c^2 = \pi_c\ \mathbb{E}[(\overline{o_t - c_{t-1}})(c_{t-1} - \eta_{t-1})]$	CHLO
Coefficients			
π_o	$\pi_o = -8 \,/\, \left(\mathbb{P}[o_t \neq h_t, \tau_t = 1] + \mathbb{P}[o_t \neq l_t, \tau_t = 1]\right)$		
π_c	$\pi_c = -8 \,/\, \left(\mathbb{P}[c_{t-1} \neq h_{t-1}, \tau_t = 1] + \mathbb{P}[c_{t-1} \neq l_{t-1}, \tau_t = 1]\right)$		
Prices			
o,h,l,c	Open, High, Low, Close log-prices.		
η	Mid-prices computed as $\eta = (l + h)/2$.		

This table reports bid–ask spread estimators obtained from several combinations of open, high, low, and close prices as described in Section 3. The OHL and OHLC estimators measure the spread at the open. The CHL and CHLO estimators measure the spread at the close. The indicator variable τ_t is defined in Eq. (7) and the de-meaned returns $\overline{r_t}$ are defined in Eq. (9).

Table 1 summarizes the four estimators derived from the combinations of prices discussed above. We call these estimators Discrete Generalized Estimators (DGEs) because they account for the fact that prices unfold in discrete time and generalize previous approaches that rely on continuously observed prices. For instance, the estimator by Abdi and Ranaldo (2017) can be regarded as a particular case of our CHL estimator in Table 1. Indeed, if we require prices to be observed continuously, they are never forward-filled and the closing price always differs from the high or low prices. Therefore, $\pi_c = -4$ and for zero-mean returns the CHL estimator becomes $S^2 = -4\mathbb{E}[(\overline{\eta_t - c_{t-1}})(c_{t-1} - \eta_{t-1})] = 4\mathbb{E}[(c_{t-1} - \eta_t)(c_{t-1} - \eta_{t-1})]$, which is identical to the estimator in Eq. (6). Thus, our CHL estimator can be regarded as a generalization of the Abdi and Ranaldo (2017) estimator that provides an analytical correction term accounting for discretely observed prices.

Next, we combine our DGEs to minimize the estimation variance and obtain the Efficient DGE (EDGE). To this end, we notice that each DGE can be written as a moment condition so that their efficient combination is obtained by applying the Generalized Methods of Moments (GMM) (Hansen, 1982). As discussed above, the OHL estimator in

Table 1 is expected to perform similarly to CHL, and OHLC is expected to perform similarly to CHLO. However, OHL and OHLC measure the spread at the open while CHL and CHLO measure the spread at the close. We thus combine OHL with CHL and OHLC with CHLO to obtain two moment conditions that measure the average spread at the open and close:

$$\mathbb{E}\left[2S^2 - \pi_o(\overline{\eta_t - o_t})(o_t - \eta_{t-1}) - \pi_c(\overline{\eta_t - c_{t-1}})(c_{t-1} - \eta_{t-1})\right] = 0\,, \tag{19}$$

$$\mathbb{E}\left[2S^2 - \pi_o(\overline{\eta_t - o_t})(o_t - c_{t-1}) - \pi_c(\overline{o_t - c_{t-1}})(c_{t-1} - \eta_{t-1})\right] = 0\,, \tag{20}$$

where we have set $S^2 = (S_o^2 + S_c^2)/2$ for notational convenience. These moment conditions can be written as $\mathbb{E}[S^2 - x_{i,t}] = 0$ where x is opportunely defined. By applying GMM, the efficient estimator is given by:

$$S^2_{\text{GMM}} = \arg\min_{S^2} \sum_{ij} (S^2 - \mu_i)\, W_{ij}\, (S^2 - \mu_j)\,, \tag{21}$$

where $\mu_i = \mathbb{E}[x_{i,t}]$ and $W = \Omega^{-1}$ is the inverse of the covariance matrix $\Omega = \mathbb{V}\text{ar}[S^2 - x_{i,t}]$, which simplifies to $\Omega = \mathbb{V}\text{ar}[x_{i,t}]$ as the variance is translation invariant. Therefore, we have a particular case of GMM where the optimal weighting matrix does not depend on the minimizing variable, and the problem reduces to the minimization of a quadratic form. By differentiating Eq. (21), setting the derivative equal to zero, and solving for S^2, we obtain:

$$S^2_{\text{GMM}} = \sum_i w_i \mu_i \quad \text{with} \quad w_i = \frac{\sum_j W_{ij}}{\sum_{i,j} W_{ij}}\,. \tag{22}$$

Finally, applying GMM in Eq. (22) with the two moment conditions above and a diagonal covariance matrix Ω gives our *Efficient Discrete Generalized Estimator* (EDGE):

$$S^2_{\text{EDGE}} = w_1 \mathbb{E}[x_{1,t}] + w_2 \mathbb{E}[x_{2,t}]\,, \tag{23}$$

$$\begin{aligned} x_{1,t} &= \frac{\pi_o}{2}(\overline{\eta_t - o_t})(o_t - \eta_{t-1}) + \frac{\pi_c}{2}(\overline{\eta_t - c_{t-1}})(c_{t-1} - \eta_{t-1})\,, \\ x_{2,t} &= \frac{\pi_o}{2}(\overline{\eta_t - o_t})(o_t - c_{t-1}) + \frac{\pi_c}{2}(\overline{o_t - c_{t-1}})(c_{t-1} - \eta_{t-1})\,, \end{aligned} \tag{24}$$

$$w_1 = \frac{\mathbb{V}\text{ar}[x_{2,t}]}{\mathbb{V}\text{ar}[x_{1,t}] + \mathbb{V}\text{ar}[x_{2,t}]}\,, \quad w_2 = \frac{\mathbb{V}\text{ar}[x_{1,t}]}{\mathbb{V}\text{ar}[x_{1,t}] + \mathbb{V}\text{ar}[x_{2,t}]}\,. \tag{25}$$

For estimation, the usual sample counterparts replace the expectations and variances, respectively.

3.2. Negative estimates

Our estimators and those of Roll (1984), Corwin and Schultz (2012), and Abdi and Ranaldo (2017) are formal estimators for the mean squared spread S^2. However, the estimate $\hat{S}^2$ may become negative in small samples due to statistical fluctuations. This is an issue because a negative squared spread is not mathematically nor economically meaningful.

To guarantee the non-negativity of spread estimates, it is common to reset negative values to zero by applying the transformation:

$$\hat{S} = \sqrt{\max\left\{0\,, \hat{S}^2\right\}}\,. \tag{26}$$

Although this approach maintains non-negativity, it can lead to a substantial number of zero estimates, which can be problematic for certain applications like portfolio sorting. In an effort to mitigate this drawback, earlier studies have explored calculating the squared spread across rolling time intervals, resetting negative estimates to zero within these intervals, and subsequently computing the average across the entire estimation period. However, Jahan-Parvar and Zikes (2023) have shown that this strategy introduces a strong upward bias that does not decline as the sample size increases, making the estimates inconsistent. They document that volatility is the primary driver of the bias and that inconsistent measures fail to replicate some well-known results in empirical finance. Following their recommendations, we apply the transformation in Eq. (26) to the final estimate. This ensures that the bias declines as the sample size increases and the estimate $\hat{S}$ is consistent.

Another way to produce consistent estimates while avoiding zero values is to take the square root of the modulus of the final estimate $\hat{S}^2$. This proposition is motivated by the positive correlation between negative estimates and minus the spread that we have found empirically (see Internet Appendix I.2). Another possibility is to use the tick size as a reasonable lower bound for the estimates.[4] However, for the sake of comparability with prior studies, we reset negative estimates to zero within this paper, leaving the exploration of alternative approaches for future research.

4. Simulation results

In this section, we perform Monte Carlo simulations to study the accuracy of EDGE and its building blocks. We compare the results with the seminal Roll (1984) estimator and with the estimators proposed more recently by Corwin and Schultz (2012) and Abdi and Ranaldo (2017). Throughout the paper, we refer to these estimators with ROLL, CS, and AR, respectively. The CS estimator is adjusted for overnight returns as described in Corwin and Schultz (2012).

4.1. Setup

For ease of comparison, we use the simulation setup of Corwin and Schultz (2012), also used in Abdi and Ranaldo (2017). Specifically, we simulate 10,000 months of data where each month consists of 21 trading days and each day consists of 390 min. For each minute of the day, the fundamental price $\tilde{P}_m$ is simulated as $\tilde{P}_m = \tilde{P}_{m-1} e^{\sigma z}$ with $\tilde{P}_0 = 1$, where σ is the standard deviation per minute and z is a random draw from a standard Gaussian distribution. The daily standard deviation equals 3%, and the standard deviation per minute equals 3% divided by $\sqrt{390}$. Trade prices are defined as $\tilde{P}_m$ multiplied by one minus (plus) half the assumed bid–ask spread, and we use a 50% chance for bid (ask) prices. Prices are assumed to be observed with a given probability. Daily high and low prices equal the highest and lowest prices observed during the day. Open and close prices equal the first and the last prices observed in the day. If no trade is observed for a given day, then the previous day's closing price is used as the open, high, low, and close prices for that day.

4.2. Results

We start by studying the bias of the various estimators. To this end, we simulate 10,000 months of daily prices and estimate the spread using the whole time series. These simulations use a constant spread of 1%, and the probability of observing a trade ranges from 1/390 to 1, such that the expected number of daily trades ranges from 1 to 390.

Fig. 2 shows how the spread estimate varies in function of the trading frequency. We find that all estimators are unbiased and correctly estimate a spread equal to 1% when we use 390 trades per day. However, their behavior is substantially different when the trading frequency declines. Indeed, CS estimates a spread of 0.75% in the simulation using 100 trades per day. Moreover, its downward bias increases rapidly as the trading frequency declines further, and it returns an estimate of zero in the simulations that use less than ten trades per day. AR is less sensitive to the trading frequency, but it is still significantly biased in the simulations that use only a few trades per day. Instead, EDGE produces unbiased estimates regardless of the number of trades, suggesting it works well in practice even for assets that trade infrequently. These results demonstrate how CS and AR strongly rely on the assumption that assets are traded continuously and produce

[4] We thank an anonymous referee for this suggestion.

