The Bias of Simple Bid-Ask Spread Estimators*

Pedro Tremacoldi-Rossi
Columbia University

Scott H. Irwin
University of Illinois

September 6, 2022

Abstract

We study why widely used low-frequency liquidity cost estimators based on high, low, and close prices perform well in some markets and poorly in others, often yield negative or indeterminate estimates, and how to quantify estimation bias empirically. Using the high-low spread estimator as our main setting, we show that the measure is biased due to two common bias factors. These bias sources contribute to performance loss significantly more than idiosyncratic factors, including model assumption violations and specific market microstructure characteristics, and are common to different spread estimators. Estimation bias increases in liquid assets or when price volatility is high. This relationship implies that evaluation studies in US equities inflate performance by almost 25%, as a few illiquid stocks disproportionately drive cross-sectional correlations. We then develop a theory-consistent method to calculate estimation bias bounds in any empirical context. Our bias bounds provide greater credibility to the use of the high-low estimator and address several practical issues introduced by its misbehavior.

Keywords: liquidity costs, bid-ask spreads, estimation bias
JEL Classification: C10, G12, G13.

*We are grateful to Torben Andersen, Dan Bernhardt, Adam Clark-Joseph, Phil Garcia, Sida Li, Tatiana Mocanu, Esen Onur, Michel Robe, Teresa Serra, Quanbiao Shang, Simon Schmickler, and numerous seminar participants for comments. Any errors are our own. Corresponding author: pt2614@columbia.edu.

1 Introduction

Following Roll (1984), the use of bid-ask spread estimators has become widely popular in empirical finance. Often times, such estimators are the simplest, or possibly the only way, to measure liquidity costs in emerging and voice-based markets (Li et al. (2018)), during historic periods (Patton and Weller (2020), Bernstein et al. (2019)), or when researchers predominantly use low-frequency data (e.g., McLean and Pontiff (2016)), as is common in the asset pricing literature. Despite the usefulness and popularity of low-frequency spread estimators, results of performance evaluation in different markets reveal puzzling results.

Consider for instance one of the best-known spread measures, the high-low estimator from Corwin and Schultz (2012). While it correlates well with effective spread levels in US stocks (Abdi and Ranaldo (2017)) and bonds (Chakravarty and Sarkar (2003)), it only marginally captures spread changes in FX markets (Karnaukh et al. (2015)) and ETFs (Marshall et al. (2018)), and has no correlation with commodity trading costs (Marshall et al. (2011)). Even when the measure remains superior to alternatives in global stock markets (Fong et al. (2017)), its proxy quality varies substantially across countries. Other simple spread estimators show the same variability in performance.

A second set of issues arises regardless of the application context. Estimators often yield spreads that are negative or implausible. Researchers are required to implement a number of *ad hoc* steps or arbitrary imputations, such as replacing negative spreads with zeros. Discrepancies in performance and problematic estimates are usually attributed to differences in microstructure across markets, which at first seems a reasonable explanation. Model assumptions may hold in certain settings and be violated in others. However, the same estimation shortcomings persist even in simulated trading data with ideal characteristics or in very large samples.[1]

What explains differences in performance and long-standing estimation issues of bid-ask spread measures and can these issues be addressed empirically? This paper derives a theoretical framework to determine the causes and implications of misbehavior in spread estimators, how to empirically measure estimation bias, and how to increase the degree of confidence in out-of-sample use of spread benchmarks. We focus our analysis on one of the most widely used spread measures, the high-low estimator of Corwin and Schultz (2012) and show how our conclusions apply to other spread estimators based on daily price data.

[1] An additional hypothesis that may account for performance issues is the changing nature of liquidity provision in modern financial markets. For example, Barardehi et al. (2018) argue that temporal dependence of orders submitted by low-latency traders make calendar-fixed liquidity proxies fail to capture relevant trading costs in algorithmic markets. Even traditional bid-ask spread estimators may capture broader liquidity dimensions than just the spread, including price pressure. Despite that, Easley et al. (2020) find that traditional microstructure measures, including price-based liquidity proxies, remain useful in a wide number of markets.

We begin our analysis by revisiting the empirical setting used in the literature to evaluate the performance of bid-ask spread estimators. These proxies are tested in standard US stock data, where benchmark effective spreads are calculated from high-frequency trade and quote data and compared to estimated spreads. Performance is generally measured as how well an estimator correlates *on average* with effective spreads in the cross-section and time-series. Considerable skewness in the distribution of liquidity in US stocks, however, plays a crucial role in inflating the measured performance of spread estimators.

We show that cross-sectional correlations between the high-low estimator and effective spreads in US stocks are heavily driven by a few very illiquid stocks whose removal from the cross-section decreases the average monthly correlation by 25% (from 0.62 to 0.48). These patterns are robust even in annual spread averages, in spread levels or changes, with or without *ad hoc* adjustments, and in other estimators such as Roll (1984) and Abdi and Ranaldo (2017).[2]

While some previous results in papers deriving spread estimators have suggested the existence of some relationship between the underlying effective spread size and proxy performance, the extent to which reported performance depends on the shape of the underlying distribution of trading costs in stocks is surprising. This persistent relationship motivates us to develop a framework that focuses on common bias sources — irrespective of application context — as opposed to idiosyncratic market features difficult to pinpoint and unlikely to account for puzzling performance results across several distinct settings.

We first derive a closed-form solution for the estimation bias in the high-low spread estimator, which comprises two terms: a moment bias and a small sample bias. Spread proxies estimate the bid-ask spread from a combination of daily high, low, and close transaction prices, which contain microstructure noise. To separate the spread from latent price information and achieve a closed-form solution, estimators are generally derived under the validity of asymptotic relationships, resulting in nonlinear functional forms of population moments. Additional assumptions usually impose restrictions that are plausible only in short data intervals (e.g., stationarity or constant spreads), so that in practice population moments are replaced with sample analogues calculated with a small number of data points.

Especially important, implementation of the high-low estimator requires replacing moments of the range with transaction-based price ranges over a two-day period. Crucially, our framework shows how this introduces a moment bias in the estimator due to its nonlinear functional form. We further establish that this bias is equivalent to the well-known Jensen's inequality bias in the Roll measure (Roll (1984), Hasbrouck (2004)), which is also biased because of the use of the sample estimator of the autocovariance. The size of this moment bias

[2]This disproportionate impact of low liquidity stocks in driving cross-sectional results is similar to the well-known effect of microcaps (more likely to be illiquid) in validating return anomalies (Hou et al. (2018)).

in the high-low estimator can be expressed as a function of observable data and thus retrieved empirically. Nonetheless, even with a bias correction, the estimator still deviates from the underlying spread due to a small sample bias, which stems from an empirical convergence failure assumed in the estimator's derivation.

The high-low spread proxy relies on the validity of the asymptotic distribution of the range, which is rarely observed in practice or even in large samples of simulated ideal prices. We show that, as a consequence, high-low estimates remain biased in sample sizes typical for financial data. This bias source is equivalent to the small sample bias in the Roll measure. In that case, the sample covariance estimator is biased in small samples (Harris (1990)), which introduces further bias in the proxy. These two common sources combine to drive a bias in the high-low measure of same nature as in the Roll measure, which relies on different transaction variables for identification and has a distinct set of assumptions.

Next, we explain the determinants of the two sources of bias. By writing the total bias as a sum of both moment and small sample terms, we prove that the high-low estimation bias is decreasing in the spread and increasing in volatility. These two latent variables in the data generating process affect the estimation bias in the high-low measure the same way they do with the Roll estimator, whose performance also worsens as the signal-to-noise ratio falls. This theoretical association rationalizes the previous patterns found between underlying effective spread sizes and estimation bias in the US stock sample. To eliminate potential idiosyncratic factors, we simulate ideal trading data that cleanly confirms the same relationship.

The fact that estimation bias depends on the underlying bid-ask spread magnitude has important implications for empirical applications using spread estimators to proxy for liquidity costs. For instance, in low-frequency cross-sectional regressions where researchers include spread estimates as independent variables, coefficient estimates will suffer from non-classical measurement error. In cases where the actual measured spread value matters, for example, to sort a portfolio on liquidity or to exploit changes in liquidity costs over time, researchers need to know how much trust to place on a spread estimate of 1% provided by a low-frequency estimator. Is this a good proxy of an underlying spread of 1% or is the actual effective spread 0.2% or 2%?

To equip researchers using the high-low estimator to address these issues, we develop an approach to obtain empirical bounds on the estimation bias. Exploiting the fact that the moment bias is always non-negative and can be computed from daily data, we show how to sign the small sample bias, allowing us to bound the effective spread using only the high-low estimate and moment bias. These bounds are sufficiently tight so that they are informative in a variety of settings — to both quantify the bias and control for non-classical measurement error in linear regressions, in the cross-section and time-series. The bounds on the underlying spread can also be used to sort portfolios based on liquidity much more precisely than simply relying on the high-low measure. Moreover, the bias bounds are derived under the same assumptions as the high-low estimator and despite of potential idiosyncratic biases remaining, are empirically-consistent in the vast majority of cases, as we show in three different applications using US stock data.

It is important to emphasize that the causes and implications of misbehavior we discuss are not exclusive to the high-low measure. We focus on the proxy because of its popularity, perceived dominance relative to other estimators, and innovative identification strategy of the spread. But the lessons we draw about its shortcomings are more general. The same factors resulting in poor performance of the measure will also negatively affect other spread estimators, such as the Roll measure and the close-high-low estimator from Abdi and Ranaldo (2017), as we show in the Online Appendix C. Indeed, even with several modifications and overall superior performance compared to the Roll measure, the high-low estimator suffers from the same sources of misbehavior as the seminal estimator.

The main contribution we make is to establish the importance of structural factors common to simple bid-ask spread estimators in affecting performance and how to empirically quantify estimation bias in the high-low measure. Our estimation bias bounds should prove useful for researchers in market microstructure, asset pricing, or that require direct low-frequency measures of trading liquidity costs.

Recent work by Lou and Shu (2017) and Jahan-Parvar and Zikes (2019) challenges the pricing of liquidity measures in certain market circumstances. Our paper takes a more fundamental approach by analyzing the structure and source of bias in spread proxies, which informs researchers whether they are likely to face implementation issues in their own setting. Rather than suggesting the abandonment of simple spread estimators altogether, we recognize that these are crucial tools to measure liquidity in many circumstances, and provide researchers with tools to increase the credibility of their results.

Our results demonstrate that even though both volatility and the underlying spread size affect the estimator's behavior, spread size is more important in generating and accounting for estimation bias. We show that volatility only begins to affect estimation bias when its magnitude is at least four times that of the underlying spread. Although this ratio is likely met in modern equity markets, which have a median effective spread of 0.3%, centering the analytical response of the bias to the spread has a crucial advantage. By exploring the relationship between spread size (the quantity the researcher aims to estimate) and estimator bias, we are able to jointly assess the empirical size of bias and the underlying magnitude of the effective spread.

Our initial set of findings microfounds misbehavior in the popular high-low spread measure in the same fashion as the work in Harris (1990) for the Roll estimator. By focusing on common factors inducing estimation bias and negative or indeterminate estimates and how these factors react to changes in volatility and, in particular, the underlying true spread, we are able to establish a direct link between misbehavior in the Roll and high-low measures, and apply our framework to another simple bid-ask spread estimator, the close-high-low proxy. This contrasts our work from previous studies that relied on idiosyncratic factors, like Bleaney and Li (2015) and Nieto (2018), or in mechanical noise from truncating averages with zeros (Jahan-Parvar and Zikes (2019)).[3]

We argue that both the moment and small sample bias sources account for almost the entirety of empirical misbehavior in the high-low measure. To test that, we derive a counterfactual version of the high-low measure which completely corrects for all common bias sources. This corrected estimator version performs extremely well, correlating with the effective spread by almost 95% in the cross-section and 80% in the time-series. A back-of-the-envelope calculation implies that, in the cross-section, common factors contribute to a loss of over 30 percentage points in performance, while all other imperfections combined imply a loss of performance of only 5 percentage points.

We also draw several recommendations from our analysis for the development of new estimators and performance evaluation. To more credibly establish the quality and recommend a certain spread proxy, performance evaluation studies using tick data should include more detailed subsampling analysis, particularly by analyzing the relationship of correlation coefficients in the cross-section across different effective spread groups. Moreover, we also suggest in sensitivity analyses that performance is measured after dropping the most illiquid units from the sample as well as in random subsamples from the data when the cross-section is sufficiently large. Coefficient stability across these different tests indicate a greater degree of out-of-sample confidence.[4]

Going forward, newly developed spread estimators using transaction data that share the structure of the Roll framework should be subjected to rigorously designed simulation environments with true spreads in the data generating process well below 0.5%, the usual minimum spread used in the literature, and verify the role of price volatility and trading costs in potential misbehavior. More preferably, a formal analysis of the existence of the moment or small sample bias in new measures is also suggested, as these two sources can severely limit performance of simple bid-ask spread estimators. Finally, even if after rigorous testing new estimators perform

[3]We refer to these bid-ask spread estimators as "simple" following the designation from Roll (1984), Corwin and Schultz (2012), Abdi and Ranaldo (2017).

[4]In addition to a more robust and extensive empirical performance evaluation, the choice of the underlying spread benchmark is important, as effective bid-ask spreads might be inappropriate (Hagströmer (2021)).

better than existing proxies, they should still accompany practical tools to quantify empirical bias and reduce its importance in applications, as we do for the high-low measure.

This paper is organized as follows. Section 2 introduces the structure of simple bid-ask spread estimators, underscoring deriving assumptions and potential idiosyncratic factors important for empirical performance. Section 3 revisits the standard performance evaluation setting of spread estimators, analyzing the robustness of previous findings. Section 4 introduces our framework, where we study the determinants and implications of estimation bias in the high-low estimator. Section 5 implements our diagnostic test and shows how to empirically assess estimation bias and spread size. Section 6 concludes.

2 The Structure of Simple Spread Estimators

This section introduces the framework common to most simple bid-ask spread estimators, focusing on the derivations and assumptions of the high-low measure by Corwin and Schultz (2012). The building blocks and identification approach of the framework were introduced by Roll (1984), especially basing the estimation of the spread on daily transaction prices. Many estimators introduced later followed or extended the Roll measure (French and Roll (1986), Thompson and Waller (1987), Lesmond et al. (1999), Hasbrouck (2004), and Fong et al. (2017)). We underscore the two sources of bias in the Roll measure previously shown in the literature. Later in the paper, we will show that the same bias sources also affect the high-low proxy, and more generally, how simple bid-ask spread estimators employing different technologies have performance issues based on the same biases.

2.1 Setting

The derivation of simple bid-ask spread estimators begins with the usual assumption that log prices follow a geometric Brownian motion (GBM). Hence, continuous existence of prices during trading hours is implicitly assumed, though continuous observation thereof is not necessary. The relationship between observed (p_t) and true prices (P_t) is given by:

$$p_t = P_t + \frac{S}{2} Q_t \tag{1}$$

where S is the total effective spread (observed) and Q_t is the order flow indicator ($Q_t = 1$ for buyer-initiated orders and $Q_t = -1$ for seller-initiated orders). Given that the latent price evolves as a random walk, $\Delta P_t = \varepsilon_t$, with $\Delta P_t = P_t - P_{t-1}$, $\varepsilon_t \sim \text{NID}\left(0, \sigma^2\right)$, trade direction is assumed to be independent of efficient prices. We refer to the true spread or the effective spread calculated from trading data interchangeably by S.[5]

In spite of its simplicity, (1) represents the cornerstone expression shared by all spread proxies we characterize as *simple estimators*. Within this category, time is indexed by $n(t)$, $t = 1, ..., T$, days, so that only end-of-day low frequency data is necessary to recover empirical estimates of the true spread S. Daily spread estimates may be averaged at the desired frequency, say months or years, in order to compute average spread estimates.