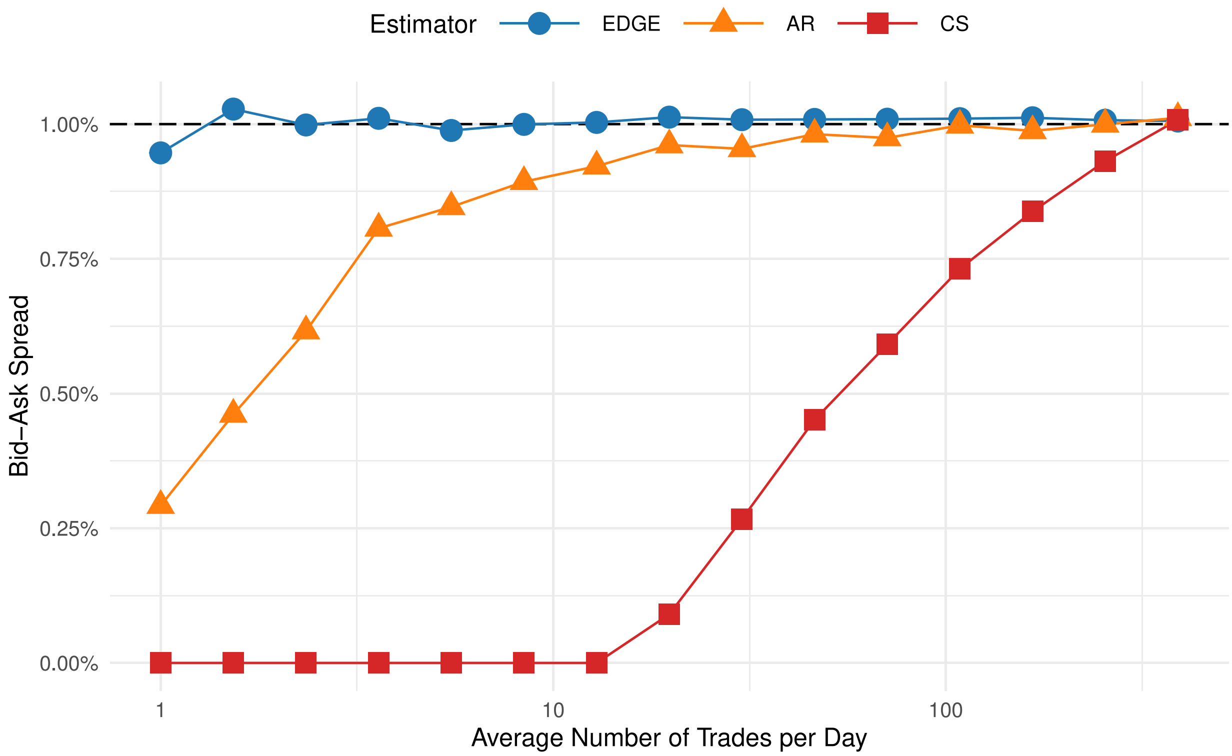

Fig. 2. Bias of simulated spread estimates.
The figure reports spreads estimated from 10,000 months of simulated data where each month consists of 21 trading days and each day consists of 390 min. For each minute of the day, the fundamental value $\tilde{P}_m$ is simulated as $\tilde{P}_m = \tilde{P}_{m-1}e^{\sigma z}$ with $\tilde{P}_0 = 1$, where σ is the standard deviation per minute and z is a random draw from a standard Gaussian distribution. The daily standard deviation equals 3%, and the standard deviation per minute equals 3% divided by $\sqrt{390}$. Trade prices are defined as $\tilde{P}_m$ multiplied by one minus (plus) half the bid–ask spread, where the spread equals 1.00%, and we assume a 50% chance that a bid (ask) is observed. The probability of observing a trade ranges from 1/390 to 1, and the average number of trades per day is reported on the x-axis. Daily high and low prices equal the highest and lowest prices observed during the day. Open and close prices equal the first and the last prices observed in the day. If no trade is observed for a given day, then the previous day's closing price is used as the open, high, low, and close prices for that day. Negative spread estimates are set to zero.

downward biased estimates when that assumption is not satisfied. Our more general methodology provides an analytical correction term that accounts for infrequent trading and produces asymptotically unbiased estimates.

Next, we study the variance of the estimators by computing the standard deviation of monthly spread estimates across 10,000 simulations, where each month consists of 21 trading days. These simulations use 390 trades per day to ensure that all estimators are unbiased. In this setting, an estimator with lower variance is strictly preferable to one with higher variance because it produces unbiased estimates with higher precision.

Fig. 3 reports the standard deviation of spread estimates in simulations that use a constant spread ranging from 0.50% to 8.00%. CS is preferable to AR for smaller spreads, while AR is for larger spreads. EDGE is always the best estimator, providing the most precise estimates with minimum variance uniformly across small and large spreads.

To shed light on the performance of EDGE, we also report the behavior of its building blocks. In agreement with the discussion in Section 3.1, Fig. 3 shows that OHL is equivalent to CHL, and OHLC is equivalent to CHLO. Moreover, the variance of OHLC and CHLO decreases for smaller spreads. On the contrary, the variance of OHL and CHL decreases for larger spreads. EDGE exploits the opposite behaviors of these estimators to produce estimates with minimum variance uniformly across small and large spreads. Indeed, Eq. (25) shows that EDGE puts more weight on the OHLC and CHLO estimators for smaller spreads, while it puts more weight on the OHL and CHL estimators for larger spreads. The result is an estimator that achieves minimum variance across small and large spreads. For an additional comparison, we also report the results for the GMM estimator in Eq. (22) where we set the weighting matrix equal to the identity matrix. This estimator has roughly the same variance of EDGE for spreads between 2.00% and 5.00%, but its variance is worse for smaller and larger spreads. We conclude that the weighting matrix used for EDGE is effective in minimizing the estimation variance.

Finally, Table 2 reports the mean and standard deviation of monthly spread estimates from daily prices across 10,000 simulations. Panel A uses 390 trades per day to simulate frequent trading. Here, estimators other than ROLL produce mean spreads close to the actual values used in the simulation and are essentially unbiased. ROLL is affected by an upward bias for small spreads that arises from truncating negative estimates and is exacerbated by the large estimation variance. EDGE outperforms all other estimators in these simulations by producing unbiased estimates with the lowest variance across small and large spreads. Panel B introduces infrequent trading in the simulations. We find that EDGE outperforms its building blocks by producing estimates with lower variance and other estimators by producing estimates with lower bias. AR seems to perform similarly to EDGE for simulations that use a spread of 0.50%, but this is due to the downward bias for infrequent trading being counterbalanced by the upward bias induced by truncating negative estimates. Although CS produces estimates with low variance, these estimates are strongly biased. For instance, CS estimates a spread of 0.04% where the actual spread used in the simulation is 1.00%.

In summary, our estimator yields unbiased estimates when other estimators are biased and achieves minimum variance when all estimators are unbiased.

5. Empirical results

In this section, we investigate the performance of the estimators on empirical data. To evaluate the performance, we first need to define

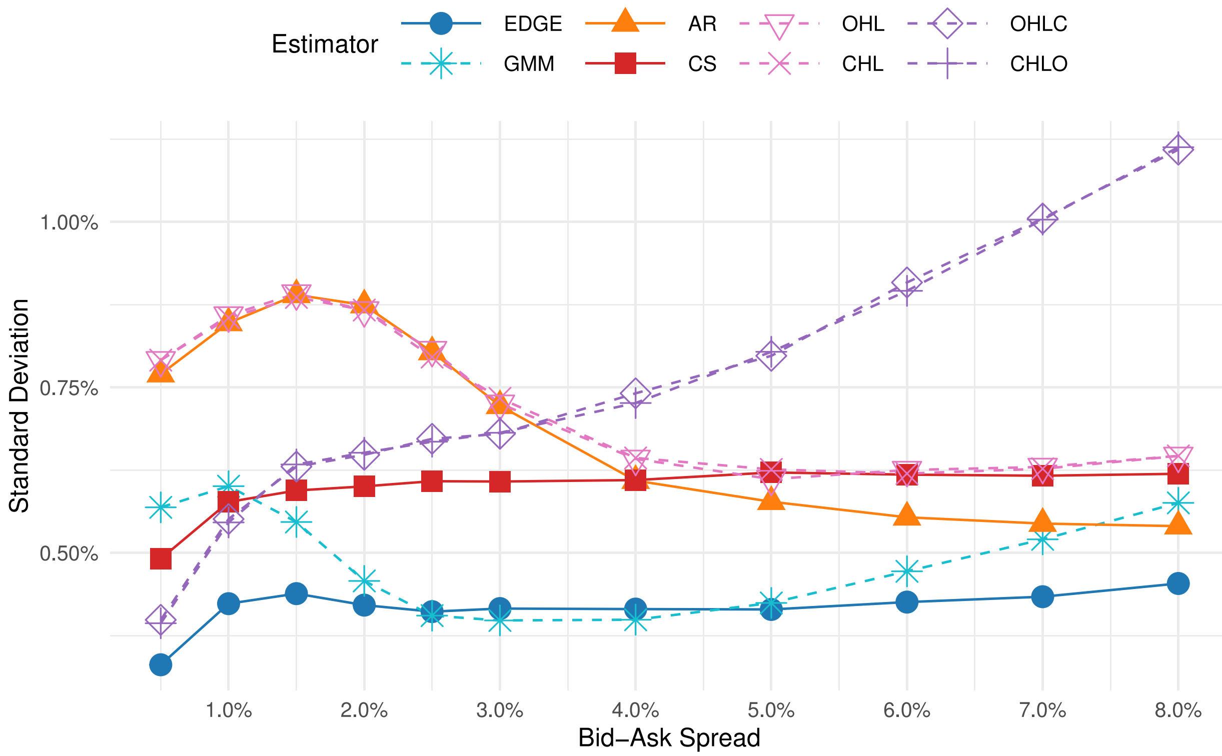

Fig. 3. Variance of simulated spread estimates.
The figure reports the standard deviation of monthly spread estimates across 10,000 simulations, where each month consists of 21 trading days and each day consists of 390 min. For each minute of the day, the fundamental value $\tilde{P}_m$ is simulated as $\tilde{P}_m = \tilde{P}_{m-1}e^{\sigma z}$ with $\tilde{P}_0 = 1$, where σ is the standard deviation per minute and z is a random draw from a standard Gaussian distribution. The daily standard deviation equals 3%, and the standard deviation per minute equals 3% divided by $\sqrt{390}$. Trade prices are defined as $\tilde{P}_m$ multiplied by one minus (plus) half the bid–ask spread, where the spread is reported on the x-axis, and we assume a 50% chance that a bid (ask) is observed. Daily high and low prices equal the highest and lowest prices observed during the day. Open and close prices equal the first and the last prices observed in the day. Negative spread estimates are set to zero. The OHL, CHL, OHLC, and CHLO estimators are defined in Table 1, and GMM is their GMM-combination in Eq. (22) where the weighting matrix is the identity matrix.