2.2 The Roll Estimator

Roll (1984) considers daily closing prices, $c_t = C_t + Q_t S/2$, as the relevant price realization for empirical use. By assuming that the distribution of price changes in (1) is plausibly stationary *in short time periods*, he arrives at the well-known Roll estimator formula:

$$S^{Roll} = 2\sqrt{-\mathrm{Cov}\left(\Delta p_t, \Delta p_{t+1}\right)} \tag{2}$$

2.2.1 The two bias sources in the Roll estimator

In practice, (2) is implemented by replacing the population autocovariance of price changes with its sample version, effectively leading to the estimator:

$$\widehat{S}^{Roll} = 2\sqrt{-\widehat{\mathrm{Cov}}\left(\Delta p_t, \Delta p_{t+1}\right)}$$

While this replacement is seemingly innocuous, the Roll measure suffers from a Jensen's inequality problem, due to its nonlinear functional form. Average spread estimates are downward-biased with (absolute) bias size decreasing in the true spread level S, increasing in price volatility σ^2, and increasing in the sample size $n(t)$: $\mathbb{E}\left[\widehat{S}^{Roll}\right] - S = -S\left(\dfrac{\mu_4/\sigma^4 + 4}{8\left(n(t) - 1\right)}\right)$.

To complicate things further, Harris (1990) shows that another source of bias affects the behavior of the Roll measure. The sample covariance statistic used to calculate $\widehat{S}^{Roll}$ is biased in small samples, so that even if the Roll measure were a linear function of the sample covariance, so that the Jensen's inequality bias would vanish, the estimator will remain biased in small samples. Unfortunately, "small sample" in this context covers the temporal dimension of financial data usually available, making the issue pervasive. Considering both bias sources in total, the Roll measure bias decreases in the spread and increases in volatility (overestimating

[5]The most common cost-based high-frequency liquidity benchmarks used empirically are the effective and quoted spreads. Effective spreads measure "immediate" execution, i.e., faced traders executing at observed prices.

S in liquid markets). These two bias sources are well-known, but strongly associated with the functional form in (2). We refer to them as a *moment bias* and a *small sample bias*. Finally, a separate issue with the Roll estimator is the presence of indeterminate daily spread estimates due to frequent positive price autocovariance.[6]

The structural link between spread size, bias, and frequency of indeterminate spreads in the Roll estimator has been relatively unexplored and often overlooked in the context of other spread estimators. Newer proxies have been shown to more accurately capture effective spreads in simulated and empirical environments relative to the Roll measure, with little attention being paid to estimation bias or other systematic issues.

2.3 The High-Low Estimator

Corwin and Schultz (2012) exploit the empirical regularity that the sign of the trade indicator Q_t is defined as positive in daily high prices and negative in daily low prices. Thus, under the assumption that high and low prices are buyer- and seller-initiated, respectively, (1) translates as the pair:

$$(h_t, l_t) = \left(H_t + \frac{S}{2}, L_t - \frac{S}{2} \right) \tag{3}$$

and more specifically, the daily log range $r_t \equiv h_t - l_t$ can be manipulated by using the moments derived in Parkinson (1980) and Garman and Klass (1980) for the true range R:

$$\mathbb{E}\left[T^{-1} \sum_{t=1}^{T} R_t \right] = k_1 \sigma \text{ and } \mathbb{E}\left[T^{-1} \sum_{t=1}^{T} R_t^2 \right] = k_2 \sigma^2 \tag{4}$$

where $k_1 \equiv \sqrt{8/\pi}$ and $k_2 \equiv 4 \ln 2$, and σ^2 is an unbiased (under the assumption of no drift in the GBM) estimator of daily variance based on the range when $\mathbb{E}\left[T^{-1} \sum_{t=1}^{T} R_t \right]$ and $\mathbb{E}\left[T^{-1} \sum_{t=1}^{T} R_t^2 \right]$ are replaced with R_t and R_t^2, respectively.

The appealing argument that using additional information contained in the daily range improves volatility estimation underscores the innovative path taken by Corwin and Schultz (2012). Since the range represents the boundaries of observed daily price oscillation, i.e., the volatility measured in σ^2, it also includes the effective spread at those boundary points accord-

[6]Negative daily serial covariance is necessary to compute (1) or its implementable version. Harris (1990) argues that small sample bias alone explains the prevalence of positive covariance estimates. Nearly half of daily spread estimates turn out negative, requiring researchers to discard substantial information. Importantly, because the small sample bias in the sample covariance decreases in the true spread size, more liquid markets frequently induce positive autovariance, ultimately generating indeterminate spread estimates more often.

ing to (3). The additional assumption that the spread is constant over every pair of consecutive days, $\bigcup_{k=1}^{n(t)-1}[t_k, t_{k+1}]$, while volatility remains proportional to time, allows for uniquely determining S from the following expressions:

$$\mathbb{E}\left[\sum_{t=1}^{2} r_t^2\right] = (8\ln 2)\,\sigma^2 + \left(8\sqrt{\frac{2}{\pi}}\right)\sigma\widehat{S} + 2\widehat{S}^2 \tag{5}$$

$$\mathbb{E}\left[r_t^{*2}\right] = (8\ln 2)\,\sigma^2 + \left(8\sqrt{\frac{1}{\pi}}\right)\sigma\widehat{S} + \widehat{S}^2 \tag{6}$$

which relate the daily range (r_t) and the two-day range (r_t^*), where $r_t^* \equiv \max\{h_t, h_{t+1}\} - \min\{l_t, l_{t+1}\}$, to both the spread and daily volatility. The two-day range captures the overall volatility in each pair of days, and if time proportionality holds, it should correspond to twice the volatility of a single day.

After solving for σ^2 in (5) and (6), the high-low spread estimator is obtained as:

$$S^{HL} = 2\left(L\left(\alpha\right) - 1\right) \tag{7}$$

for

$$\alpha \equiv \frac{\sqrt{2\beta} - \sqrt{\beta}}{3 - 2\sqrt{2}} - \sqrt{\frac{\gamma}{3 - 2\sqrt{2}}}, \quad \beta \equiv \sum_{t=1}^{2} \mathbb{E}\left[r_t^2\right], \quad \gamma \equiv r_t^{*2} \tag{8}$$

where $L(\cdot)$ is the logistic function. We transform the functional forms for S and α from Corwin and Schultz (2012) for ease of exposition using the steps in Appendix (II).

2.3.1 Specific model assumptions

Before moving on to the re-evaluation of the empirical performance of the high-low measure, we underscore important identifying assumptions of the estimator. First, daily high prices are buyer-initiated and low prices are seller-initiated. Second, the price process is continuous and that volatility is proportional to the trading horizon. Third, the asymptotic distribution of the true range holds. Fourth, the latent spread is constant over the trading horizon.

Abdi and Ranaldo (2017) check that the first assumption is verified in 95% of US stock-days, therefore it is a stylized fact in practice. The second assumption has devoted most of the attention from the analysis in Corwin and Schultz (2012) on model assumption failures, particularly time-varying volatility. We show later that, combined with time-varying spreads, these assumptions play a small role in affecting the empirical performance of the estimator when compared to the bias sources we highlight.

3 Reevaluating Simple Bid-Ask Spread Estimators

This section revisits the standard evaluation environment of bid-ask spread estimators by investigating the robustness of previous findings in the cross-section and time-series of US stocks. Our analysis will show that a skewness effect in US stocks — many stocks with small spreads and few that are extremely illiquid — produces cross-sectional correlations with a low degree of out-of-sample confidence. Removing few large effective spread stocks from the sample used in horse-race exercises decreases the measured performance of the high-low, close-high-low, and Roll estimators by almost 25%. This lack of robustness associated with the underlying level of effective spreads will motivate our theoretical framework developed later in the paper.

3.1 Benchmark Stock Data

Our baseline empirical setting uses the standard US stock sample employed in horse-race type papers like Goyenko et al. (2009). This is the sample used to benchmark the performance of simple bid-ask spread estimators (e.g., Corwin and Schultz (2012) and Abdi and Ranaldo (2017)). We use Daily TAQ (DTAQ) stock data from 2009 through 2013 with stocks listed at NYSE, AMEX, and NASDAQ, applying the Holden and Jacobsen (2014) procedure to compute daily effective spreads. We describe these steps in more detail in Appendix (I).[7] Estimation of low-frequency proxies uses CRSP daily data, again closely tracking the setting in previous studies. We also apply the data cleaning steps in Corwin and Schultz (2012) for anomalous trading days, though the effects of carrying out these changes on the overall analysis are minimal.

Figure (1) plots the distribution of daily effective spreads across all US stocks. The median effective spread in the sample is 0.21% and more than 70% of stocks have effective spreads lower than 1%. The long right tail shows that only about 250 stocks have spreads higher than 3%, less than 10% of all stocks. This prominent skewness of effective spreads in US stocks, therefore, is a stylized fact in the data.

[7]We thank Craig Holden and Stacey Jacobsen for making their code publicly available. We also thank Shane Corwin, Paul Schultz, Farshid Abdi, and Angelo Ranaldo for making their code publicly available for comparison.

3.2 Reassessing the Performance of Simple Spread Estimators

3.2.1 Notation

The performance of low-frequency bid-ask spread estimators is usually evaluated by comparing monthly averages of effective spreads and the estimated proxy in the cross-section and time-series. Formally, the monthly average of high-low estimates is given by:

$$\widehat{S}_n^{HL} \equiv \frac{\sum_{t(n)} \widehat{S}_t^{HL} \cdot \mathbb{1}\{adjustment\}}{\sum t\,(n)} \tag{9}$$

where daily HL estimates are averaged over t trading days available in month n. We denote the high-low estimator used empirically as $\widehat{S}^{HL}$.

Daily spread estimates may be discarded or imputed based on *ad hoc* adjustments used in the literature: i) a zero adjustment, which replaces negative daily estimates with zeros, ii) using only positive spreads, or iii) using all raw estimates without any correction. We refer to monthly average versions of the high-low measure as follows: *Raw* for using both negative and positive estimates (that is, no adjustment); *Zeros*, for the replacement of negative estimates with zero ($\mathbb{1}\{\widehat{S}_t^{HL} < 0\} = 0$); and *Positives* when discarding negative or indeterminate estimates. We also calculate annual averages of spread estimates, which are then compared to annual effective spread averages, and replicate some of our analysis using *changes* in monthly and annual estimates. Computed sample averages of the Roll and close-high-low estimators follow the same approach in (9).

3.2.2 Analysis

With the previous cross-sectional effective spread patterns just illustrated, we now dissect the performance of simple bid-ask spread estimators. Figure (2) begins by calculating the average cross-sectional correlation between monthly high-low spread estimates and effective spreads. The HL measure performs relatively well, tracking the cross-section of over 3,300 stocks with an average correlation of 62%.[8]

Next, we break down the stock sample according to effective spread deciles and recalculate the cross-sectional correlation coefficients within each liquidity level decile. Given the heterogeneity in effective spreads in US stocks, particularly the right skewness, this analysis verifies whether performance might be driven by a particular subset of stocks. We display results for each decile for three computed versions of high-low estimates: using the zero ad-

[8]The correlation point-estimate is identical to the value in Table 5 in Abdi and Ranaldo (2017), who use the same stock sample and similar period. Estimates for the Roll and close-high-low proxies are also identical.

justment (which is the recommended *ad hoc* modification by Corwin and Schultz (2012), in blue in the figure), only positive daily estimates (in gray), and without any adjustments (raw, in brown). Although these other versions of the high-low measure are not recommended by Corwin and Schultz (2012), we include them in our analysis to show that the patterns we discuss next are not an artifact of the *ad hoc* zero adjustment.[9]

The distribution of performance across effective spread levels reveals important patterns masked by an evaluation of the full cross-section. First, the HL measure (zero adjustment or positive estimates) only performs modestly in the top effective spread decile, which has an average monthly spread of 4.3%, 20 times larger than the median effective spread for overall US stocks. Even in this most favorable case, both adjustments only yield average correlations below 50%. More importantly, dropping stocks with the largest underlying effective spread in the top decile yields a cross-sectional correlation across all remaining stocks (deciles 1 to 9) of 48%, 14 percentage points lower than the sample-wide performance. To put this magnitude of performance loss in perspective, 14 percentage points is larger than the measured improvement in performance from the close-high-low estimator over the high-low measure in Abdi and Ranaldo (2017).[10]

The same "skewness effect" impacts the Roll and close-high-low estimators, both for monthly and annual analyses (Online Appendix Figures (D.1) and (D.2)). The overall loss in performance from removing the top stock decile from the cross-section is of similar magnitude to the monthly high-low estimator, about 25%. Online Appendix Figures (D.3), (D.4), and (D.5) indicate that the cross-sectional correlation between changes of monthly and annual high-low, close-high-low, and Roll spread estimates and effective spreads are also inflated by a few illiquid stocks in the cross-section. We underscore this specific context as previous papers have found that the high-low measure may track well only changes in liquidity in certain markets, and not necessarily levels as in Figure (2).

Finally, a simple "placebo" test of the potential relationship between estimation performance and the underlying effective spread size is to drop the decile with the *most* liquid stocks. The idea here is to cleanly separate the effect of decreasing the sample size (in the cross-section) from dropping *certain* stocks from the sample (those few stocks with very large

[9]Note that imputing zeros or using only positive spread estimates introduces bias in sample averages of any parameter estimate. In the context of spread estimation, this mechanical (or sampling) bias from truncation or censoring trivially inflates the average estimated spread, especially relative to smaller true spreads and when the price variance is large. While the moment bias and the small sample bias are structural biases affecting *point* estimates of the Roll measure, zero imputation or other *ad hoc* empirical adjustments introduce an additional bias source. Within our framework in Section 4, the potential bias introduced from the zero adjustment is absorbed as part of an idiosyncratic bias.

[10]The same general pattern is observed in the performance of annual average spreads, also in Figure (2). Using longer time periods improves the relationship between the high-low measure and effective spreads in the cross-section, but a large contributor to performance is again the top decile of trading costs.

spreads). Across monthly and annual averages of levels and changes and for all estimators, removing the first stock decile leaves the performance in the remaining sample unchanged relative to the full cross-section.[11]

3.2.3 Implications

The fact that the standard benchmark evaluation of simple bid-ask spread estimators overstates performance by almost 25% has important empirical and theoretical implications. On the empirical side, the widespread use of low-frequency spread estimators in markets outside US equities relies entirely on out-of-sample extrapolation — if an estimator "robustly" benchmarks effective spreads in US stocks, so it should in markets like futures, foreign exchange, bonds, and ETFs. While it is reasonable to expect that performance may not replicate due to idiosyncratic factors in each application context, the role of skewness in effective spreads in inflating performance raises concerns to applications in *any* market.[12]

On the theoretical side, the fact that three estimators using different technologies to recover the spread all have a much worse performance once few illiquid stocks are removed from the cross-section points to a common factor of misbehavior — possibly associated with the underlying latent spread. This conjecture motivates us to develop a framework focusing on structural forces determining the estimation bias of the high-low measure and how the true spread and volatility influence these biases.