Table 2
Monthly estimates from simulated daily prices.

		EDGE	OHLC	CHLO	OHL	CHL	AR	CS	ROLL
Panel A: Frequent Trading									
$S = 0.50\%$	Mean	0.44	0.46	0.46	0.79	0.79	0.70	0.60	1.44
	(sd)	(0.33)	(0.40)	(0.39)	(0.79)	(0.79)	(0.77)	(0.49)	(1.43)
$S = 1.00\%$	Mean	0.90	0.88	0.88	1.03	1.03	0.95	1.03	1.59
	(sd)	(0.42)	(0.55)	(0.55)	(0.86)	(0.86)	(0.85)	(0.58)	(1.49)
$S = 3.00\%$	Mean	2.88	2.87	2.88	2.92	2.93	2.92	2.93	2.95
	(sd)	(0.41)	(0.69)	(0.69)	(0.73)	(0.72)	(0.70)	(0.61)	(1.83)
$S = 5.00\%$	Mean	4.87	4.86	4.87	4.92	4.93	4.97	4.90	4.90
	(sd)	(0.42)	(0.81)	(0.81)	(0.62)	(0.62)	(0.58)	(0.61)	(2.14)
$S = 8.00\%$	Mean	7.84	7.78	7.79	7.88	7.89	7.99	7.86	7.93
	(sd)	(0.45)	(1.11)	(1.10)	(0.64)	(0.64)	(0.54)	(0.62)	(2.63)
Panel B: Infrequent Trading									
$S = 0.50\%$	Mean	0.71	0.77	0.79	0.89	0.91	0.65	0.02	1.44
	(sd)	(0.75)	(0.87)	(0.88)	(0.96)	(0.97)	(0.73)	(0.07)	(1.42)
$S = 1.00\%$	Mean	0.95	0.99	0.99	1.11	1.10	0.81	0.04	1.56
	(sd)	(0.83)	(0.97)	(0.96)	(1.03)	(1.04)	(0.80)	(0.10)	(1.47)
$S = 3.00\%$	Mean	2.89	2.76	2.76	2.86	2.86	2.26	0.35	2.89
	(sd)	(0.83)	(1.23)	(1.23)	(1.20)	(1.19)	(0.92)	(0.36)	(1.82)
$S = 5.00\%$	Mean	5.02	4.89	4.92	5.01	5.04	4.04	1.17	4.83
	(sd)	(0.81)	(1.32)	(1.33)	(1.13)	(1.13)	(0.85)	(0.62)	(2.12)
$S = 8.00\%$	Mean	8.19	8.10	8.06	8.23	8.20	6.59	2.66	7.71
	(sd)	(0.96)	(1.59)	(1.62)	(1.24)	(1.26)	(0.94)	(0.96)	(2.65)

The table reports means and standard deviations (in %) of monthly spread estimates across 10,000 simulations, where each month consists of 21 trading days and each day consists of 390 min. For each minute of the day, the fundamental value $\tilde{P}_m$ is simulated as $\tilde{P}_m = \tilde{P}_{m-1}e^{\sigma z}$ with $\tilde{P}_0 = 1$, where σ is the standard deviation per minute and z is a random draw from a standard Gaussian distribution. The daily standard deviation equals 3%, and the standard deviation per minute equals 3% divided by $\sqrt{390}$. Trade prices are defined as $\tilde{P}_m$ multiplied by one minus (plus) half the bid–ask spread S, and we assume a 50% chance that a bid (ask) is observed. Panel A reports the results where the probability of observing a trade is 100%. In Panel B, that probability equals 1%. Daily high and low prices equal the highest and lowest prices observed during the day. Open and close prices equal the first and the last prices observed in the day. If no trade is observed for a given day, then the previous day's closing price is used as the open, high, low, and close prices for that day. Negative spread estimates are set to zero.

the ground truth, that is, the spread that serves as the benchmark for the evaluation. Following the literature, we use the effective spread obtained by matching high-frequency trade and quote data to evaluate the performance of the various estimators that only require commonly available daily price data.

5.1. Data

To compute bid–ask spread estimates (*i.e.*, EDGE, AR, CS, ROLL), we obtain daily prices from the CRSP US Stock Database in the period 1926–2021 for all NYSE, AMEX, and NASDAQ stocks with CRSP share codes of 10 or 11 (*i.e.*, U.S. common shares). To ensure that all the estimates are obtained from transaction prices, we keep only observations for which the open, high, low, and close prices are directly available. CRSP reports quotes derived from bid and ask prices if transaction prices are unavailable, and a dash in front of the price marks these values. We consider these non-transaction-based prices as missing values. Then, we drop the days where the high, low, or close price is missing. We also drop days where the open or close prices are outside the high-low range or where the low price is higher than the high price.

We match CRSP and TAQ daily data using CUSIP identifiers and tickers. First, we reconstruct the time series of `CUSIP`s for each `KYPERMNO` in CRSP. Similarly, we reconstruct the time series of `TICKER`s for each `KYPERMNO` in CRSP. Then, we compute the time series of `CUSIP`s for each `SYMBOL` in TAQ using the Monthly TAQ Master files for 1993–2009 and the Daily TAQ Master files for 2010–2021. Finally, we merge the daily datasets by matching observations with the same date, with the same CUSIP, and where the TAQ's `SYMBOL` is equal to the `TICKER` in CRSP. Our identification strategy allows us to match 99% of the stocks in CRSP.

For each stock-month, we estimate the spread from daily prices with EDGE, AR, CS, and ROLL and drop the estimate for all the estimators if it is missing for any of them. For instance, EDGE cannot be computed if open prices are missing, and ROLL cannot be computed if a stock-month contains only two daily observations. In such cases, we drop the corresponding estimate for all estimators. We use no explicit cutoff for the number of observations in a given stock-month. The cutoff is implicitly determined by the requirements of the most stringent estimator. Ultimately, the covariance requires at least two returns to be computed, meaning we need at least three daily observations in a stock-month. In our CRSP-TAQ merged sample, the frequency of missing estimates for each of the estimators is 1.24% for EDGE, 0.48% for CHL, 1.02% for OHL, 1.17% for CHLO, 1.02% for OHLC, 0.03% for AR, 0.03% for CS, and 0.14% for ROLL. Moreover, when CHL is missing and CS is not, the CS estimate is zero in 90% of the cases. When CHL is missing and AR is not, the AR estimate is zero in 100% of the cases. These are mostly cases when the stock always trades at the same price so that the denominator of our estimators is zero and the estimate is undefined. In such cases, a missing estimate should be preferable to an implicit imputation of zero produced by the other estimators.

We rely on the TAQ database from May 1993 to December 2021 to compute the benchmark effective spread. Daily spreads are obtained via the Wharton Research Data Services (WRDS) Intraday Indicators using Monthly TAQ from 1993 to 2003 and Daily TAQ from 2004 onward, according to the methodology described in Holden and Jacobsen (2014). For each month, we winsorize the daily spreads at 99.5% (one-sided) and compute the root mean squared spread for each stock. We refer to this measure as HJ.

To ensure that our results are robust to the choice of the benchmark, we also compute spreads using the weighted midpoint as described in Hagströmer (2021). First, we replicate the daily spread measures from the WRDS Intraday Indicators using the Daily TAQ database in the period 2004–2021 and we recompute our monthly HJ benchmark. The benchmark achieves 99.5% correlation with the one obtained using the estimates pre-computed by WRDS. Next, we replace the midpoint with the weighted midpoint to generate the effective spreads described in Hagströmer (2021). The correlation between the monthly benchmarks using the midpoint and weight-midpoint effective spreads is 99.1%. We have evaluated the estimators using both benchmarks, and all the results are fully consistent. Throughout the paper, we use the midpoint benchmark as it is pre-computed by WRDS also for the Monthly TAQ database in the period 1993–2003, where the National Best Bid and Offer (NBBO) is not directly available and matching trades with quotes poses several challenges (Holden and Jacobsen, 2014).

5.2. Results

Our CRSP-TAQ merged sample consists of about 1.6 million stock-month spread estimates for each estimator in the sample period from May 1993 to December 2021. In Table 3, we report summary statistics and several evaluation metrics for the estimates. EDGE achieves the highest correlation with the HJ benchmark, the lowest mean absolute percentage error (MAPE), root mean squared error (RMSE), and the smallest fraction of zero estimates.[5]

The remainder of this section is dedicated to a deeper comparison across the estimators in a cross-sectional, time-series, and panel-data setting.

5.3. Cross-sectional correlation

Looking at cross-sectional correlations on a month-by-month basis allows us to evaluate the estimators' ability to capture the cross-sectional distribution of spreads in different time periods. Given the effective spread benchmark $S_{i,t}$ for stock i at time t and the corresponding estimate $\hat{S}_{i,t}$, we compute the cross-sectional correlation at time t as $\rho_t = \mathbb{C}\text{or}_i[S_{i,t}, \hat{S}_{i,t}]$. The month-by-month cross-sectional correlations for the various estimators are displayed in Fig. 4. The correlation between EDGE and the effective spread benchmark is consistently higher than the correlations achieved by any other estimator throughout the whole period considered in the analysis.

5.4. Time-series correlation

Looking at time-series correlations on a stock-by-stock basis allows us to evaluate the ability of the estimators to capture the time-series distribution of spreads for different kinds of stocks. To this end, we split all stocks in deciles based on their market capitalization. The size deciles are sorted by increasing market capitalization of each stock as its last listing date in CRSP. Then, given the effective spread benchmark $S_{i,t}$ for stock i at time t and the corresponding estimate $\hat{S}_{i,t}$, we compute the time-series correlation for decile d as $\rho_d = \mathbb{C}\text{or}_{t,i\in d}[S_{i,t}, \hat{S}_{i,t}]$. The time-series correlations for each decile obtained with the various estimators are displayed in Fig. 5. The correlation between EDGE and the effective spread benchmark is consistently higher than the correlations achieved by any other estimator for all types of stocks.

5.5. Panel-data correlation

Next, we analyze the performances across five dimensions: market venues, time periods, market capitalization, spread size, and trading frequency. When analyzing market venues, the groups correspond to NYSE, AMEX, and NASDAQ. For the time periods, we use those defined in Corwin and Schultz (2012) and Abdi and Ranaldo (2017). In addition, we extend the sample and include the more recent subperiod 2016–2021. For size groups, we sort stocks in quintiles based on their market capitalization at their last listing date in CRSP. Spread quintiles are sorted on the average effective spread throughout the life

[5] The MAPE and RMSE are computed on log-spreads as described in Internet Appendix I.4.