4 Estimation Bias & Performance

This section provides a theoretical framework to evaluate the determinants of the structure and behavior of the high-low measure and its empirical consequences. We first show that replacing the population moment in the estimator's parameter β with daily ranges introduces a moment bias, which makes the implementable version of the high-low estimator biased com-

[11] Apart from the inconsistent performance across different markets we previously mentioned and discuss in detail later, existing analysis suggests some degree of systematic variation in the high-low proxy quality even for US stock data. Abdi and Ranaldo (2017) show in their Table 5 a breakdown of correlation quality by effective spread quintiles, suggesting a much lower performance of considered spread proxies in more liquid stock groups. This holds for both the high-low estimator and their measure, the close-high-low spread proxy, as well as the Roll measure. A similar point is inferred from Table 6 in Corwin and Schultz (2012), where the cross-sectional correlation between the effective spread and the high-low measure decreases from 70% in low cap stocks to 18% in high market-capitalization firms, which are more liquid. While suggestive, there is no indication in previous papers that these subsample patterns are robust or even theoretically founded.

[12] Further analysis in the Online Appendix A demonstrates this point further by showing the high sensitivity of findings in horse-race type analysis in US stocks to changes in the sample size and other simple tests.

pared to its theoretical counterpart. This moment bias is structurally equivalent to the Jensen's inequality bias in the Roll estimator.

More importantly, even if we remedy this moment bias, the estimator is still biased in practice because of its estimation horizon. The high-low measure relies on the asymptotic distribution of the range, which fails to hold for consecutive trading days. More generally, the relationship required for unbiased estimates fails to hold even over longer trading horizons, making the spread estimator biased irrespective of the moment bias. This small sample bias in the high-low measure corresponds to the small sample bias in the Roll estimator, where the sample covariance estimator introduces additional bias in the proxy. These two common biases almost completely accounts for performance loss in high-low estimator.

With a closed-form expression for the bias, we further show that the high-low bias increases in small spreads and high volatility, which explains our previous findings on poorer performance in more liquid stocks. We confirm those patterns in generated data that simulates an ideal setting for all other model assumptions. In the same simulated environment, we also find that the frequency of negative estimates increases when we shock the spread towards zero. Online Appendix C applies this framework to the close-high-low measure of Abdi and Ranaldo (2017), which yields by and large the same conclusions with respect to estimation bias size, latent spread and volatility, and negative estimates.

Finally, we conduct a counterfactual exercise to quantify the contribution of the two common bias sources — moment and small sample — to the overall performance of the high-low measure. This exercise isolates the loss in performance attributed to the two common bias sources from *all* other idiosyncratic bias source, including US stock microstructure effects and model assumption failures (e.g., constant spreads). Counterfactually removing the two common biases improves empirical performance by over 30 percentage points (62% to 94%) in the cross-section and in the time-series (48% to 80%), jointly accounting for 90% of total performance loss in the high-low estimator.

4.1 Framework

Start by rewriting the high-low estimator formula in Equation (7) as

$$\widetilde{S}^{HL} = \alpha$$

after linearizing the expression $2\frac{\exp\alpha - 1}{\exp\alpha + 1}$ around $\alpha_o = 0$. This approximation generates negligible error for spread values between -100% and 100%. Going forward, we denote

$\widetilde{S}^{HL}$ as the theory-implied high-low estimator, which uses the moment-based parameter $\beta = \mathbb{E}\left[\sum_{j=0}^{1} r_{t+j}^2\right]$.

Next, the following equivalence allows us to solve for the parameter α in the difference of square roots involving β and γ:

$$\widetilde{S}^{HL} = \underbrace{\frac{\sqrt{2\beta}-\sqrt{\beta}}{3-2\sqrt{2}} - \sqrt{\frac{\gamma}{3-2\sqrt{2}}}}_{\alpha} \equiv \left(1+\sqrt{2}\right)\left(\sqrt{\beta}-\sqrt{\gamma}\right) \tag{10}$$

We denote the expression above as the **theoretical high-low estimator**. The functional form in (10) provides a direct way to express how the high-low spread estimator is jointly — and fully — determined by the latent spread and price volatility

$$\widetilde{S}^{HL} = S + \underbrace{\left(1+\sqrt{2}\right)\left(\sqrt{2}\mathbb{E}\left[R_t\right] - R_t^*\right)}_{g(\sigma)} \tag{11}$$

The high-low estimator is therefore additively separable in (S, σ), where the function $g(\sigma) \in \mathbb{R}$ drives the sign of the daily spread estimate. We return to the determinants of negative spreads in the Online Appendix B. Finally, though we show it formally below, it is clear from (11) that unless g is zero, the theoretical estimator is biased. In practice, $\widetilde{S}^{HL}$ is estimated using a sample-based version of β, defined as $\widehat{\beta} \equiv r_t^2 + r_{t+1}^2$. This replacement results in a different version of the functional form in (11). First, we rewrite $\widehat{\beta}$ as[13]

$$\sqrt{\widehat{\beta}} = \sqrt{r_t^2 + r_{t+1}^2} = \left(R_t^{min} + S\right)\left(\sqrt{1+\kappa^2}\right)$$

for $\kappa \equiv \dfrac{\max\{r_t, r_{t+1}\}}{\min\{r_t, r_{t+1}\}} \equiv \dfrac{r_t^{max}}{r_t^{min}}$. Then, the **implementable high-low estimator** is

$$\widehat{S}^{HL} = \omega S + \left(1+\sqrt{2}\right)\left(\phi R_t^{min} - R_t^*\right) \tag{12}$$

where $\phi \equiv \sqrt{1+\kappa^2}$ and $\omega \equiv \left(\phi + \sqrt{2}\phi - 1 - \sqrt{2}\right)$. In practical terms, the implementable version of the estimator scales the underlying true spread by a factor influenced by the variation between daily ranges, which both determine ω and ϕ. The linear functions for the theoretical and implementable high-low spread estimators above make the computation of the high-low estimator bias straightforward.

[13] This follows from the trigonometric relation $\sqrt{(\min\{r_t, r_{t+1}\})^2 + (\kappa \min\{r_t, r_{t+1}\})^2} = \min\{r_t, r_{t+1}\}\sqrt{1+\kappa^2}$.

4.2 Bias and Its Sources

The first important result we establish in this section leverages the convenience of the functional forms we derive in (11) and (12), to compute the estimation bias of the theory-based HL measure and from the actual bias from its empirical implementation. Proposition 1 shows the unbiasedness of the theoretical estimator and then demonstrates that the implementable high-low measure can only be unbiased under the unlikely empirical case of identical consecutive observed ranges. In all plausible application cases, it will necessarily suffer from estimation bias.

PROPOSITION 1. The theoretical high-low spread estimator $\widetilde{S}^{HL}$ is an unbiased measure of the true spread S. The implementable estimator $\widehat{S}^{HL}$ is unbiased only if $r_t = r_{t+1}$.

Proof. Appendix (II). ■

4.2.1 Unbiasedness requires identical daily price ranges

Proposition 1 establishes a crucial result: while the theoretical high-low spread estimator is an unbiased measure of the true spread, the implementable version of the estimator almost surely results in biased daily spread estimates, except for days with exactly identical observed ranges. Bringing this result to light is a direct consequence of our rewriting of the high-low estimator's functional form in (12). This additively separable form further shows that computing sample averages of $\widehat{\beta}$ and γ and then computing high-low estimates from (10) does not eliminate the bias. Importantly, note that the bias in Proposition 1 is of the same nature as the Jensen's inequality bias in the Roll measure. We refer to this source of bias in the context of both estimators as moment bias to stress the issue at hand: the use of a sample statistic or daily data points like in the high-low and close-high-low estimators, instead of the theory-derived moment.

4.2.2 Small sample validity of the asymptotic range distribution

To the extent that all other model assumptions hold, the size of the bias from the empirical use of the high-low estimator is completely determined by some function of the distance between r_t and r_{t+1}. The greater the variation between ranges over a two-day interval, the greater the moment bias. In practice, however, even if daily ranges *were equal*, resulting in an identity between the theoretical and implementable versions of the high-low estimator, daily spreads would still remain biased empirically. That is because the unbiasedness established in

Proposition 1 relies on $T^{-1}\sum_{t=1}^{T} R_t^* = T^{-1}\sum_{t=1}^{T}\sqrt{2}R_t$, which rarely occurs during the estimation time interval of financial data. This small sample bias is again of the same nature as the small sample bias that the sample autocovariance suffers from in the Roll measure.

The computation of bid-ask spreads from short estimation windows, e.g., two consecutive days, is consistent with derivation assumptions only likely holding in short time-series. For instance, Roll (1984) requires the distribution of returns to be stationary, while Corwin and Schultz (2012) and Abdi and Ranaldo (2017) assume that the latent spread is constant.[14] Defending these assumptions becomes increasingly difficult for longer data intervals. However, moment-based estimators are derived using relationships that may not hold in small samples. Particularly, finding a closed-form solution for the high-low estimator (and the close-high-low estimator) relies on the asymptotic distribution of the range from Feller (1951).

The unbiasedness of the theoretical high-low measure rests on the average two-day range being approximately equal to the average range, scaled by $\sqrt{2}$. The same must hold for the implementable HL estimator (in addition to the requirement of constant ranges). To verify how often this condition is met even in an ideal setting, in Figure (3) we simulate 10,000 months of trading data with a constant volatility value of $\sigma = 0.03$ and compare daily values and monthly averages of $\sqrt{2}R_t$ and R_t^* from the data. The sample distributions of the two ranges are distinct almost everywhere, as even monthly averages of each quantity seldomly fall on the 45-degree line (in blue in the figure).

4.2.3 Bias size

Next, we bring the two bias sources together and decompose the total bias in the implementable high-low measure into closed-forms of the moment and small sample biases.

Proposition 2. The bias in the implementable high-low estimator ($\widehat{S}^{HL}$) can be decomposed as

$$\mathbb{E}\left[\widehat{S}^{HL}\right] - S = \underbrace{\left(1+\sqrt{2}\right)\left[\underbrace{r_t^{min}\left(\phi-\sqrt{2}\right)}_{Moment\ Bias} + \underbrace{\left(\sqrt{2}R_t^{min} - R_t^*\right)}_{Small\ Sample\ Bias}\right]}_{Bias}$$

Proof. Appendix (II). ■

[14] Of course, deriving each estimator requires a number of additional assumptions, such as time-constant volatility, or continuous observation of trades, as mentioned before.

Proposition 2 separates out the bias in the high-low spread estimator bias two components: the moment bias, which depends on observed variables, thereby being quantifiable empirically, and small sample bias, which depends on two latent variables. Consistent with our previous discussion, the *Moment Bias* term vanishes when consecutive observed ranges are identical ($\phi = \sqrt{2}$) and the *Small Sample Bias* term goes to zero under the asymptotic validity of the range distribution in the estimating interval ($R_t^* \longrightarrow \sqrt{2}R_t$).

4.3 Common Biases Account for 90% of Performance Loss

We now take the bias expression in Proposition 2 to the US stock data. Specifically, we investigate how much the small sample and moment biases account for the empirical performance of the high-low estimator. To answer that, we calculate a "best-performing" version of the high-low measure where we correct daily spread estimates $\widehat{S}^{HL}$ as:

$$\widehat{S^*}^{HL} = \widehat{S}^{HL} - Moment\ Bias - Small\ \widehat{Sample}\ Bias \tag{13}$$

in which we empirically retrieve the term $Small\ \widehat{Sample}\ Bias$ by computing "true" ranges by subtracting daily effective spreads from daily observed ranges. *Moment Bias* is computed directly from the data. The bias-corrected version of the high-low estimator above nets out both theoretically-derived bias sources. The difference in performance between the measure and the standard high-low estimator (which tracks stocks by 63% in the cross-section and 48% in the time-series) therefore captures the performance loss in the implementable high-low spread attributed to the moment and small sample biases.[15]

Figure (4) replicates Figure (2) using monthly averages of the bias-adjusted HL estimator and monthly effective spreads. Once again, we analyze cross-sectional correlations for the deciles of stocks sorted according to liquidity levels, plotting in the figure the within-subsample cross-sectional correlation between the measure and the effective spread as well as the average frequency of daily negative estimates. By removing common biases from high-low spread estimates, the adjusted proxy performs in the overall cross-section significantly better than the standard implementable estimator.

The overall cross-sectional correlation increases from about 63% to 94%. More strikingly, this substantial performance improvement occurs in all spread deciles whose correlation with

[15]In the empirical applications using non-simulated data, we consider the effective spread — the underlying variable the simple bid-ask spread estimators we study attempt to recover — to correspond to the latent or "true" spread. We refer to the difference between a spread estimate and the effective spread in such contexts as empirical bias to maintain the language concise. Later in the paper, we refer to this difference interchangeably as measurement error in the linear regression framework to stress the econometric issues associated with the estimator's bias in that case.

effective spreads increases anywhere from two-fold to almost four-fold. This generalized improvement in performance significantly reduces the influence of few illiquid stocks in determining the cross-sectional correlation, which drops only by 5% after removing the top stock decile, instead of 25% as in the exercise in the previous section.

Since our derived bias expression assumes all high-low model assumptions hold, the residual noise in the cross-sectional correlation (6%) captures any estimation bias solely attributable to the microstructure of the US stock sample — idiosyncratic bias. That includes, for example, direct empirical violations from assuming constant spreads within two consecutive days or constant volatility. Counterfactually, this implies that if we were able to make the high-low estimator perfectly accurate, 90% of the performance improvement necessary would be attributed to eliminating the two biases that are also common to the Roll measure.[16]

Conducting a similar counterfactual exercise to evaluate the time-series performance of the high-low estimator corrected for common biases reveals similar results. The sample-wide time-series correlation improves from around 48% for the high-low estimator to 80% after adjusting for the moment and small sample biases. We select two stocks, each on opposite sides of the liquidity spectrum of US stocks, to contrast the change in performance in the time-series. Figure (5) plots estimated monthly averages of effective spreads, high-low estimates, and the best-performing HL measure for Tesla ($TSLA) and ATAC Resources ($ATC), a Canadian mining company. Tesla monthly effective spreads are on average 0.14% and ATAC's over 6%. We use the zero adjustment for comparison as the other high-low *ad hoc* adjustments or using both negative and positive estimates make the measure's baseline time-series correlation null or even negative.

High-low estimates are over ten-fold the size of Tesla trading costs, correlating only 16% in the time-series. After adjusting for the two common biases, the estimator tracks the underlying spread very closely, with a correlation coefficient of 80%. Among the stocks with largest effective spreads, ATAC offers a more favorable empirical setting for the high-low estimator, which tracks effective spread levels more closely and correlate at about 50% — still somewhat poorly, nonetheless. Performance gains from correcting the biases in this case are relatively smaller, but the post-correction time-series correlation is again about 80%.

[16]Note that the unexplained performance improvement of 10% would not only include underlying properties of stock prices, but also to mechanical biases introduced by any *ad hoc* adjustment, such as zero imputation, whenever those are implemented. This underscores the overwhelming importance of the moment and small sample biases in affecting the performance of the spread estimator.