Table 3
Summary statistics.

	N (1)	Mean (%)	Med (%)	Sd (%)	Cor_P (%)	Cor_S (%)	MAPE (%)	RMSE (1)	FNPE (%)
EDGE	1,637,621	2.11	1.00	3.37	**78.86**	**66.68**	**16.21**	**1.23**	**25.63**
OHLC	1,637,621	2.22	1.05	3.57	69.87	57.17	18.83	1.37	29.74
CHLO	1,637,621	2.01	0.85	3.56	74.26	58.77	17.08	1.27	30.97
OHL	1,637,621	2.35	1.21	3.65	69.95	54.64	20.47	1.49	29.97
CHL	1,637,621	2.16	1.03	3.67	73.83	55.44	18.93	1.41	31.30
AR	1,637,621	1.70	0.95	2.50	68.13	53.55	19.90	1.41	31.87
CS	1,637,621	0.66	0.28	1.10	45.55	33.77	35.90	2.61	29.18
ROLL	1,637,621	2.47	1.39	4.09	55.22	41.38	24.53	1.80	32.60
HJ	1,637,621	1.89	0.75	2.73	–	–	–	–	–

The table reports summary statistics of stock-month spread estimates from daily prices in the sample period 1993–2021 (CRSP-TAQ merged sample). Negative spread estimates are set to zero, and we drop the stock-month estimate for all the estimators if it is missing for any of them. The table reports the number of stock-months (N), the mean (Mean), median (Med), and standard deviation (Sd) of the estimates, their Pearson's (Cor_P) and Spearman's (Cor_S) correlation with the HJ benchmark, the mean absolute percentage error (MAPE) and the root mean squared error (RMSE) computed on the log-spreads (see Internet Appendix I.4), and the Fraction of Non-Positive Estimates (FNPE). The highest correlations, the lowest errors, and the lowest fraction of non-positive estimates are in bold.

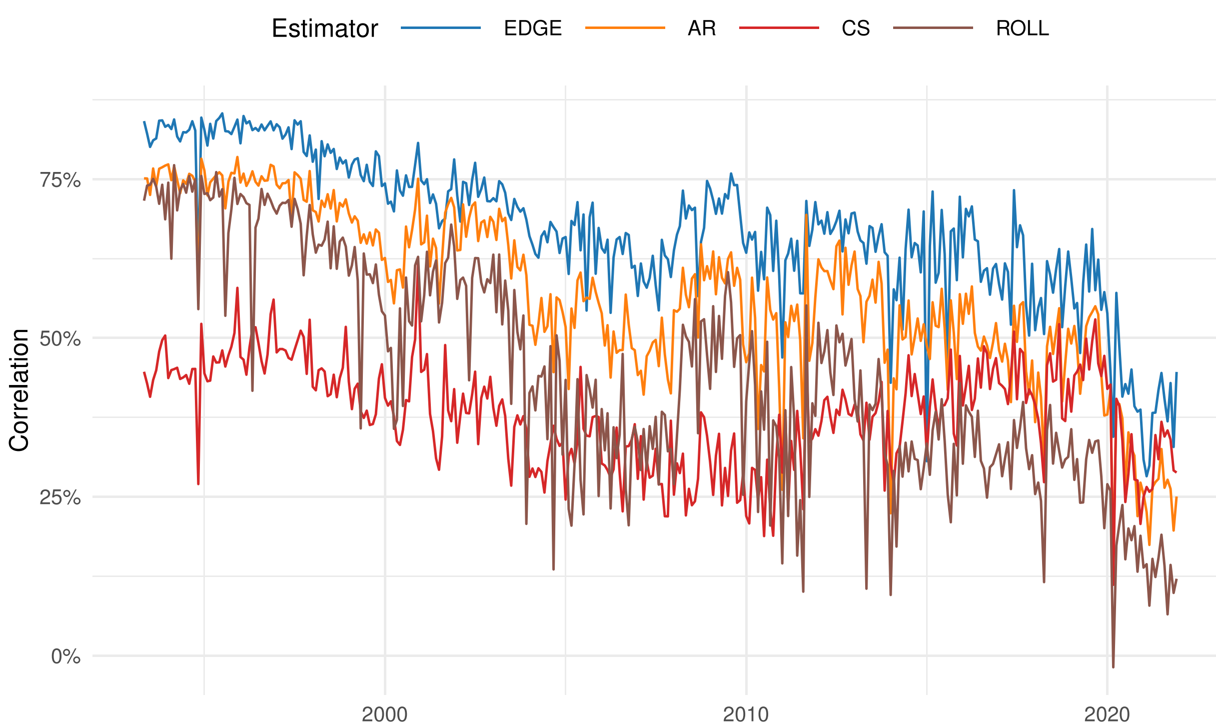

Fig. 4. Cross-sectional correlation with the HJ benchmark.
The figure shows cross-sectional Pearson's correlations between stock-month spread estimates from daily prices and the HJ benchmark for each month in the sample period 1993–2021 (CRSP-TAQ merged sample). Negative spread estimates are set to zero, and we drop the stock-month estimate for all the estimators if it is missing for any of them.

of the stock. For the trading frequency, we split stocks based on their average number of daily trades during the whole sample period. Then, given the effective spread benchmark $S_{i,t}$ for stock i at time t and the corresponding estimate $\hat{S}_{i,t}$, we compute the correlation for group g as $\rho_g = \mathbb{C}\text{or}_{(i,t)\in g}[S_{i,t}, \hat{S}_{i,t}]$.

The results are summarized in Table 4 for market venues (Panel A), time periods (Panel B), market capitalization (Panel C), spread size (Panel D), and trading frequency (Panel E). One clear result emerges: EDGE outperforms all the alternative estimators in each market venue, sub-period, market capitalization, spread size, and for each trading frequency by consistently achieving the highest correlation with the effective spread benchmark. To shed light on the performance of EDGE, we also report the behavior of its building blocks. In particular, it is natural to compare CHL with AR as they use the same information set of high, low, and close prices. As the main difference between the two estimators is that CHL accounts for infrequent trading, the outperformance of CHL compared to AR demonstrates the importance of relaxing the assumption that prices are observed continuously to ultimately improve empirical results. Table 4 further shows that any single building block OHL, CHL, OHLC, CHLO outperforms AR, CS, and ROLL. Finally, EDGE optimally combines its building blocks to provide an estimator that is superior to any of them taken individually.

In the Internet Appendix, we provide representative illustrations for individual stocks to investigate the estimators' performance further. We also compare the estimators using additional evaluation metrics such as Spearman's (rank) correlation, MAPE and RMSE, and the fraction of zero estimates. Overall, EDGE achieves the highest rank correlation with the benchmark, the lowest MAPE and RMSE, and generates the lowest fraction of non-positive estimates. We also find that it achieves the best results when estimating first differences instead of spread levels and when increasing the estimation window from one month to one year. It is also interesting to note that CS achieves a slightly lower MAPE and RMSE compared to EDGE in the following cases: (a) NYSE stocks, (b) recent periods, (c) large stocks, (d) small spreads, and (e)

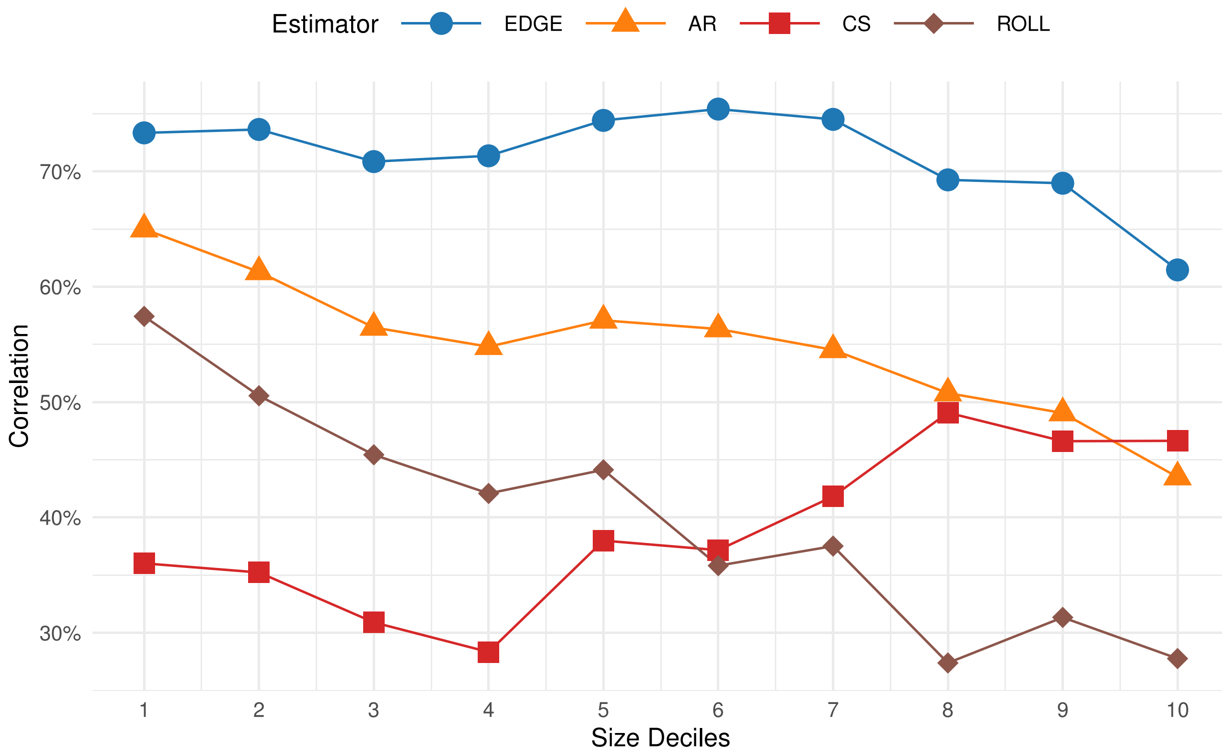

Fig. 5. Time-series correlation with the HJ benchmark.
The figure shows time-series Pearson's correlations between stock-month spread estimates from daily prices and the HJ benchmark for size deciles in the sample period 1993–2021 (CRSP-TAQ merged sample). Size deciles are sorted by increasing market capitalization at the last observed period for each individual stock. Negative spread estimates are set to zero, and we drop the stock-month estimate for all the estimators if it is missing for any of them.

frequent trading. Taken together, these are all cases where the bid–ask spread is expected to be small and where the downward bias of the CS estimator may improve the estimate. Indeed, if the spread is expected to be small, then an estimator biased towards zero may yield better results. This observation suggests that, generally, a Bayesian approach may further improve the estimate when a good prior is available for specific applications. We leave such possibility for future research as we focus on an estimator of general applicability here.

6. Applications

To demonstrate the broad applicability of EDGE, we provide three representative examples. The first revisits historical spread estimates from daily prices in the U.S. stock market since 1926. The second studies spread estimates obtained from intraday prices for U.S. stocks. Finally, the third applies the estimator outside the U.S. stock market and compares low- and high-frequency estimates for cryptocurrencies.

6.1. Low-frequency estimates for the U.S. stock market

Using CRSP data since 1926, we construct, for each month, three portfolios based on size according to the following procedure. First, we sort the stocks based on their market capitalization at the end of each month. Then, we select small-caps, mid-caps, and large-caps using the 50th and 80th percentiles as breakpoints. Finally, we compute monthly spread estimates for individual stocks and construct the average spread for each of the three portfolios in each month between 1926–1992 (CRSP sample) and 1993–2021 (CRSP-TAQ merged sample).[6] The results are reported in Fig. 6.

Panel A displays the cross-sectional mean of spread estimates for small stocks. According to EDGE, the spread was high in the 1930s, spiked in 1933 with peaks between 10%–15%, decreased until the 1960s and increased again with a first peak of about 5% in 1963, a second peak of 7.5% in 1975, and a third peak of 10% in the early 1990s. In line with the idea that liquidity evaporates in times of crisis (Nagel, 2012), these years coincide with periods of financial downturn and economic recession, such as the Great Depression between 1929–41, the U.S. Banking Crisis of 1933, the Kennedy Slide of 1962, the 1973–1975 recession following the oil crisis, and the early 1990s recession in the United States. Following the electronization of financial markets in the 2000s, the spread decreased significantly until the global financial crisis, when it spiked again in 2009 with a peak close to 5%. The spread has continued to reduce in the last decade, reaching the lowest level ever as of December 2021. In the CRSP-TAQ merged sample after 1993, the HJ benchmark closely follows this trend and overlaps with EDGE. Instead, CS and AR tend to underestimate the spread, particularly for older periods, mirroring that these estimators are biased when trading becomes increasingly infrequent. In the historical sample before 1993, we find that the gap between EDGE and the alternative estimators widens. EDGE is often larger than AR by a factor of two, and the difference is even more pronounced compared to CS. Given our benchmark result from the recent sample, we conjecture that the alternative estimators considerably underestimate the effective spread in the early sample.

Panels B and C report the results for medium and large stocks, respectively. As expected, we find that larger stocks tend to have lower spreads than smaller stocks. Indeed, EDGE estimates an average spread that is typically below 2.5% for medium stocks and below 1% for large stocks. The gap with AR and CS decreases for larger stocks, mirroring that their bias reduces for stocks presumably traded more frequently.

Fig. 6 also reports end-of-day quoted spreads derived from CRSP. These spreads are significantly higher than the effective spread benchmark in the sample period between 1993 and the early 2000s. The

[6] When EDGE cannot be computed, we use the CHL estimator in Table 1 that does not need open prices. Open prices are missing in CRSP from July 1962 through June 1992.

Table 4
Pearson's correlation with the HJ benchmark.

	EDGE	OHLC	CHLO	OHL	CHL	AR	CS	ROLL
Panel A: Analysis by Market Exchange								
NYSE	**64.94**	53.22	61.84	52.15	58.43	46.79	45.87	29.59
AMEX	**68.99**	57.77	67.84	59.26	68.23	61.05	38.32	48.15
NASDAQ	**78.16**	68.62	73.24	68.87	72.99	67.03	41.09	54.81
Panel B: Analysis by Time Period								
1993–1996	**82.93**	75.94	76.83	77.68	78.57	75.31	46.94	70.23
1997–2000	**78.47**	68.24	73.40	68.95	73.41	69.20	45.06	60.19
2001–2002	**73.04**	60.47	70.32	61.44	69.59	67.31	40.70	59.00
2003–2007	**67.65**	57.56	63.45	57.26	61.97	57.34	33.87	38.10
2008–2011	**69.89**	62.16	64.00	61.49	62.26	59.17	33.99	43.70
2012–2015	**60.78**	51.41	55.93	52.09	55.64	53.14	37.29	29.24
2016–2021	**53.98**	46.72	43.97	46.48	43.11	40.78	39.34	22.11
Panel C: Analysis by Market Capitalization								
Size quintile 1	**74.35**	63.60	69.90	64.86	70.44	65.00	37.16	56.08
Size quintile 2	**71.29**	60.36	66.68	60.39	66.33	56.08	30.20	44.31
Size quintile 3	**75.13**	65.09	70.32	63.12	67.28	57.22	38.41	40.11
Size quintile 4	**72.55**	62.93	68.07	59.63	63.44	53.02	44.60	32.90
Size quintile 5	**66.65**	57.77	61.32	54.24	56.17	47.31	47.24	30.31
Panel D: Analysis by Spread Size								
Spread quintile 1	**17.84**	15.62	16.56	15.43	15.21	14.18	12.64	9.60
Spread quintile 2	**45.66**	39.59	41.73	34.67	34.15	30.35	32.79	15.06
Spread quintile 3	**61.98**	52.28	57.80	48.88	52.46	44.72	40.82	24.64
Spread quintile 4	**67.76**	55.55	64.44	55.32	63.21	55.22	37.74	38.98
Spread quintile 5	**71.38**	60.78	66.08	62.57	67.24	61.83	33.15	55.06
Panel E: Analysis by Trading Frequency								
Numtrd quintile 1	**74.77**	65.88	69.12	67.68	70.37	67.81	40.02	65.10
Numtrd quintile 2	**79.15**	69.32	74.45	69.55	74.17	69.58	51.59	52.00
Numtrd quintile 3	**75.41**	65.77	70.94	63.92	67.92	60.98	50.42	40.53
Numtrd quintile 4	**67.17**	58.53	62.78	56.00	58.71	52.54	48.98	32.41
Numtrd quintile 5	**55.48**	48.05	50.02	45.23	45.78	39.36	43.91	22.29

The table reports Pearson's correlations (in %) with the HJ benchmark for stock-month spread estimates from daily prices in the sample period 1993–2021 (CRSP-TAQ merged sample). The highest correlation per group is in bold. Negative spread estimates are set to zero, and we drop the stock-month estimate for all the estimators if it is missing for any of them. The size quintiles are sorted by increasing market capitalization at the last observed period for each individual stock. The spread quintiles are sorted by increasing average HJ spreads during the whole sample period. The trade quintiles are sorted by increasing average number of daily trades during the whole sample period.

historical sample also supports this finding before 1993, where quoted spreads are often higher than EDGE by a factor of two. We thus confirm earlier studies that the quoted spread overstates the effective spread finally paid by traders by up to 100% (Huang and Stoll, 1994; Petersen and Fialkowski, 1994; Bessembinder and Kaufman, 1997; Bacidorea et al., 2003), due to dealers offering a better price than the quotes, also known as trading inside the spread (Lee, 1993). We also find that quoted and effective spreads closely overlap in the last two decades, suggesting that this phenomenon has reduced over time and quoted spreads have become a better proxy of effective spreads following the electronization of financial markets.

Finally, we notice that estimating spreads from daily prices leads to an upward bias that becomes increasingly evident in more recent periods and for larger stocks. For instance, EDGE estimates an average spread of 0.42% in December 2021 for large-caps, while the HJ benchmark is 0.06%. This bias arises in small samples due to the practice of resetting negative estimates to zero, which leads, on average, to overstating the spread, especially when the spread is small compared to volatility (Jahan-Parvar and Zikes, 2023). A way to mitigate this small-sample bias is to extend the estimation window with more daily observations (see Internet Appendix I.5). Another way to improve the estimation is using intraday prices whenever they are available, as discussed in the next section.

6.2. High-frequency estimates for the U.S. stock market

While the variance component of an asset return is proportional to the return interval, the spread component is not. Hence, we can rely on high-frequency prices to reduce the asset variance without altering the spread component and achieve a better signal-to-noise ratio to improve the spread estimate.

For instance, let N be the sample size and consider estimates derived from a monthly sample of daily data ($N = 21$), a yearly sample of daily data ($N = 252$), or a monthly sample of minute data ($N = 21 \times 390$). According to Eq. (17), the estimation variance is roughly proportional to σ_1^4/N where σ_1 is the volatility per period, and the standard error is proportional to $\sigma_1^2/\sqrt{N}$. For daily prices, $\sigma_1 = \sigma/\sqrt{252}$ where σ is the volatility per year. For minute prices, $\sigma_1 = \sigma/\sqrt{252 \times 390}$. Thus, estimates derived from a yearly sample of daily data have a standard error that is $\sqrt{252/21} = 3.5$ times smaller than that obtained from a monthly sample of daily data. Estimates derived from a monthly sample of minute data have a standard error that is $390^{3/2} = 7702$ times smaller. To put this in perspective, the enhancement factor of the sample using minute prices would be achieved by a sample of 1,245,699,037 daily prices, equivalent to approximately 5 million years of trading. From this analysis, we conclude that using intraday prices offers a more effective way to improve the spread estimate than increasing the sample size with more daily data.

To illustrate how EDGE can substantially improve the estimation of bid–ask spreads using intraday prices, we proceed as follows. First, we aggregate trades into open, high, low, and close prices every minute using the Daily TAQ database from October 2003 to December 2021. Then, we estimate the spread with EDGE from the minute data for each stock-month. Finally, we compare the estimates derived this way with those derived from daily prices and the HJ benchmark.