4.4 Determinants of Bias

Our framework thus far provides a decomposition of the bias in the high-low estimator into two components, which we show are structurally identical to the biases in the Roll measure. These common biases account almost entirely for the performance loss in the high-low measure. However, one puzzle remains — why is the performance of the high-low and other simple spread estimators much worse for relatively liquid stocks? The answer to this question is not only crucial to design better performance evaluation settings for new bid-ask spread estimators, but to increase the degree of confidence in out-of-sample use. If simple bid-ask spread estimators perform worse in liquid markets, without ex-ante knowledge on whether their specific setting is "liquid", how confident can researchers be in the use of low-frequency spread proxies?

Answering why spread estimators perform worse in liquid assets empirically is challenging because stocks like $TSLA and $ATC might systematically differ from each other along other dimensions than just effective spread levels. Any of these unobservable factors could drive differences in estimation performance. To cleanly observe whether there exists a relationship between bias size and the latent spread, we use a simulation environment with "ideal" financial data where once at a time, we allow for only the true spread or price variance to change, while holding constant all other pricing features. This is equivalent to imposing that $TSLA and $ATC are identical in all dimensions other than the true spread and volatility. The simulated environment has the additional benefit of controlling for potential model violations (idiosyncratic biases), as we set the data generating process of stock prices to closely match derivation assumptions.

4.4.1 Simulation evidence: bias, spread size, and frequency of negative estimates

To isolate potential model assumption violations — what we term idiosyncratic bias — in driving the observed relationship in US stock data between bias, negative estimates, and the effective spread size in the stock data, we use a simulation environment with financial price series data providing an "ideal" data generating process. This exercise confirms the importance of the underlying spread and volatility magnitude in determining the bias, which we then formally prove. It also indicates a similar relationship with respect to the frequency of indeterminate estimates, which we explore in the the Online Appendix B.

We generate 10,000 twenty-one-day months of efficient price data. Prices are created for 390 minutes m on each day, following $\mathcal{P}_m = \mathcal{P}_{m-1} + x\sigma_m$, where $x \sim \mathrm{N}(0,1)$, and $\sigma_m = \sigma/\sqrt{390}$ is the standard deviation per minute. The daily standard deviation of efficient price returns is set constant to $\sigma = 3\%$. Observed prices are then calculated after compounding or discounting true prices by half the spread S, as in equation (1). For every disjoint interval of 390 minutes, we compute daily high, low and closing observed prices. At the beginning of each month, we normalize the true price as \$ln 100. This data generating process is identical to the simulation environment in Corwin and Schultz (2012) and reflects an ideal "empirical" setting since both volatility and spread levels are held constant, and prices are always observed.

Our main goal is to analyze the behavior of estimation bias with respect to the underlying latent spread, particularly for spread sizes below 0.5%, as these correspond to the vast majority of US stock liquidity costs. Spreads below 0.5% are generally omitted from similar simulation exercises in the literature. The first panel in Figure (6) compares average monthly high-low estimates (with the zero adjustment) with the underlying, true spread level in the data generating process from 10,000 months of trading data replicated with true spread increments of 0.1%. Two salient features emerge. Estimated spreads are almost always different from the underlying true spread in levels and smaller trading costs are associated with overstated estimates, which then become downward-biased for larger spread levels. This is confirmed in the second panel, where an effective spread of 0.1% generates a bias ($\widehat{S}^{HL} - S$) of 1.2%. As we increase the latent spread size, the bias size gradually declines, until eventually becoming negative for large effective spreads.

4.4.2 Effect of spread and volatility on the bias size

The relationships unveiled in the simulated environment suggest a strong relationship between the high-low estimator bias size and the latent spread. To properly gauge such relationships as structural, however, we must again turn to theory. Proposition 3 exploits the convenience of the bias formula derived in Proposition 2 to establish how the bias magnitude behaves as we shock the underlying true spread size.

PROPOSITION 3. The bias in the implementable high-low spread estimator is decreasing in the true spread. Given sufficiently high values of volatility relative to the the latent spread (about 4 times), the bias increases in price volatility.

Proof. Appendix (II). ■

Proposition 3 shows that increasing the latent spread size decreases the *Moment Bias* term. Intuitively, the term is zero when $r_t = r_{t+1}$, which implies ϕ must equal $\sqrt{2}$. The parameter ϕ is in turn given by $\sqrt{1 + \frac{r_{t+1}^2}{r_t^2}}$, where the observed range is an additively separable linear function of S. Since the spread over the two-day trading interval is assumed constant, very large spreads move the ratio closer to one, regardless of the values of each day's true range. That is, large spreads make distinct observed ranges closer in magnitude, therefore decreasing the estimation bias size.

4.4.3 The role of volatility

Proposition 3 further establishes that the high-low estimation bias increases in volatility when σ is sufficiently large relative to the underlying spread ($\sigma/S \approx 4$). The current effective spread levels of US stocks likely attain this noise-to-signal ratio, making volatility an additional driver of empirical bias in the high-low measure. Nonetheless, because the estimation bias always decreases in the spread and volatility only matters when spreads are sufficiently low, we focus our main analysis on the relationship between bias and frequency of negative spreads with the latent spread, while keeping the volatility of the DGP constant. For completeness, we show the effect of σ on negative estimates for various levels of S in the Online Appendix B.

4.4.4 Lessons from model predictions to other simple spread estimators

Note that the effect of underlying spread and volatility levels on the bias we derived in the high-low estimator generates the same predictions as in the Roll estimator. Given our earlier result that the HL proxy suffers from the same common biases as the Roll estimator, it is perhaps not surprising that both measures become less accurate when the underlying spread decreases and volatility increases. Albeit we do not derive a formal bias expression for the close-high-low estimator, we show in Online Appendix Figures (C.1) and (C.2) that by and large the same relationships hold in the simulated data, as they also do in the US stock sample.

5 Measuring Bias Empirically

The high-low estimator bias is comprised of a moment bias term and a small sample bias term. The moment bias can be directly calculated from the data. The small sample bias, however, is a function of true daily ranges which are unobservable. Even if we wanted to estimate it using $\sqrt{2}\mathbb{E}\left[R_t^{min}\right] - \mathbb{E}\left[R_t^*\right]$, this would require a complicated approach to compute these moments in small samples, as the *Small Sample Bias* should be zero in expectation. Because of that, even though empirically correcting the high-low estimator for both biases would make its performance nearly perfect, such approach is unfortunately not feasible.

Without being able to point-identify the estimator's empirical bias, in this section we pursue the next best option: deriving theory-consistent bounds on the bias. These bounds are easily computed from the same data used to estimate the high-low measure. They are also sufficiently tight so that they are informative in a variety of settings — to both quantify the bias and control for non-classical measurement error in linear regressions, in the cross-section and time-series. The bias bounds are derived under the same assumptions as the high-low estimator and despite potentially idiosyncratic biases remaining, are empirically consistent in the vast majority of cases.

5.1 Deriving Empirical Bias Bounds

Begin by rewriting the closed-form expression for the high-low bias in Proposition 2 as $S = \widehat{S}^{HL} - Moment\ Bias - Small\ Sample\ Bias$. By construction, the moment bias is always non-negative, since $\phi \geq \sqrt{2}$. Note that when $Small\ Sample\ Bias > 0$, the following expression is always valid: $S < \widehat{S}^{HL} - Moment\ Bias$, so that the true spread S is bounded from above by correcting the high-low estimate for the moment bias.

To make progress toward signing the small sample bias term, equation (14) decomposes the term $\sqrt{2}R_t^{min} - R_t^*$. This identifies the variation that generates the failure of asymptotic distribution of the range to hold in financial data. Simply put, the small sample bias is non-zero when consecutive two-day price ranges differ significantly (the price variation on one day is much higher than the other day) or when one price range shifts vertically with respect to the previous range (one day experiences a significant and sustained price jump or drop):

$$\sqrt{2}R_t^{min} - R_t^* = \frac{1}{2}\underbrace{\left(r_t^{min} - r_t^{max}\right)}_{\Delta\ range} - \underbrace{\left(\eta_t^{max} - \eta_t^{min}\right)}_{Shift\ in\ range} + \left(\sqrt{2}-1\right)R_t^{min} \tag{14}$$

By replacing the last term in (14) with the *observed* minimum range, $\left(\sqrt{2}-1\right)r_t^{min}$, we calculate a modified version of the small sample bias, $Small\ \widetilde{Sample}\ Bias$, to verify its sign empirically. Note that this substitution implies necessarily that the true small sample bias term is overestimated, so that the procedure will fail to identify some number of negative small sample bias terms by mistakenly attributing them as positive.

Figure (7) plots 40,000 randomly chosen stock-day estimates of the "true" small sample bias, calculated using effective spreads, and $Small\ \widetilde{Sample}\ Bias$ (patterns for the full stock sample are identical). Blue markers indicate a sign correspondence between the two biases. For 90.5% of stock-day data points, using the computed version of the small sample bias that only uses observed data correctly predicts the sign of the small sample bias (the accuracy is identical using all stock-days). As we show next, this high accuracy rate is more than sufficient to robustly sign the small sample bias empirically. False negatives or false positives are displayed in gray, comprising the remaining 10.5% of observations. From these erroneous predictions, the vast majority represent false positives, which aligns with the fact that $\left(\sqrt{2}-1\right) r_t^{min} > \left(\sqrt{2}-1\right) R_t^{min}$.

To derive a lower bound on the high-low estimation bias, or equivalently, the minimum empirical bias, we compute

$$\widehat{Bias_n}^{min} = \widehat{S}_n^{HL} - f_{n(t)}\left(\left(\widehat{S}_t^{HL} - Moment\ Bias_t\right) \cdot \mathbb{1}\left\{Small\ \widetilde{Sample}\ Bias_t > 0\right\}\right)$$

where the minimum empirical estimation bias $\widehat{Bias_n}^{min}$ at frequency n (e.g., monthly, annual) is obtained from discounting monthly averages of high-low spreads with a measure of central tendency — either the monthly median or average — of upper bounds on the latent spread. Using median or mean values of the bound produces very similar empirical results, although the median usually gives slightly tighter bias bounds because it reduces the contribution of false negatives or positives when calculating $Small\ \widetilde{Sample}\ Bias$.

To obtain an upper bound on the estimation bias, note that because the small sample bias is a function of price volatility and the moment bias is determined by both volatility and the latent spread, the two biases are highly correlated: above 91% in the US stock cross-section. When empirically computing the small sample bias with $Small\ \widetilde{Sample}\ Bias$, the relationship between the two common biases becomes even stronger, as the term further introduces the latent spread into the small sample bias computation. This has an important empirical consequence: positive values of $Small\ \widetilde{Sample}\ Bias$ imply corresponding moment bias terms greater than the true small sample bias in over 96% of US stock-days, such that $S > \widehat{S}^{HL} - 2 \times Moment\ Bias$, since $Moment\ Bias > Small\ Sample\ Bias$.

By determining positive values of $Small\ \widetilde{Sample}\ Bias$, we can then derive an upper bound on the high-low estimation bias — the maximum empirical bias — as:

$$\widehat{Bias_n}^{max} = \widehat{S}_n^{HL} - f_{n(t)}\left(\left(\widehat{S}_t^{HL} - 2 \times Moment\ Bias_t\right) \cdot \mathbb{1}\left\{Small\ \widetilde{Sample}\ Bias_t > 0\right\}\right)$$

The interval $\left[\widehat{Bias_n}^{min}, \widehat{Bias_n}^{max}\right]$ provides researchers a theory-consistent approach to bound the empirical estimation bias from using the high-low spread estimator. Computing the bounds requires the same daily data used to originally obtain the high-low measure — high and low daily prices — and imposes minimum computational burden. The method also involves no *ad hoc* modifications or adjustments.

5.2 Empirical Implementation

We now assess whether the derived bias bounds are empirically useful in a variety of dimensions. We state the implementation of the empirical bounds in general terms, but remain using the US stock sample throughout to keep it consistent with all of our previous results.

Assume a researcher has computed daily high-low estimates from a certain dataset with i units (e.g., stocks, futures contracts, bonds) and $n(t)$ months of trading data with t days. The researcher then calculates monthly high-low spread averages for each unit in the cross-section. The researcher may be interested in asset pricing regressions, where returns are regressed on liquidity risk — proxied with the high-low spreads — in the style of Acharya and Pedersen (2005) and Pontiff and Singla (2019) or to control for liquidity as in McLean and Pontiff (2016). Our previous results show that simply using low-frequency spread estimates in these cases is severely problematic. Because estimation bias is a function of the underlying latent spread, linear regressions using estimated spreads will suffer from non-classical measurement error.[17]

Alternatively, the researcher may be interested in the actual measured spread values, perhaps to sort a portfolio on liquidity or to exploit changes in liquidity costs as a result of an intervention or structural changes in markets over time. In any application where the level of spread estimates matters, the use of simple spread estimators would lead to misleading results, as units with lower liquidity costs would look like illiquid assets, since estimation bias decreases in the spread size. Researchers need to have tools to assess whether a simple bid-ask spread estimate of 1% is a good proxy of an underlying spread of about 1% or whether the actual effective spread is instead 0.2% or 2%.

5.2.1 Empirical Bias in the Cross-Section

Our first empirical exercise evaluates the bias bounds in the cross-section of stocks. After computing daily high-low estimates and their monthly averages, we compute daily values of the moment bias, $\left(1+\sqrt{2}\right) r_t^{min} \left(\phi-\sqrt{2}\right)$, and daily estimates of the small sample bias, $Small\ \widetilde{Sample}\ Bias$. Note that only days with positive high-low estimates can be used to compute estimation bias, as negative estimates are not theoretically-consistent. To obtain monthly empirical bias bounds, we apply the formulas for $\widehat{Bias}_t^{min}$ and $\widehat{Bias}_t^{max}$ using days with positive small sample bias estimates. We recommend choosing the median for $f_{n(t)}$ instead of

[17] More specifically, because measurement error in the context of simple bid-ask spread estimators is negatively correlated with the effective spread — that is, mean reverting — the attenuation bias decreases, which may even flip the sign of the estimated effect.

the mean as this reduces the potential influence of idiosyncratic biases in making the signs of $\widehat{Small\ Sample\ Bias}$ and the true small sample bias fail to coincide.[18]

Figure (8) plots cross-sectional averages of three variables: the actual high-low empirical estimation bias ($Bias_n$), which measures the average percentage-point deviation of the estimator from the effective spread, and the minimum and maximum empirical estimation bias bounds. We present monthly averages over our time period in three panels — all US stocks, US stocks below the median effective spread — about 0.21% — and above the median. Across all panels, the estimated biases successfully track and bound the actual high-low estimation bias. The correlation between each bound and $Bias_n$ is remarkable — over 98%. Therefore, including the bias bounds or its midpoint in linear regressions would be highly effective in controlling for non-classical measurement error when using daily or monthly spread estimates.

5.2.2 Recovering Effective Spreads

The bounds not only provide the researcher with a precise estimate of how much bias on average the high-low estimator suffers from in her particular context, but also allow her to obtain theoretically-consistent bounds on the effective spread, significantly improving the degree of confidence in the use of the measure. To see how, consider that the researcher has calculated the average monthly high-low estimate for all US stocks (top panel of the figure), which is 1.80%. Next, she obtains the average estimated lower bound on the bias — 0.54% — and the estimated maximum empirical bias of 1.58% (for reference, the actual estimation bias in the US stock sample is 0.95%, which the researcher cannot observe in practice but that is successfully bounded by our bias estimates).