Fig. 7 reports the results for large-cap stocks. These stocks are featured by tiny spreads that are difficult to estimate in small samples due to their small signal-to-noise ratio, which causes a large fraction of

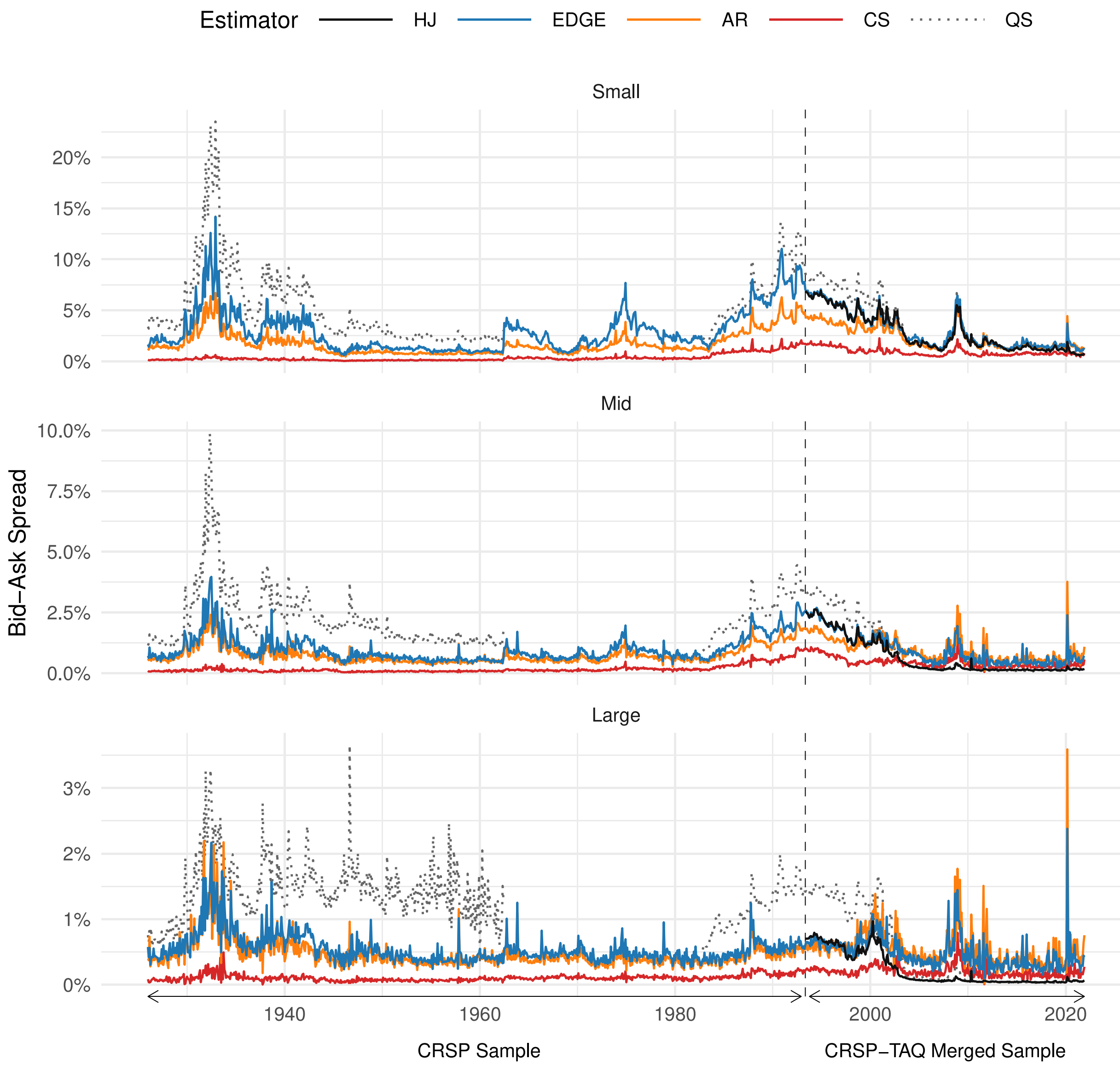

Fig. 6. Low-frequency estimates for U.S. stocks.
The figure reports the average spread across stocks for each month and size group from 1926 to 2021. Stocks are sorted into small-caps, mid-caps, and large-caps based on their market capitalization at the end of each month using the 50th and 80th percentiles as breakpoints. Spreads are estimated for each stock-month using daily prices. Negative spread estimates are set to zero, and we drop the stock-month estimate for all the estimators if it is missing for any of them. EDGE is replaced with CHL when open prices are missing in CRSP. End-of-day quoted spreads (QS) are missing from July 1962 to October 1982. The HJ benchmark obtained from TAQ data is available since May 1993.

non-positive estimates and generates an upward bias due to the practice of resetting negative estimates to zero (Jahan-Parvar and Zikes, 2023). Indeed, we find that the EDGE estimates from daily prices are negative in 41% of stock-months, and they are higher than the HJ benchmark by 0.35% (35bps) on average. Instead, estimates derived from minute prices are negative in only 0.05% of stock-months, and their upward bias shrinks to zero (1bps).

Next, we analyze the estimates for all stocks. This sample consists of 711,161 stock-month spread estimates derived from both minute and daily prices. We find that using minute prices reduces the fraction of negative estimates from 34.15% to 0.02% and significantly improves all evaluation metrics. The Pearson's (Spearman's) correlation with the HJ benchmark increases from 56.17% (43.47%) to 88.79% (97.31%). The MAPE (RMSE) reduces from 23.68 (1.80) to 5.17 (0.41).

Finally, we estimate spreads from minute prices using the Monthly TAQ database from May 1993 to July 2014. The Monthly TAQ data are identical to the Daily TAQ data except for two main differences. First, Monthly TAQ only reports raw quotes, while Daily TAQ includes an NBBO file that reports the highest bid price and lowest ask price among all available exchanges at each timestamp. Second, Monthly TAQ data are timestamped to the second while Daily TAQ data are timestamped to the millisecond. While these differences cause several problems in measuring effective spreads by matching trades with quotes (Holden and Jacobsen, 2014), they do not affect EDGE. Indeed, the correlation between the EDGE estimates obtained with Monthly TAQ and Daily TAQ in the overlapping period between October 2003 and July 2014 is 99.8%. This suggests that, by relying on transaction prices only, EDGE is more robust than measuring effective spreads by matching trades with quotes, and it is less sensitive to the quality of the data.

In summary, low-frequency estimates can be substantially improved using intraday prices. This is particularly relevant for cases where high-frequency prices are available, but quotes are not, or they cannot be reliably matched with trades. Examples include, but are not limited to, over-the-counter markets, dark pools, and crypto exchanges.

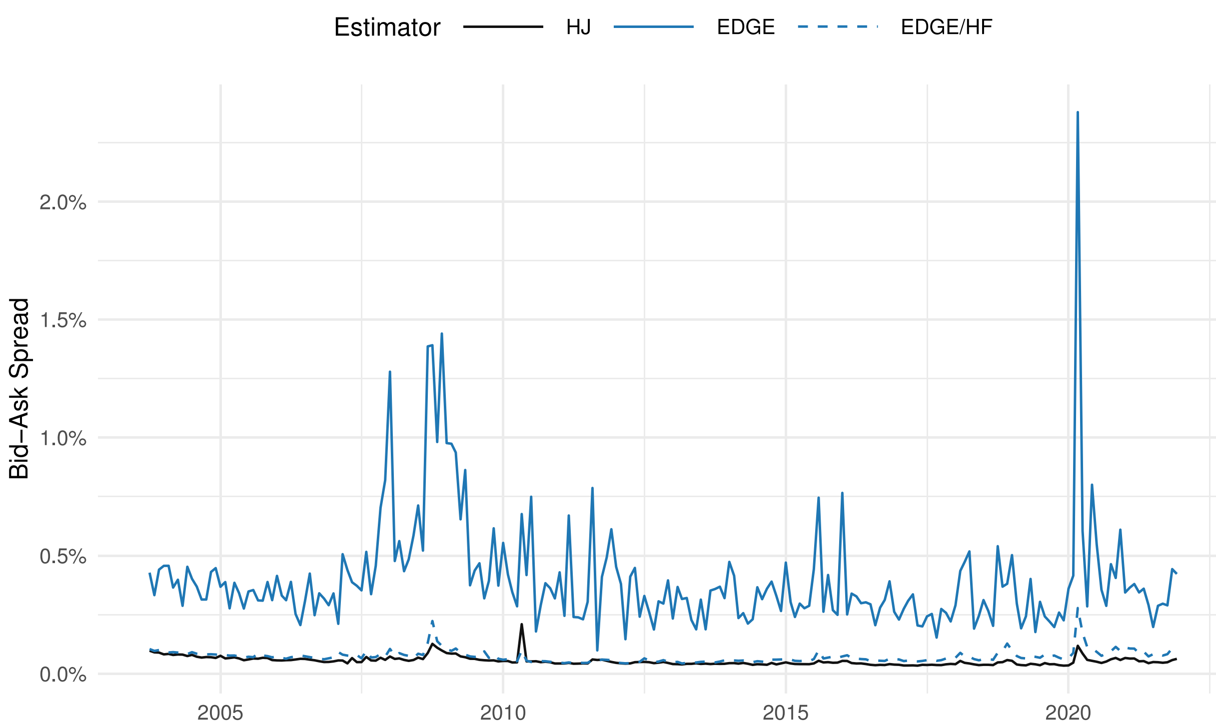

Fig. 7. High-frequency estimates for U.S. stocks.
The figure reports the average spread across large-cap stocks for each month from October 2003 to December 2021. Spreads are estimated for each stock-month using daily (EDGE) or minute (EDGE/HF) prices. Negative spread estimates are set to zero, and we drop the stock-month estimate for all the estimators if it is missing for any of them. HJ is the benchmark spread obtained from TAQ data.

6.3. Estimates for other markets

Our estimator represents a general way to estimate effective spreads, and it is designed to be applied to a variety of markets. To illustrate its applicability outside the U.S. stock market, we analyze estimates for cryptocurrency pairs listed in Binance.

Binance is a leading crypto exchange listing hundreds of cryptocurrencies that can be exchanged for one another via trading pairs. Each trading pair (*e.g.*, ETH/BTC) reports the price of the base currency (*e.g.*, ETH) in units of the quote currency (*e.g.*, BTC). Like other crypto exchanges, Binance provides historical and real-time daily and intraday prices for free, while trade and quote data are subject to subscription fees, and their historical coverage is more limited. As trade and quote data are unavailable to us, we cannot compute bid–ask spreads obtained by matching trades with quotes.

To estimate effective spreads from freely available data, we download historical open, high, low, and close prices for all cryptocurrency pairs at the minute, hourly, and daily frequency. We then compute monthly estimates with EDGE, AR, and CS for each pair and each frequency and drop the estimate for all estimators if missing for any of them. Our sample consists of 2163 crypto pairs and 53,865 pair-month spread estimates for each frequency and estimator in the sample period from July 2017 to December 2021.