Discounting the high-low estimate with each bound, that is, 1.80%−1.58%= 0.22% and 1.80%−0.54%= 1.26%, yields back implied effective spread bounds, which in this case results in the prediction that the actual effective spread is between 0.22% and 1.26%, or should be approximately equal to 0.74% at the midpoint. The average monthly effective spread in the sample is 0.85%. This means that the implied effective spread midpoint generated by our empirical bias bounds is only 15% above the effective spread, while the high-low measure is 112% higher.

Replicating the exercise for stocks below the effective spread median gives a similar high degree of precision, with an important distinction. Recall that when we derive the maximum empirical bound on the spread, the relationship $S > \widehat{S}^{HL} - 2 \times Moment\ Bias$ is implied by a positive value of the estimated small sample bias. Note that when the true spread S is very

[18]Although the choice between the mean or median in our US stock sample leads to minimal empirical differences, false positives are associate with very large bounds on the true spread, so that using the median in applications where idiosyncratic biases may be important is akin to trimming abnormal values.

small, the moment bias term approaches the small sample bias, so that the inequality becomes close to binding.[19] This can be seen in the second panel of Figure (8), where the actual estimation bias falls very close to the upper bias bound.

As a result, if the researcher has reason to believe that her application context involves liquid assets, using the upper bound to obtain the implied effective spread will more accurately approximate the effective spread *level* than the implied spread midpoint — 0.03% for an effective spread of 0.09% and a high-low estimate of 1.28%. To control for estimation bias, however, any of the bounds correlates similarly well, above 98%. As an additional example, we revisit the example of Tesla stocks we used in Figure (5). Applying the upper bias bound in the $TSLA time-series recovers an average implied effective spread of 0.26% (for an effective spread of 0.14% and a high-low estimate of 1.78%).

An inverted relationship exists in illiquid assets. The actual estimation bias will approximate more closely the lower bound on the estimation bias, as very large spreads make the inequality $S < \widehat{S}^{HL} - \textit{Moment Bias}$ closer to binding. This appears in the bottom panel of Figure (8), where the lower bound tracks the estimation bias very closely. In cases where the researcher has reason to believe that spreads are very large, using the lower bound will give more accurate implied effective spread levels than the midpoint. For the subsample of stocks above the median effective spread, this approach recovers a spread estimate of 1.64% for an effective spread of 1.46% (and a high-low estimate of 2.21%). Revisiting the time-series example of the highly illiquid stock $ATC from Figure (5) and using the lower bias bound yields an almost perfect point-estimate of the effective spread: 6.26% for an effective spread of 6.26% and a high-low of 5.72%).

5.2.3 Empirical Bias in the Time-Series

We now use the bias bounds to estimate *stock*-level high-low biases. Figure (9) shows again the actual high-low empirical estimation bias and the minimum and maximum bias bounds. This time, we average monthly estimates for each stock over the entire time period, so that each stock has a single value for each of the three bias variables. The top panel shows the distribution of the empirical bias and its bounds for all US stocks. We sort observations along the x-axis from the largest estimation bias (in percentage-points) to the smallest (that is, when the high-low measure understates the effective spread). To make the interpretation easier, we

[19] Intuitively, because the moment bias depends both on volatility and the spread and the small sample bias only on volatility, a very small spread makes both variables have a perfect correlation. Formally, the moment bias is equal to the small sample bias when $R_t^{max}/R_t^{min} \approx 1.24$, given $\lim_{S \longrightarrow 0} \left\{ r_t^{min} \left(\phi - \sqrt{2} \right) - \sqrt{2} R_t^{min} + R_t^* \right\}$.

report the x-axis as the cumulative share of stocks from the left (0% of the total sample) to the right (100% of stocks).

Consistent with the previous results in the cross-section, the performance of the empirical bias bounds in the time-series is also robust. For the vast majority of stocks, the true estimation bias lies inside the theory-consistent bounds, which behave properly in almost all stocks that generate upward-biased high-low estimates. This includes around 80% of the stocks in the cross-section. Stocks that generate downward-biased high-low spreads — primarily illiquid stocks — are the only cases where the actual empirical bias lies outside the bounds $\widehat{Bias}^{min}$ and $\widehat{Bias}^{max}$.

To gain further insight into the relationship between the estimated bounds and the actual bias, the second panel in the figure shows results only for stocks with spreads below the sample median of about 0.21%. In this case, almost all stocks have an actual high-low bias inside the estimated bounds, which almost perfectly track bias levels from 4 to about 1 percentage point. When normalized by the underlying effective spread, these biases amount anywhere from 5-fold to 20-fold the actual spread level. Similarly, the last panel in the figure shows the performance of the bias bounds for stocks above the median effective spread. This subsample includes the illiquid stocks for which the high-low estimator produces severely understated spreads, which lead to the estimation bias to lie outside the bounds. Nonetheless, for actual biases from 12 to 0.5 percentage point — about 75% of stocks in this subsample — the empirical bounds once again work consistently well.

5.2.4 Sorting Stocks Based on Liquidity

The analysis above illustrates the time-series usefulness and reliability of the empirical bias bounds. Researchers can credibly estimate the high-low bias to measure it directly or report it to establish credibility of spread estimates. To provide another example of how researchers are able to benefit from the approach, we analyze the gains in predictive power when sorting stocks on liquidity by using the bias bounds. Even coarsely sorting assets may result in severely distorted rankings and portfolios if the levels of the sorting variable fail to pick up the latent factor. For example, a researcher interested in dividing stocks in portfolios based on liquidity levels who would use the high-low proxy would assign several highly liquid assets to an illiquid portfolio, since these spread estimates will be more upward biased than estimates of actually illiquid assets.

Our minimum bias bounds can be used to "recover" theory-consistent bounds on the effective spread, lending credibility to the use of the high-low estimator in settings where estimate levels across the cross-section matter. Table (1) performs a series of exercises confirming this point. The first question we answer there is: how well does the ranking of spread estimates predict the actual effective spread ranking? In columns (1) through (4), we regress a ranking variable at the stock level — which orders observations from the lowest effective spread to the largest effective spread — on the spread ranking generated by the high-low measure without *ad hoc* adjustments and with the zero adjustment, and by correcting the high-low spread with the bias bound.

The first column shows that a researcher using only the positive spread estimates obtained from the high-low measure would predict the correct ranking of effective spreads by about 72%. Using the zero adjustment increases the predictive power to 82%, since the *ad hoc* adjustment introduces mechanical attenuation bias that coincidentally offsets some of the upward bias in liquid stocks generated by the small sample and moment biases. Sorting based on the bias bound correction significantly improves accuracy — a move of 10 positions in the liquidity rank is associated with a move of 9 positions implied by our measure. The spread implied by the bias bounds absorbs 80% of the variation in the effective spread ranking, several percentage points more than the standard high-low measures. This last point becomes even clearer when we regress the effective spread rank on both the high-low and the bound-corrected spread, where almost no additional variation is absorbed vis-à-vis only using the bound-corrected spread.

Columns (5) through (8) perform a similar exercise, now using a more coarse sorting approach. We divide stocks in 10 portfolios based on effective spread deciles and regress the decile rank on the similar measure constructed with the spread estimates. The overall direction and differences in magnitude of results are identical to the previous columns, with the spread implied by the bias bounds performing significantly better in allocating stocks to the correct portfolio based on the effective spread.

6 Conclusion

This paper studies why simple bid-ask spread estimators present conflicting performance results in different markets, often yield negative or indeterminate estimates, and how to empirically assess and address these issues. Using the popular high-low estimator as our main setting, we show that the estimator's misbehavior is almost entirely driven by two common biases: a moment and a small sample bias. These are the same sources of bias found in the seminal Roll estimator, with idiosyncratic factors dominating relative to specific market settings, or conditions, and model assumptions failure. We show how to empirically measure the minimum and maximum estimation bias from using the high-low measure, equipping researchers to remedy issues arising with the use of low-frequency spread estimators, including non-classical measurement error in regression analysis, portfolio sorting on liquidity, and tracking spread levels in the cross-section and time-series.

We also draw several recommendations from our analysis, from the development of new estimators to performance evaluation. To more credibly establish the quality and recommend a certain spread proxy, performance evaluation studies using tick data should include more detailed subsampling analysis, particularly by analyzing the relationship of correlation coefficients in the cross-section across different effective spread groups. Moreover, we also suggest in sensitivity analyses that performance is measured after dropping the most illiquid units from the sample as well as in random subsamples from the data when the cross-section is sufficiently large. Coefficient stability across these different tests indicate a greater degree of out-of-sample confidence.

Going forward, newly developed spread estimators using transaction data that share the structure of the Roll framework should be subjected to rigorously designed simulation environments with true spreads in the data generating process well below 0.5%, the usual minimum spread used in the literature, and verify the role of price volatility and trading costs in potential misbehavior. More preferably, a formal analysis of the existence of the moment or small sample bias in new measures is also suggested, as these two sources severely limit performance of simple bid-ask spread estimators.

References

Abdi, Farshid and Angelo Ranaldo (2017) "A Simple Estimation of Bid-Ask Spreads from Daily Close, High, and Low Prices," *The Review of Financial Studies*, Vol. 30, No. 12, pp. 4437–4480.

Acharya, Viral V. and Lasse Heje Pedersen (2005) "Asset pricing with liquidity risk," *Journal of Financial Economics*, Vol. 77, No. 2, pp. 375–410.

Alizadeh, Sassan, Michael W. Brandt, and Francis X. Diebold (2002) "Range-Based Estimation of Stochastic Volatility Models," *The Journal of Finance*, Vol. 57, No. 3, pp. 1047–1091.

Barardehi, Yashar H, Dan Bernhardt, and Ryan J Davies (2018) "Trade-Time Measures of Liquidity," *The Review of Financial Studies*, Vol. 32, No. 1, pp. 126–179.

Bernstein, Asaf, Eric Hughson, and Marc Weidenmier (2019) "Counterparty Risk and the Establishment of the New York Stock Exchange Clearinghouse," *Journal of Political Economy*, Vol. 127, No. 2, pp. 689–729.

Bleaney, Michael and Zhiyong Li (2015) "The Performance of Bid-Ask Spread Estimators Under Less Than Ideal Conditions," *Studies in Economics and Finance*, Vol. 32, No. 1, pp. 98–127.

Chakravarty, Sugato and Asani Sarkar (2003) "Trading Costs in Three U.S. Bond Markets," *The Journal of Fixed Income*, Vol. 13, No. 1, pp. 39–48.

Corwin, Shane A. and Paul Schultz (2012) "A Simple Way to Estimate Bid-Ask Spreads from Daily High and Low Prices," *The Journal of Finance*, Vol. 67, No. 2, pp. 719–760.

Easley, David, Marcos López de Prado, Maureen O'Hara, and Zhibai Zhang (2020) "Microstructure in the Machine Age," *The Review of Financial Studies*.

Feller, William (1951) "The Asymptotic Distribution of the Range of Sums of Independent Random Variables," *The Annals of Mathematical Statistics*, Vol. 22, No. 3, pp. 427–432.

Fong, Kingsley Y. L., Craig W. Holden, and Charles A. Trzcinka (2017) "What Are the Best Liquidity Proxies for Global Research?" *Review of Finance*, Vol. 21, No. 1.

French, Kenneth R. and Richard Roll (1986) "Stock Return Variances," *Journal of Financial Economics*, Vol. 17, No. 1, pp. 5 – 26.

Garman, Mark B. and Michael J. Klass (1980) "On the Estimation of Security Price Volatilities from Historical Data," *The Journal of Business*, Vol. 53, No. 1, pp. 67–78.

Goyenko, Ruslan Y., Craig W. Holden, and Charles A. Trzcinka (2009) "Do Liquidity Measures Measure Liquidity?" *Journal of Financial Economics*, Vol. 92, No. 2, pp. 153 – 181.

Hagströmer, Björn (2021) "Bias in the Effective Bid-Ask Spread," *Journal of Financial Economics*.

Harris, Lawrence (1990) "Statistical Properties of the Roll Serial Covariance Bid/Ask Spread Estimator," *The Journal of Finance*, Vol. 45, No. 2, pp. 579–590.

Hasbrouck, Joel (2004) "Liquidity in the Futures Pits: Inferring Market Dynamics from Incomplete Data," *The Journal of Financial and Quantitative Analysis*, Vol. 39, No. 2, pp. 305–326.

Holden, Craig W. and Stacey Jacobsen (2014) "Liquidity Measurement Problems in Fast, Competitive Markets: Expensive and Cheap Solutions," *The Journal of Finance*, Vol. 69, No. 4, pp. 1747–1785.

Hou, Kewei, Chen Xue, and Lu Zhang (2018) "Replicating Anomalies," *The Review of Financial Studies*, Vol. 33, No. 5, pp. 2019–2133.

Jahan-Parvar, Mohammad R and Filip Zikes (2019) "When do Low-Frequency Measures Really Measure Transaction Costs?".

Karnaukh, Nina, Angelo Ranaldo, and Paul Söderlind (2015) "Understanding FX Liquidity," *The Review of Financial Studies*, Vol. 28, No. 11, pp. 3073–3108.

Lesmond, David A., Joseph P. Ogden, and Charles A. Trzcinka (1999) "A New Estimate of Transaction Costs," *The Review of Financial Studies*, Vol. 12, No. 5, pp. 1113–1141.

Li, Xindan, Avanidhar Subrahmanyam, and Xuewei Yang (2018) "Can Financial Innovation Succeed by Catering to Behavioral Preferences? Evidence from a Callable Options Market," *Journal of Financial Economics*, Vol. 128, No. 1, pp. 38 – 65.

Lou, Xiaoxia and Tao Shu (2017) "Price Impact or Trading Volume: Why is the Amihud (2002) Measure Priced?" *The Review of Financial Studies*, Vol. 30, No. 12, pp. 4481–4520.

Marshall, Ben R., Nhut H. Nguyen, and Nuttawat Visaltanachoti (2011) "Commodity Liquidity Measurement and Transaction Costs," *Review of Financial Studies*.

——— (2018) "Do Liquidity Proxies Measure Liquidity Accurately in ETFs?" *Journal of International Financial Markets, Institutions and Money*, Vol. 55, pp. 94–111.

McLean, R. David and Jeffrey Pontiff (2016) "Does Academic Research Destroy Stock Return Predictability?" *The Journal of Finance*, Vol. 71, No. 1, pp. 5–32.

Nieto, Belén (2018) "Bid-Ask Spread Estimator from High and Low Daily Prices: Practical Implementation for Corporate Bonds," *Journal of Empirical Finance*, Vol. 48, pp. 36 – 57.

Parkinson, Michael (1980) "The Extreme Value Method for Estimating the Variance of the Rate of Return," *The Journal of Business*, Vol. 53, No. 1, pp. 61–65.

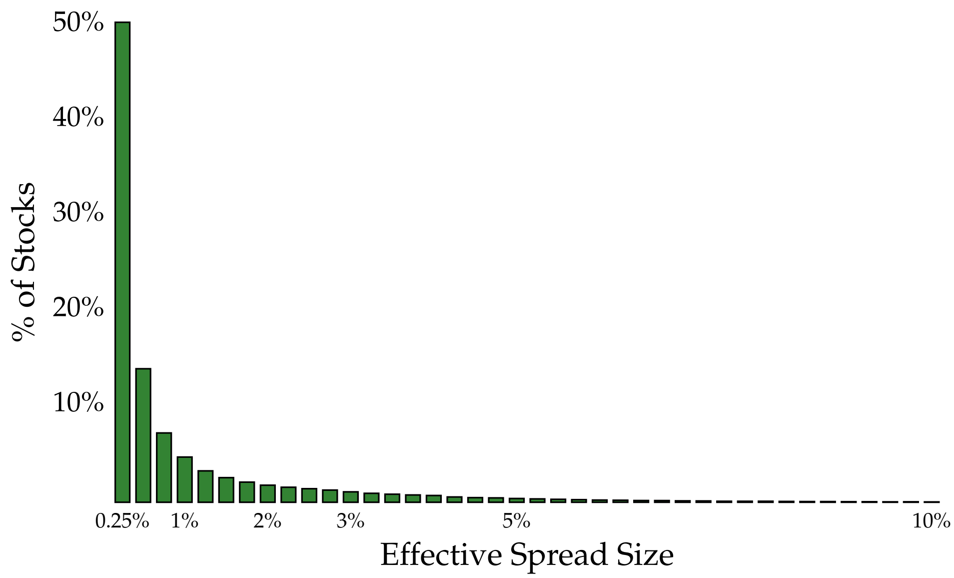
Patton, Andrew J. and Brian M. Weller (2020) "What You See is Not What You Get: The Costs of Trading Market Anomalies," *Journal of Financial Economics*.

Pontiff, Jeffrey and Rohit Singla (2019) "Liquidity Risk?" *Critical Finance Review*, Vol. 8, No. 1-2, pp. 257–276.

Roll, Richard (1984) "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market," *The Journal of Finance*, Vol. 39, No. 4, pp. 1127–1139.

Thompson, Sarahelen R. and Mark L. Waller (1987) "The Execution Cost of Trading in Commodity Futures Markets," *Food Research Institute Studies*, Vol. 20, No. 2, pp. 141. – 24.

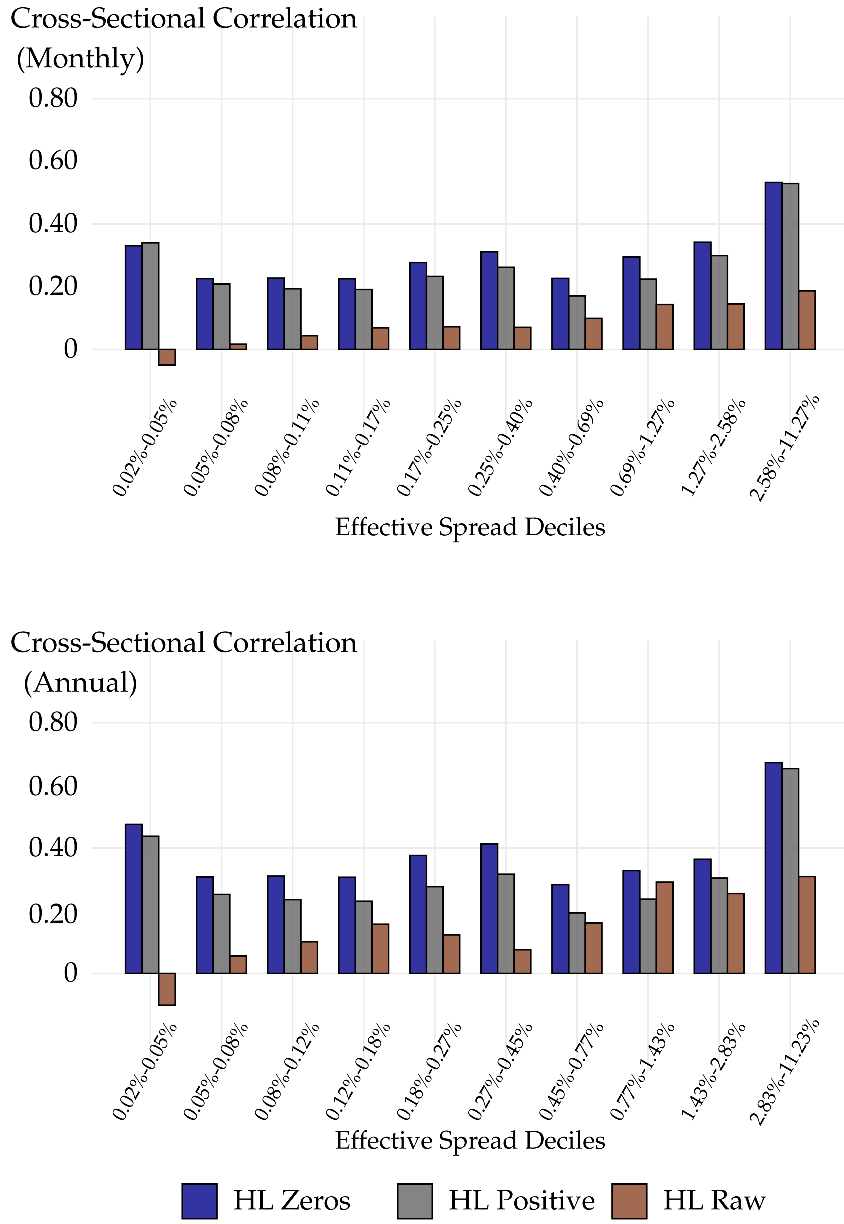
Figures



Over 70% of US stocks have effective spreads lower than 1%. There are very few stocks with spreads greater than 5%.

Notes: This figure shows percentiles of monthly effective spreads from DTAQ US stock data. There are 3,398 stocks in total and sample construction details are described in Appendix (I).

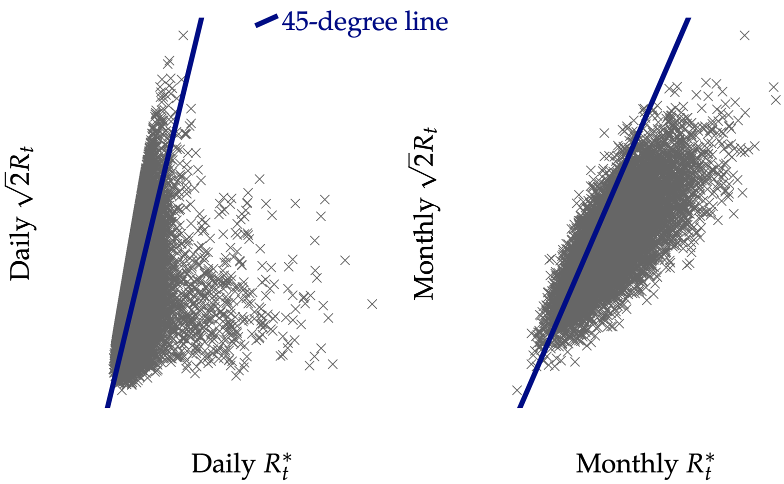
Figure 1: Distribution of Effective Bid-Ask Spreads in US Stocks



In all effective spread size deciles, monthly averages of high-low estimates without ad hoc adjustments correlate poorly with monthly average effective spreads. Correlations from the estimates with the positives and zeros adjustments behave similarly, with the best performance in the least liquid stocks (above 40% correlation with effective spreads).

Notes: This figure shows cross-sectional correlations between three versions of the high-low spread estimator and effective spreads from DTAQ US stock data. There are 3,398 stocks in total and sample construction details are described in Appendix (I).

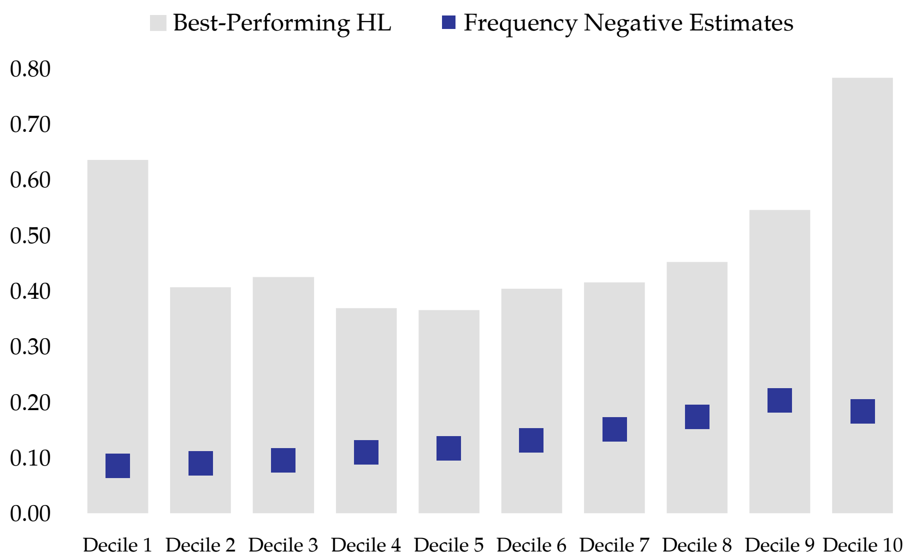
Figure 2: Cross-Sectional Correlation: HL Estimates and Effective Spread By Effective Spread Decile Size



The identity required for the small sample bias in the high-low estimator to be zero fails to hold in almost all of 210,000 trading days and 10,000 monthly averages.

Notes: This figure compares 210,000 daily values of $\sqrt{2}R$ and R^* (left) and 10,000 monthly averages of each range in trading data simulated with ideal conditions. The main text contains more information on the data generating process.

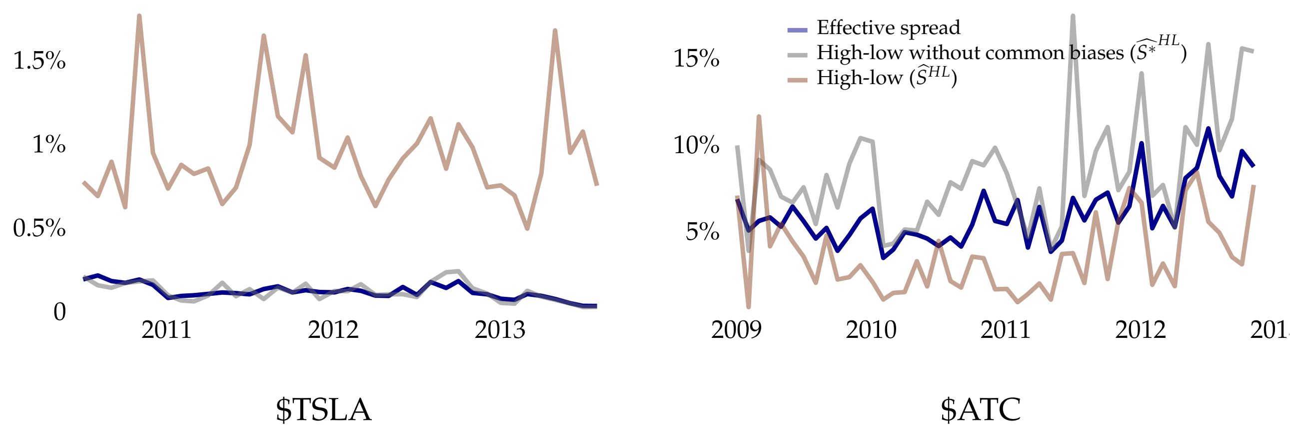
Figure 3: Relationship Between $\sqrt{2}R$ and R^* Under Ideal Conditions



In all effective spread deciles, monthly averages of the best-performing high-low estimator correlate twice as well with monthly average effective spreads as high-low estimates.

Notes: This figure shows cross-sectional correlations between our calculated best-performing high-low spread estimator and effective spreads from DTAQ US stock data. There are 3,398 stocks in total and sample construction details are described in Appendix (I). The *Best-Performing HL* is a bias-free version of the high-low measure, obtained as $\widehat{S}^{HL} - \mathit{Moment\ Bias} - \widehat{\mathit{Small\ Sample}}\ \mathit{Bias}$.

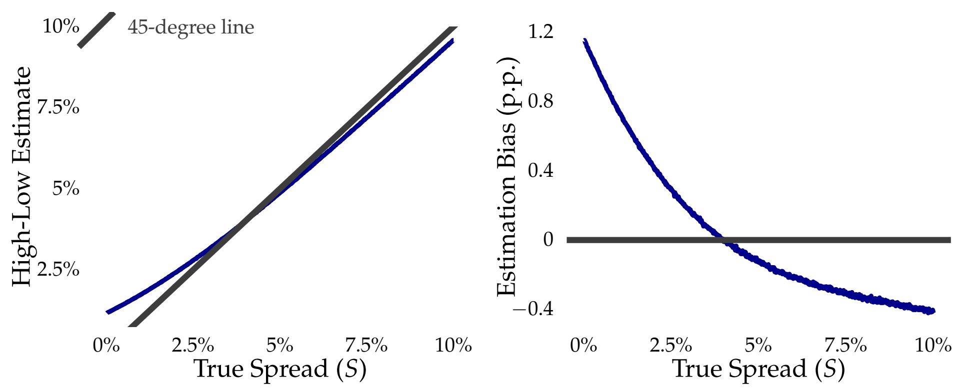
Figure 4: Cross-Sectional Correlation: Best-Performing HL & Effective Spread by Effective Spread Decile Size



The best-performing high-low measure tracks the effective spread almost perfectly in a very liquid stock ($TSLA). When the stock is relatively iliquid ($ATC), the measure still improves considerably on the standard high-low estimator.

Notes: This figure shows monthly averages of effective spreads and two versions of the high-low estimator for Tesla and ATAC Resources stocks. The *HL Zeros Adjustment* represents monthly averages of the high-low estimator with zeros imputed for negative daily estimates. The *Best-Performing HL* is a bias-free version of the high-low measure, obtained as $\widehat{S}^{HL} - \mathit{Moment\ Bias} - \mathit{Small}\ \widehat{\mathit{Sample}}\ \mathit{Bias}$.

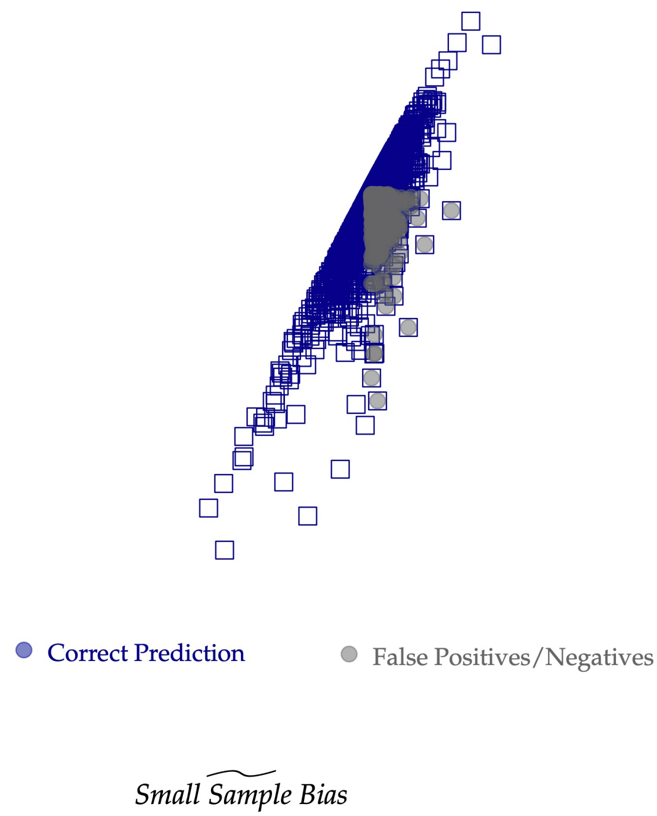
Figure 5: Two Examples of Performance: Best-Performing HL Estimator



For small spreads (below 1%), the high-low estimator suffers from an upward bias. As the spread becomes larger, the bias decreases. For reference, the median effective spread in US stocks during 2003-2015 is 0.3%.