We expect AR and CS to overstate the spread when using daily prices due to the upward bias induced by resetting negative estimates to zero. When using intraday prices, we expect them to understate the spread because the number of trades per period reduces at higher frequencies, and their downward bias shrinks the estimate to zero. Instead, we expect EDGE to mitigate these two concerns because its lower variance reduces the upward bias, and the estimator is unaffected by the downward bias due to infrequent trading.

Fig. 8 reports the time evolution of the average spread across all trading pairs for each estimator. As expected, AR and CS produce different estimates depending on the sampling frequency. Estimates derived from daily prices are significantly higher than those derived from hourly prices, which, in turn, are higher than those derived from minute prices. Depending on the frequency used, the average spread in the whole sample period ranges anywhere between 0.18% (0.02%) and 1.85% (1.45%) according to AR (CS). This tenfold difference makes it impossible to estimate the spread reliably because it is unclear which sampling frequency should be preferred in principle. Instead, EDGE produces estimates less sensitive to the sampling frequency, and estimates from daily prices closely overlap with those from hourly and minute prices. The average spread in the whole sample period remains in the narrow range between 0.68% and 0.70%, depending on whether minute, hourly, or daily prices are used. In 2021, we find that the average spread for crypto pairs is between 0.35%–0.45%.

In summary, EDGE is less sensitive to the sampling frequency than other estimators and can potentially reduce a large source of non-standard errors (Menkveld et al., 2024) in the measurement of transaction costs.

7. Conclusion

Historically, the development of bid–ask spread estimators has evolved along two complementary paths that consider either high-frequency or low-frequency data. The former exploits trade and quote data to obtain an explicit proxy of the fundamental price and measure the distance of transaction prices from it. The latter introduces assumptions about the fundamental price to derive an estimator from transaction prices only. While estimates derived from trades and quotes are typically more accurate, low-frequency estimates are more readily available and are becoming increasingly popular. However, low-frequency estimators assume that prices are observed continuously. Here, we document that these approaches lead to understating effective spreads, especially for infrequently traded assets that should

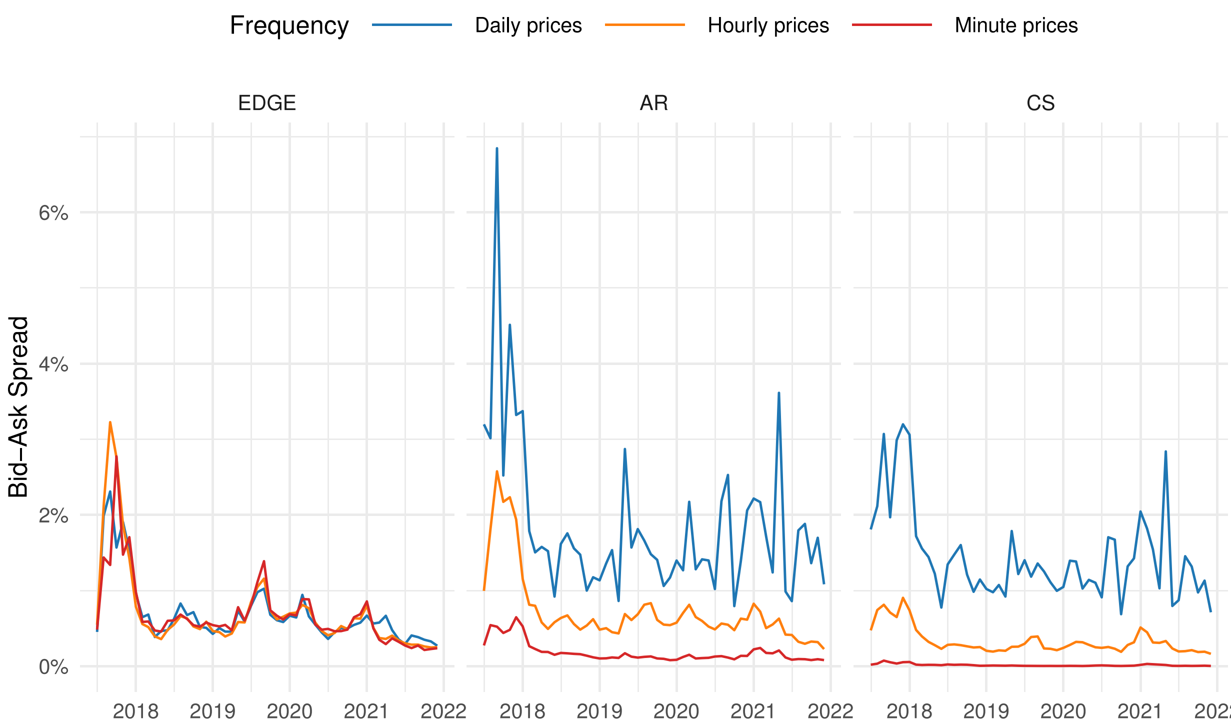

Fig. 8. Low- and high-frequency estimates for Cryptocurrencies.
The figure reports the average spread across trading pairs listed in Binance for each month from July 2017 to December 2021. Spreads are estimated for each pair-month using daily, hourly, or minute prices. Negative spread estimates are set to zero, and we drop the pair-month estimate for all the estimators if it is missing for any of them.

presumably be associated with high transaction costs. We then develop a novel methodology relaxing the assumption that prices are observed continuously and derive generalized estimators that correct this downward bias analytically. We show that different estimators are preferable depending on whether the spread is large or small compared to volatility, and we combine them efficiently to produce an unbiased estimator with minimum variance. Through theoretical analyses, numerical simulations, and empirical evaluations, we find that our efficient estimator dominates each generalized estimator taken individually and other estimators from transaction prices.

Our efficient estimator has broad applicability for several reasons. First, it is derived under more general assumptions than other approaches and extends the domain of applicability to various assets and time periods. Second, the estimator is unaffected by the downward bias due to infrequent trading and makes it possible to estimate effective spreads for assets traded infrequently, for historical periods, or using high-frequency prices when quotes are unreliable or unavailable. Third, the estimator minimizes the estimation variance and thus also minimizes the upward bias that arises from resetting negative estimates to zero in small samples (Jahan-Parvar and Zikes, 2023).

Our results show that other estimators significantly understate effective spreads in the 20th century, while end-of-day quoted spreads overstate effective spreads by up to 100%. Thus, this work makes available the most realistic effective spread estimates for the U.S. stock market from 1926 to the advent of high-frequency data. We further show that our estimator can substantially improve estimates from daily prices using intraday prices, while other estimators are dominated by their downward bias because trading becomes sparse in high frequency. To demonstrate the generalizability of these results outside the U.S. stock market, we estimate bid–ask spreads for cryptocurrencies. Our efficient estimator produces consistent estimates regardless of whether daily or intraday prices are used, while other estimators produce a tenfold difference between daily and intraday estimates. We conclude that our estimator may reduce a significant source of non-standard errors in applied research (Menkveld et al., 2024).

Finally, we provide guidance for future research aimed at estimating transaction costs. First, we have shown that the assumption that prices are observed continuously has far-reaching implications and causes biases that generally vary in the cross-section and time series, and they also depend on the sampling frequency of open, high, low, and close prices. Future work should explicitly account for discretely observed prices to avoid this source of bias. Second, our estimator can be applied at any frequency, and, in this sense, it reconciles the high-frequency and low-frequency literature. For this reason, we argue that a better classification is distinguishing between methods that require trade and quote data and those that require transaction prices only. Third, we have constructed an efficient estimator in the class of covariance-based estimators from open, high, low, and close prices. To design more efficient estimators, future work could either consider approaches that are not based on the serial covariance of returns or exploit information other than prices, such as, for instance, the trading volume or a suitable Bayesian prior.

CRediT authorship contribution statement

David Ardia: Funding Acquisition, Resources, Validation, Writing – Original Draft. **Emanuele Guidotti:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing. **Tim A. Kroencke:** Conceptualization, Project Administration, Resources, Supervision, Validation, Writing – Original Draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. Tim A. Kroencke is Member of the Academic Advisory Council of Vontobel Bank, Zurich.

Data availability

Replication Package for "Efficient Estimation of Bid-Ask Spreads from Open, High, Low, and Close Prices" (Harvard Dataverse)

Appendix A. Proofs

A.1. Proof of Eq. (10)

The de-meaned returns defined in Eq. (9) have mean zero conditional on τ_t, for any return r_t computed in the time interval between the end of period $t-1$ and the end of period t. Indeed, r_t is identically zero conditional on $\tau_t = 0$ because $h_t = l_t = c_{t-1}$ and thus $\mathbb{E}[\overline{r}_t \mid \tau_t = 0] = 0$. Moreover, $\mathbb{E}[\overline{r}_t \mid \tau_t = 1] = \mathbb{E}[r_t \mid \tau_t = 1] - \mathbb{E}[r_t]/\mathbb{E}[\tau_t] = 0$ because $\mathbb{E}[r_t \mid \tau_t = 0] = 0$. In summary, it holds that $\mathbb{E}[\overline{r}_t \mid \tau_t] = 0$ and using the law of total covariance we have:

$$\begin{aligned}\mathbb{C}\mathrm{ov}[\overline{r}_t, r_s] &= \mathbb{E}[\mathbb{C}\mathrm{ov}[\overline{r}_t, r_s \mid \tau_t]] + \mathbb{C}\mathrm{ov}[\mathbb{E}[\overline{r}_t \mid \tau_t], \mathbb{E}[r_s \mid \tau_t]] \\ &= \mathbb{E}[\mathbb{C}\mathrm{ov}[\overline{r}_t, r_s \mid \tau_t]] \\ &= \mathbb{C}\mathrm{ov}[\overline{r}_t, r_s \mid \tau_t = 1]\mathbb{P}[\tau_t = 1] + \mathbb{C}\mathrm{ov}[\overline{r}_t, r_s \mid \tau_t = 0]\mathbb{P}[\tau_t = 0] \\ &= \mathbb{C}\mathrm{ov}[r_t, r_s \mid \tau_t = 1]\mathbb{P}[\tau_t = 1]\,.\end{aligned} \tag{A.1}$$

The last equality follows from the fact that $\overline{r}_t = r_t = 0$ conditional on $\tau_t = 0$, while $\overline{r}_t = r_t + const.$ conditional on $\tau_t = 1$ and the constant is irrelevant for the calculation of the covariance.