Notes: This figure computes 10,000 monthly averages of the high-low spread estimator (with zeros adjustment) for each 0.1-percentage-point increment in the spread between 0.1% and 10% from simulated data. The data generating process is described in detail in the text and is designed to maintain all model assumptions, including constant true spreads in each sample. Bias is defined as the difference between the estimated spread and the true spread.

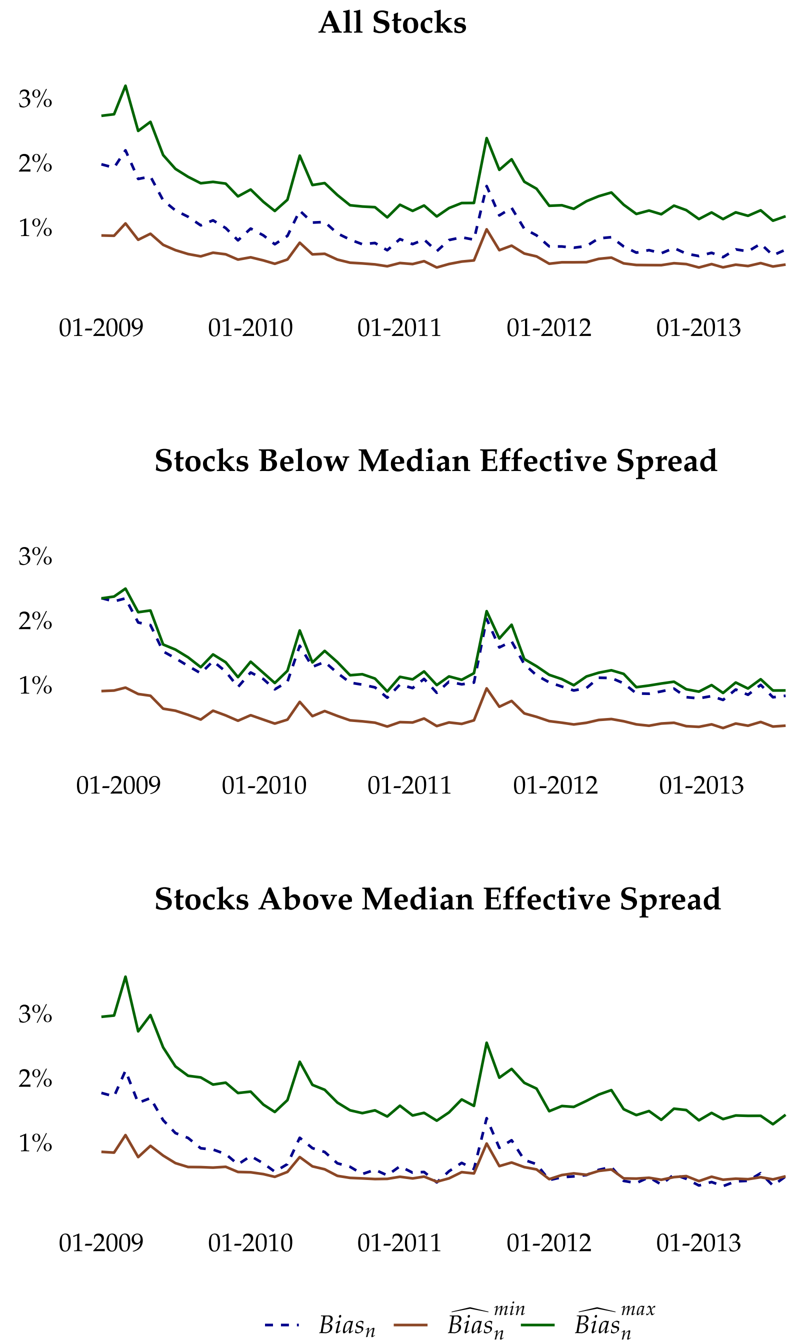
Figure 6: Bias and True Spread Under Ideal Conditions



The empirical approach to predict the sign of the small sample bias has over 90% accuracy.

Notes: This figure plots 40,000 randomly chosen stock-day estimates of an empirical approach to measure the small sample bias in the high-low measure. $\widetilde{Small\ Sample}\ Bias$ is calculated as $r_t^{min} + \left(1 + \sqrt{2}\right)(0.5\Delta r_t - \Delta \eta_t)$. Blue markers indicate the correct correspondence between the respective derived small sample bias and its theory-implied value, thereby correctly predicting the sign of the small sample bias. Gray markets give cases where such correspondence fails, resulting in false positives or negatives.

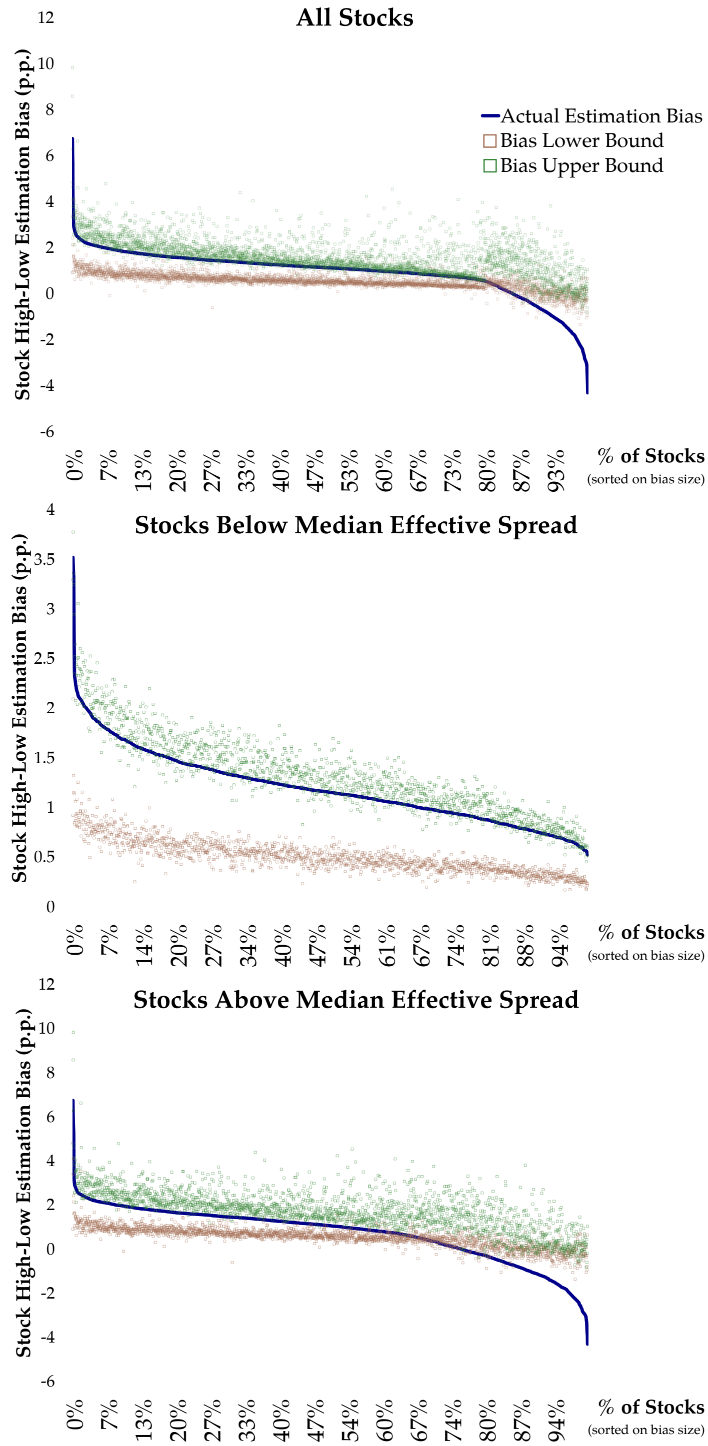
Figure 7: Testing the Accuracy of the Empirical Approach to Sign the Small Sample Bias



The empirical estimation bias bounds track the actual estimation bias almost perfectly.

Notes: This figure compares the average cross-sectional high-low estimation bias with three measures: $Bias_n$, which is the actual empirical bias of the high-low spread estimator, the maximum empirical bias estimate $\widehat{Bias}_n^{max}$ and the minimum empirical bias estimate $\widehat{Bias}_n^{min}$. For a given stock×month, we calculate the empirical estimation bias of the high-low measure by discounting a stock's monthly effective spread with the monthly high-low estimate. We then compute monthly averages across all stocks. Empirical estimation bias bounds are obtained following the steps described in the text. The y-axis gives the empirical bias in percentage points.

Figure 8: Bounding the High-Low Empirical Bias in US Stocks



Estimated bias bounds are empirically consistent for over 80% of stocks.

Notes: This figure shows stock-level average actual estimation bias of the high-low measure, as well as estimates of average maximum and minimum empirical biases. Each panel orders stocks from the largest upward estimation bias in percentage points on the left to the largest downward estimation bias to the right. Therefore, gradually moving toward the right end of the x-axis completes the US stock cross-section sorted according to the high-low bias. The top panel shows all stocks, the middle panel subsets the stock sample and plots only stocks below the median effective spread during the period of 0.21%. The bottom panel plots stocks above the median spread.

Figure 9: Stock-level Estimation Bias Bounds

Tables

Table 1: Ranking Stocks Based on Liquidity

	Effective Spread Rank (1, . . . , I stocks)				**Effective Spread Portfolios** (1, . . . , 10 deciles)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High-low (positive estimates)	**0.724**				**0.713**			
	[*70.71*]				[*68.76*]			
High-low (zero adjustment)		**0.822**		**0.237**		**0.810**		**0.229**
		[*103.72*]		[*13.39*]		[*98.14*]		[*13.59*]
Spread implied by bias bounds			**0.902**	**0.701**			**0.890**	**0.697**
			[*133.03*]	[*39.33*]			[*131.34*]	[*41.26*]
R-squared	0.52	0.67	0.80	0.82	0.51	0.65	0.79	0.80

Notes: This table compares how well each different estimation method to compute stock-level bid-ask spreads correlates with orderings based on actual effective spreads. Columns (1) through (4) regress the effective spread rank of each stock i ($1 =$ smallest spread, $I =$ largest spread) on a ranking variable obtained with high-low estimates (without *ad hoc* adjustments) in (1), high-low estimates with the zero adjustment (columns (2) and (4)), and using the estimation bias bounds derived in the main text in (3) and (4). A coefficient estimate of 1 indicates that the liquidity ranking constructed with the given spread proxy perfectly predicts the actual ordering of stocks based on effective spreads. In columns (5)—(8), we sort stocks in 10 portfolios depending on their effective spreads (effective spread deciles) and regress a stock's liquidity decile on the same measure constructed using the liquidity proxies. A coefficient estimate equal to 1 indicates that the spread estimator perfectly assigns stocks to their correct liquidity portfolio. Robust standard errors are omitted in the table and we report [*t-stat*] values.

Appendix

I Sample Handling & Construction

We briefly describe the data cleaning and computation of effective spreads from DTAQ US stock data. We closely follow the procedures in Holden and Jacobsen (2014), using Daily TAQ data. We only retain bids and asks generating bid-ask spreads within five dollars and 2.5 dollars above or below the previous midpoint. Intraday effective spreads are calculated as:

$$\frac{2|P_k - M_k|}{M_k}$$

where $M_k = (B_k + A_t)/2$ is the midpoint of the national best bid-offer quotes of the k-th trade, given by P_k. We then compute dollar-weighted averages of the above to compute daily effective spreads used in the main paper. We then merge daily effective spreads to daily CRSP data, which includes end-of-day prices used to compute low-frequency spread estimators.

II Proofs & Other Derivations

This section presents proofs and auxiliary derivations of the main theoretical results and propositions in the paper.

II.1 Useful Relationships

Simplified Form of α. Let α be a function of β and γ without explicitly considering each parameter's own arguments (observed high and low prices). The formula for α derived by Corwin and Schultz (2012) is:

$$\alpha(\beta, \gamma) = \frac{\sqrt{2\beta} - \sqrt{\beta}}{3 - 2\sqrt{2}} - \sqrt{\frac{\gamma}{3 - 2\sqrt{2}}}.$$

Call $\varphi \equiv 3 - 2\sqrt{2}$, $\varphi > 0$. After rewriting the above, we have

$$\alpha = \frac{\sqrt{\varphi}\left[\sqrt{\beta}\left(\sqrt{2} - 1\right) - \sqrt{\gamma}\sqrt{\varphi}\right]}{\varphi\sqrt{\varphi}} = \frac{\sqrt{\beta}\left(\sqrt{2} - 1\right) - \sqrt{\gamma}\left(\sqrt{3 - 2\sqrt{2}}\right)}{3 - 2\sqrt{2}}$$

which can be further simplified into

$$\alpha = \frac{\left(\sqrt{2}-1\right)\left(\sqrt{\beta}-\sqrt{\gamma}\right)}{3-2\sqrt{2}}$$

since $\sqrt{3-2\sqrt{2}} = \sqrt{2}-1$. After employing a similar replacement for φ, we arrive at the simplified version of α used in the main text:

$$\alpha = \frac{\left(3+2\sqrt{2}\right)\left(\sqrt{2}-1\right)\left(\sqrt{\beta}-\sqrt{\gamma}\right)}{\left(3+2\sqrt{2}\right)\left(3-2\sqrt{2}\right)} = \left(\sqrt{2}+1\right)\left(\sqrt{\beta}-\sqrt{\gamma}\right).$$

SIMPLIFIED FORM OF THE HIGH-LOW ESTIMATOR. The original closed-form of the high-low spread estimator is

$$S^{HL} = 2\left(\frac{e^{\alpha}-1}{e^{\alpha}+1}\right).$$

A well-known representation of the hyperbolic tangent function $\tanh(\cdot)$ is given by

$$\tanh x \equiv \frac{e^{2x}-1}{e^{2x}+1}$$

for which we can set $x \equiv \alpha/2$, and the simplified version of the HL estimator follows

$$S^{HL} = 2\tanh\left(\frac{\alpha}{2}\right)$$

or, equivalently, the logistic representation below:

$$S^{HL} = 2\left(L\left(\alpha\right)-1\right).$$

II.2 Main Proofs

PROOF OF PROPOSITION 1. Since $\mathbb{E}\left[R_t^*\right] = \sqrt{2}\mathbb{E}\left[R_t\right]$, unbiasedness of $\widetilde{S}^{HL}$ immediately follows. For $\widehat{S}^{HL}$ to be unbiased, it must be that $\left(1+\sqrt{2}\right)\left(\phi R_t^{min} - R_t^*\right) = \left(1-\omega\right)S$. Expanding both sides yields

$$\mathbb{E}\left[\phi\right]\mathbb{E}\left[R_t\right] - \sqrt{2}\mathbb{E}\left[R_t\right] = \sqrt{2}S - \mathbb{E}\left[\phi\right]S$$

which holds if and only if $\phi = \sqrt{2}$, which in turn is implied by $r_t = r_{t+1}$, $\forall t$. Note that $\phi = \sqrt{2}$ results in $\omega = 1$. This completes the proof.