A.2. Proof of Eq. (14)

We need to compute:

$$\mathbb{E}[Z_{\eta_t} Z_{o_t} \mid \tau_t = 1] = \frac{\mathbb{E}[Z_{h_t} Z_{o_t} \mid \tau_t = 1] + \mathbb{E}[Z_{l_t} Z_{o_t} \mid \tau_t = 1]}{2}\,. \tag{A.2}$$

We start by considering high prices, and we condition on whether or not the opening price o_t is equal to the highest price h_t:

$$\begin{aligned}\mathbb{E}[Z_{h_t} Z_{o_t} \mid \tau_t = 1] &= \mathbb{E}[Z_{h_t} Z_{o_t} \mid o_t = h_t, \tau_t = 1]\mathbb{P}[o_t = h_t \mid \tau_t = 1] \\ &\quad + \mathbb{E}[Z_{h_t} Z_{o_t} \mid o_t \neq h_t, \tau_t = 1]\mathbb{P}[o_t \neq h_t \mid \tau_t = 1]\,.\end{aligned} \tag{A.3}$$

If $o_t = h_t$, then the opening price is selected as the highest price in the period, and the bid–ask bounces $Z_{h_t} = Z_{o_t} = S_{o_t}/2D_{o_t}$ coincide. Thus, we have:

$$\mathbb{E}[Z_{h_t} Z_{o_t} \mid o_t = h_t, \tau_t = 1] = \mathbb{E}[S^2_{o_t}]/4\,. \tag{A.4}$$

If $o_t \neq h_t$, then Z_{h_t} and Z_{o_t} are uncorrelated by Assumption 3 and:

$$\mathbb{E}[Z_{h_t} Z_{o_t} \mid o_t \neq h_t, \tau_t = 1] = \mathbb{E}[Z_{h_t} \mid o_t \neq h_t, \tau_t = 1]\mathbb{E}[Z_{o_t} \mid o_t \neq h_t, \tau_t = 1] = 0\,, \tag{A.5}$$

because $\mathbb{E}[Z_{o_t} \mid o_t \neq h_t, \tau_t = 1] = \mathbb{E}[Z_{o_t}] = 0$ if we consider that the bid–ask bounce at the open is independent from whether the opening price is the highest price in the period. Substituting Eqs. (A.4)–(A.5) into Eq. (A.3) gives:

$$\mathbb{E}[Z_{h_t} Z_{o_t} \mid \tau_t = 1] = \mathbb{E}[S^2_{o_t}]\mathbb{P}[o_t = h_t \mid \tau_t = 1]/4\,. \tag{A.6}$$

The same equation holds for low prices by replacing Z_{h_t} with Z_{l_t} and h_t with l_t. Substituting Eq. (A.6) for high and low prices into Eq. (A.2) yields Eq. (14).

A.3. Proof of Eq. (15)

Substituting Eqs. (13)–(14) in Eq. (12) yields:

$$\begin{aligned}&\mathbb{E}[(\overline{\eta_t - o_t})(o_t - c_{t-1})] \\ &= \frac{\mathbb{E}[S^2_{o_t}]}{4}\left(\frac{\mathbb{P}[o_t = h_t \mid \tau_t = 1] + \mathbb{P}[o_t = l_t \mid \tau_t = 1]}{2} - 1\right)\mathbb{P}[\tau_t = 1] \\ &= \frac{\mathbb{E}[S^2_{o_t}]}{4}\left(-\frac{\mathbb{P}[o_t \neq h_t \mid \tau_t = 1] + \mathbb{P}[o_t \neq l_t \mid \tau_t = 1]}{2}\right)\mathbb{P}[\tau_t = 1] \\ &= \frac{\mathbb{E}[S^2_{o_t}]}{4}\left(-\frac{\mathbb{P}[o_t \neq h_t, \tau_t = 1] + \mathbb{P}[o_t \neq l_t, \tau_t = 1]}{2}\right).\end{aligned} \tag{A.7}$$

Solving Eq. (A.7) for $\mathbb{E}[S^2_{o_t}]$ gives Eq. (15).

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jfineco.2024.103916.

References

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

Amihud, Y., Noh, J., 2020. Illiquidity and stock returns II: Cross-section and time-series effects. Rev. Financ. Stud. 34 (4), 2101–2123.

Bacidorea, J., Ross, K., Sofianosa, G., 2003. Quantifying market order execution quality at the New York stock exchange. J. Financial Mark. 6, 281–307.

Bali, T.G., Subrahmanyam, A., Wen, Q., 2021. Long-term reversals in the corporate bond market. J. Financ. Econ. 139 (2), 656–677.

Bessembinder, H., Kaufman, H.M., 1997. A comparison of trade execution costs for NYSE and NASDAQ-listed stocks. J. Financ. Quant. Anal. 32 (3), 287–310.

Birru, J., 2018. Day of the week and the cross-section of returns. J. Financ. Econ. 130 (1), 182–214.

Bongaerts, D., de Jong, F., Driessen, J., 2017. An asset pricing approach to liquidity effects in corporate bond markets. Rev. Financ. Stud. 30 (4), 1229–1269.

Brogaard, J., Pan, J., 2021. Dark pool trading and information acquisition. Rev. Financ. Stud. 35 (5), 2625–2666.

Cai, F., Han, S., Li, D., Li, Y., 2019. Institutional herding and its price impact: Evidence from the corporate bond market. J. Financ. Econ. 131 (1), 139–167.

Chaieb, I., Errunza, V., Langlois, H., 2020. How is liquidity priced in global markets? Rev. Financ. Stud. 34 (9), 4216–4268.

Chen, Y., Eaton, G.W., Paye, B.S., 2018. Micro(structure) before macro? The predictive power of aggregate illiquidity for stock returns and economic activity. J. Financ. Econ. 130 (1), 48–73.

Chen, X., Linton, O., Yi, Y., 2017. Semiparametric identification of the bid–ask spread in extended Roll models. J. Econometrics 200 (2), 312–325.

Choi, J., Hoseinzade, S., Shin, S.S., Tehranian, H., 2020. Corporate bond mutual funds and asset fire sales. J. Financ. Econ. 138 (2), 432–457.

Corwin, S.A., Schultz, P., 2012. A simple way to estimate bid-ask spreads from daily high and low prices. J. Finance 67 (2), 719–760.

Ding, Y., Xiong, W., Zhang, J., 2022. Issuance overpricing of China's corporate debt securities. J. Financ. Econ. 144 (1), 328–346.

Easley, D., Lopez de Prado, M., O Hara, M., Zhang, Z., 2020. Microstructure in the machine age. Rev. Financ. Stud. 34 (7), 3316–3363.

Eaton, G.W., Irvine, P.J., Liu, T., 2021. Measuring institutional trading costs and the implications for finance research: The case of tick size reductions. J. Financ. Econ. 139 (3), 832–851.

Garman, M.B., Klass, M.J., 1980. On the estimation of security price volatilities from historical data. J. Bus. 53, 67–78.

Goldstein, I., Jiang, H., Ng, D.T., 2017. Investor flows and fragility in corporate bond funds. J. Financ. Econ. 126 (3), 592–613.

Grosse-Rueschkamp, B., Steffen, S., Streitz, D., 2019. A capital structure channel of monetary policy. J. Financ. Econ. 133 (2), 357–378.

Hagströmer, B., 2021. Bias in the effective bid-ask spread. J. Financ. Econ. 142 (1), 314–337.

Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.

Hasbrouck, J., 2009. Trading costs and returns for U.S. equities: Estimating effective costs from daily data. J. Finance 64 (3), 1445–1477.

Holden, C.W., Jacobsen, S., 2014. Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. J. Finance 69 (4), 1747–1785.

Hou, K., Xue, C., Zhang, L., 2018. Replicating anomalies. Rev. Financ. Stud. 33 (5), 2019–2133.

Hua, J., Peng, L., Schwartz, R.A., Alan, N.S., 2019. Resiliency and stock returns. Rev. Financ. Stud. 33 (2), 747–782.

Huang, R.D., Stoll, H.R., 1994. Market microstructure and stock return predictions. Rev. Financ. Stud. 7 (1), 179–213.

Jacobs, H., Müller, S., 2020. Anomalies across the globe: Once public, no longer existent? J. Financ. Econ. 135 (1), 213–230.

Jahan-Parvar, M.R., Zikes, F., 2023. When do low-frequency measures really measure effective spreads? evidence from equity and foreign exchange markets. Rev. Financ. Stud. 36, 4190–4232.

Kaviani, M.S., Kryzanowski, L., Maleki, H., Savor, P., 2020. Policy uncertainty and corporate credit spreads. J. Financ. Econ. 138 (3), 838–865.

Lee, C.M.C., 1993. Market integration and price execution for NYSE-listed securities. J. Finance 48 (3), 1009–1038.

Li, X., Subrahmanyam, A., Yang, X., 2018. Can financial innovation succeed by catering to behavioral preferences? evidence from a callable options market. J. Financ. Econ. 128 (1), 38–65.

Loon, Y.C., Zhong, Z.K., 2016. Does dodd-frank affect OTC transaction costs and liquidity? evidence from real-time CDS trade reports. J. Financ. Econ. 119 (3), 645–672.

McLean, R.D., Pontiff, J., 2016. Does academic research destroy stock return predictability? J. Finance 71 (1), 5–32.

Menkveld, A.J., Dreber, A., Holzmeister, F., Huber, J., Johannesson, M., Kirchler, M., Razen, M., Weitzel, U., Abad, D., Abudy, M.M., et al., 2024. Nonstandard errors. J. Finance 79 (3), 2339–2390.

Michaelides, A., Milidonis, A., Nishiotis, G.P., 2019. Private information in currency markets. J. Financ. Econ. 131 (3), 643–665.

Nagel, S., 2012. Evaporating liquidity. Rev. Financ. Stud. 25 (7), 2005–2039.

Parkinson, M., 1980. The extreme value method for estimating the variance of the rate of return. J. Bus. 53, 61–65.

Patton, A.J., Weller, B.M., 2020. What you see is not what you get: The costs of trading market anomalies. J. Financ. Econ. 137 (2), 515–549.

Petersen, M.A., Fialkowski, D., 1994. Posted versus effective spreads. Good prices or bad quotes? J. Financ. Econ. 35 (3), 269–292.

R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Ranaldo, A., de Magistris, P.S., 2022. Liquidity in the global currency market. J. Financ. Econ. 146 (3), 859–883.

Ranaldo, A., Schaffner, P., Vasios, M., 2021. Regulatory effects on short-term interest rates. J. Financ. Econ. 141 (2), 750–770.

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

Schwert, M., 2017. Municipal bond liquidity and default risk. J. Finance 72 (4), 1683–1722.

Stoikov, S., 2018. The micro-price: A high-frequency estimator of future prices. Quant. Finance 18 (12), 1959–1966.