Proof of Proposition 2. Start with

$$\left[\left(\phi + \sqrt{2}\phi - 1 - \sqrt{2}\right) S + \left(1 + \sqrt{2}\right)\left(\phi R_t^{min} - R_t^*\right)\right] - S$$

which yields the following after straightforward algebra and by noting that $r_t = R_t + S$:

$$-2r_t^{min} + 2R_t^{min} - \sqrt{2}S + \phi r_t^{min} + \sqrt{2}\phi r_t^{min} - R_t^*\left(1 + \sqrt{2}\right)$$

Adding and subtracting $\sqrt{2}r_t^{min}$ from the above and further algebra finally gives

$$\left(1 + \sqrt{2}\right)\left[r_t^{min}\left(\phi - \sqrt{2}\right) + \left(\sqrt{2}R_t^{min} - R_t^*\right)\right]. \tag{15}$$

We know must show that the expression in (15), which measures the daily estimation error in the high-low proxy, is equivalent in expectation to the estimator's bias. First, consider the implementable high-low spread formula and its expected value:

$$\widehat{S}^{HL} = \left(1 + \sqrt{2}\right)\left(\phi r_t^{min} - r_t^*\right)$$

and then

$$\mathbb{E}\left[\widehat{S}^{HL}\right] - S = \left(1 + \sqrt{2}\right)\left(\mathbb{E}\left[\phi r_t^{min}\right] - \mathbb{E}\left[R_t^*\right]\right) - S\left(2 + \sqrt{2}\right)$$

which gives the estimator's bias and is identical to the expectation of (15).

Proof of Proposition 3. For simplicity and without loss of generality, assume that $R_t < R_{t+1}$. Therefore, $\partial\phi/\partial S < 0$ and thus the following holds for the bias in Proposition 2:

$$\frac{\partial\left(\widehat{S}^{HL} - S\right)}{\partial S} = \left(1 + \sqrt{2}\right)\left(R_t^{min}\frac{\partial\phi}{\partial S} + \phi + S\frac{\partial\phi}{\partial S} - \sqrt{2}\right) =$$

$$-\frac{\left(1 + \sqrt{2}\right)\left(R_t^{min}\sqrt{\frac{2\left(R_t^{max} + S\right)^2}{\left(R_t^{min} + S\right)^2} + 2} + S\sqrt{\frac{2\left(R_t^{max} + S\right)^2}{\left(R_t^{min} + S\right)^2} + 2} - R_t^{max} - R_t^{min} - 2S\right)}{\left(R_t^{min} + S\right)\phi} < 0$$

since $R_t^{min} = R_t < R_{t+1} = R_t^{max}$ and $S > 0$. For the volatility result, first remember that $\phi r_t^{min} \equiv \sqrt{\widehat{\beta}}$. Thus, the expected value of the bias expression is given by:

$$\left(1+\sqrt{2}\right)\left(\mathbb{E}\left[\sqrt{(r_t^{max})^2 + \left(r_t^{min}\right)^2}\right] - \sqrt{2}\mathbb{E}\left[r_t^{min}\right]\right) \leq \left(1+\sqrt{2}\right)\left[\sqrt{8\ln 2\sigma^2 + 4S\sqrt{\frac{8}{\pi}}\sigma + 2S^2} - \sqrt{2}\sqrt{\frac{8}{\pi}}\sigma - \sqrt{2}S\right]$$

where the inequality holds from the Jensen's inequality. The derivative of the above with respect to σ is positive if and only if:

$$\frac{\sigma}{S} > 4.20 - \varepsilon \tag{16}$$

where $\varepsilon \geq 0$ is the error induced by the concavity of β. This completes the proof.

Online Appendix

The Bias of Simple Bid-Ask Spread Estimators

Pedro Tremacoldi-Rossi
Columbia University

Scott H. Irwin
University of Illinois

Contents

A Additional Performance Evaluation in US Stocks

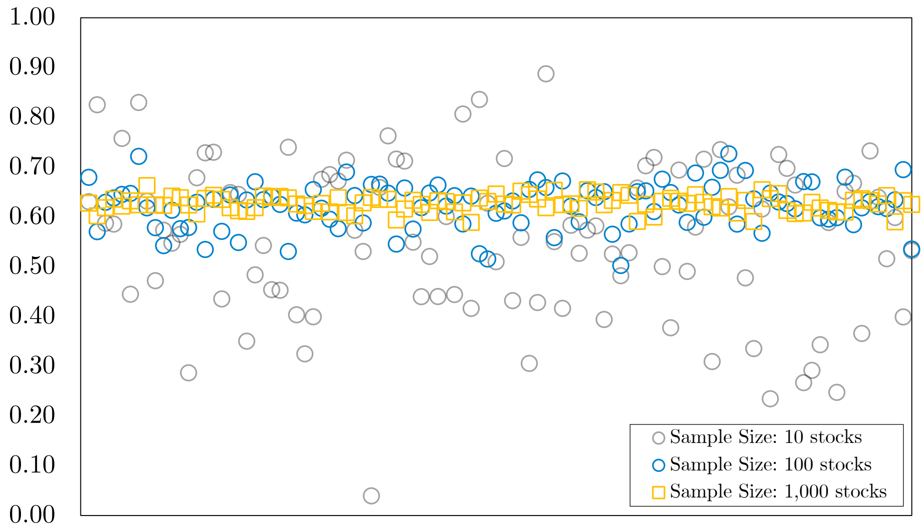
This section builds on Section 3 in the main paper. We conduct additional tests designed to assess the robustness and stability of the performance of the high-low spread estimator in US stocks. This sample is the standard setting used in horse-race type studies that evaluate the performance of multiple bid-ask spread estimators and recommend out-of-sample use — i.e., in other settings — based on in-sample performance.

Cross-section size and performance stability. We begin with a second factor that accounts for the performance of high-low spread estimates benchmarked in stock data. Our goal is to consider how the estimator's performance varies were a practitioner to observe random samples of an underlying asset class - i.e., the cross-section - with different sizes. In particular, we are interested in draws that are "small" compared to the full sample used in performance evaluation studies, which includes thousands of units (stocks). In most applications, a cross-section available to researchers includes no more than few dozen assets, as with US commodity or treasury futures. Thus, the stability of correlation estimates in samples of different sizes is a desired feature, since inference over the performance of spread estimators in the evaluation sample implies out-of-sample extrapolation.

To do so, Figure (A.1) repeats the following experiment under three different settings. We draw 100 random samples of 10 stocks from the full DTAQ dataset, calculating the cross-section correlation between monthly HL estimates and effective spreads in each sample. We repeat the process for another 100 random draws with 100 and 1,000 stocks each time. The variability in performance in samples of 10 stocks is remarkable: the high-low estimator generates both exceptional cross-sectional correlations upwards of 80% and no correlation in other cases. Larger sample draws lead to more stable average cross-sectional correlation estimates, quickly converging to the full sample correlation of about 62%.

Liquidity heterogeneity in the sample. A robust finding for *all* moment-based bid-ask spread estimators that rely on a combination of closing, high, and low prices is a much poorer performance in the time-series compared to the cross-section. The larger the number of assets relative to the number of periods (usually months), the wider this gap. Although we are not interested in the reasons for this difference in performance, we exploit the fact that spread size heterogeneity is much higher between units than within to demonstrate the importance of heterogeneity in spread levels in the application context. This also purges the analysis from potential time-invariant unobserved factors within each stock.

First, from Figure (2) in the main paper, we learn that the high-low estimator performs better when effective spreads are larger. From Figure (A.1), we also know that larger cross-sections have higher and more stable correlations, suggesting that the proxy requires large variations in the spread *level* to be more accurately estimated. Mechanically, larger cross-sections tend to have more variation in spread levels



A researcher with a random sample of 10 US stocks would obtain very distinct average cross-sectional correlations between monthly high-low estimates and monthly effective spreads every time a new sample would be drawn. Correlations become stable across randomly chosen samples when the stock data includes more than 100 units.

Notes: This figure shows cross-sectional correlations between the high-low spread estimator (using the zeros adjustment) and the effective spreads from 100 random samples of 10, 100, and 1,000 stocks from DTAQ US stock data. There are 3,398 stocks in total and sample construction details are described in Appendix (I).

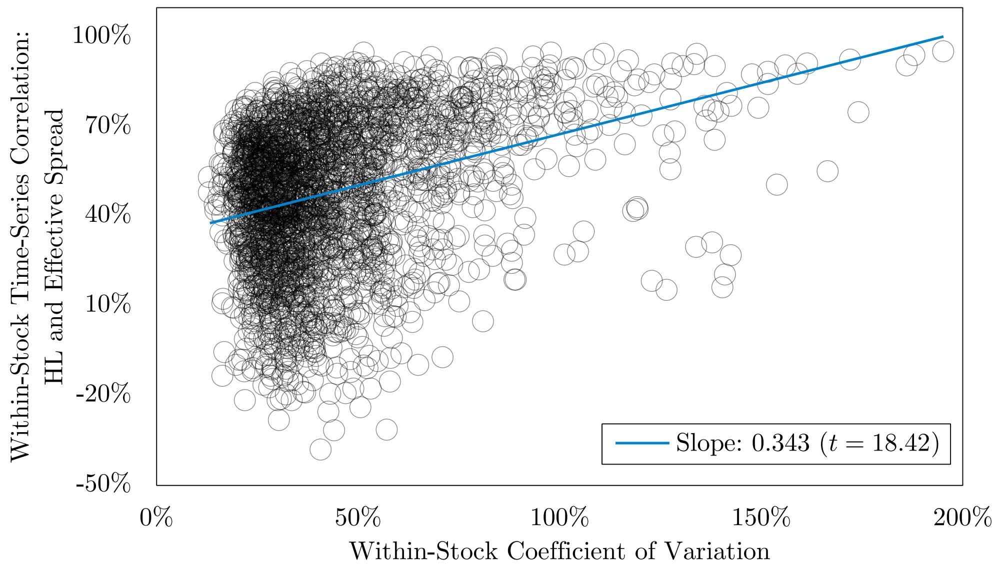
Figure A.1: Cross-Sectional Correlation: Stability Across Different Sample Sizes

than spread levels do when varying over time for the same stock, which might offer one explanation as to why performance is worse in the time series.

We test that below. Figure (A.2) plots within-stock average monthly time-series correlations between the high-low estimator and effective spreads against within-stock coefficients of variation of monthly effective spreads (standard deviation of spreads divided by the sample mean across months). The strong positive estimated slope confirms that the spread measure performs better in stocks with more dispersion in underlying effective spreads. Applying this rationale to the full cross-section generates an important result. Samples of assets with relatively homogeneous trading costs are likely to result in poor accuracy for the high low estimator.

Suggestions for performance validation. The exercises we perform in this section are simple and would greatly increase confidence in application of estimators across different markets and when benchmarking performance of new spread measures. Some suggestions for implementation are the following. When the cross-section is sufficiently large, researchers should conduct a systematic analysis of the performance of the estimator in subsamples with different levels of liquidity.

Most preferably, newly-developed estimators should address how the measure may suffer from systemic bias with respect to the latent spread, and empirically analyze the extent to which such de-



Stocks with greater dispersion in effective spreads are associated with higher time series correlation between monthly high-low estimates and the spread.

Notes: This figure compares within-stock dispersion in the spread (sample standard deviation divided by mean) with time series correlations between the high-low spread estimator (using the zeros adjustment) and effective spreads.There are 3,398 stocks in total from DTAQ US stock data and sample construction details are described in Appendix (I).

Figure A.2: HL Performance in the Time Series and Dispersion of Effective Spread

pendence may exist. They should also carefully report the stability of correlation coefficients in random subsamples from the full cross-section. This exercise confronts the fact that often times researchers in empirical finance may only have partial or sample data available. In such “small sample” applications, proxies with large variability of estimated spreads may yield somewhat spurious performance benchmarks, incorrectly concluding that the measure performs adequately in that context or market.

B Negative Spread Estimates: Determinants & Empirical Use

Why do high-low estimates turn out negative? Negative high-low spread estimates are inconsistent with theory. Like many of other simple spread estimators, negative (or indeterminate) daily computed spreads with the high-low estimator are pervasive. The standard approach is to consider these point estimates as "meaningless", either replacing them with zeros or discarding them. This implies that in many instances nearly half of the estimating sample is excluded from the analysis and with that, potential relevant information.

Negative spread estimates are not "random" or purely the result of model assumption failures. This is clear in Figure (B.1), which replicates the ideal simulated data we use in the main paper and computes the average frequency of negative high-low estimates by each effective spread bin. For small spreads the frequency of spread estimates that turn out negative is as high as 40%, even without any idiosyncratic source of bias. Based on that, we build on the following simple insight. If the intensity of negative estimates is a result of any aspect in the underlying data, we may use how often negative estimates arise to learn about the latent price process.

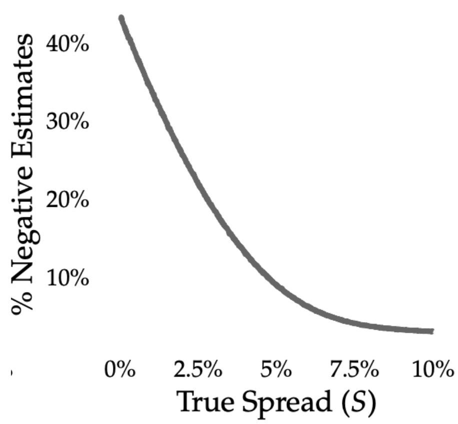


Figure B.1: High-Low Negative Estimates Under Ideal Conditions

To fix ideas, we begin with mechanical relationships that account for generated negative spread estimates, but that offer no insight on mechanisms. We do that next. The high-low estimator depends upon both the first and second-day ranges, r_t and r_{t+1}, and the two-day range r_t^*. Small perturbations at price boundaries have small effects on daily range values, but may alter the domain of the parameter γ and lead to considerable swings in the HL point-estimate. Let $t \in \{1, 2\}$. If changes in the second-day high, for example, are sufficiently large so that h_2 becomes higher than the previous day high, the mapping of γ shifts from h_1 to h_2, and the relationship between β and γ is updated. From the expression for α in the main paper or in Appendix (II), it is clear that the high-low estimator is positively defined only when $\beta > \gamma$.

Negative spread estimates do not arise only if the volatility-proportionality assumption in the high-low estimator is violated. Consider the following example. Suppose the first-day range is $r_1 = 0.01$, which is obtained when $(h_1, l_1) = (350, 345)$, and the second-day range is $r_2 = 0.27$, for $(h_2, l_2) = (401, 305)$. Clearly, the second day is much more volatile than the previous day, but the HL estimate is $0.09\% > 0$ over the two days. The simple explanation is that the square of the two-day range cannot be larger than itself added to the square of the first day range: $\gamma = r_2^2 < \beta = r_1^2 + r_2^2$. This happens because r_1 is enclosed in the second-day range. Note that it follows from this reasoning that the HL estimator can only be negative when the parameter γ combines boundaries from both consecutive days (i.e., either $h_1 - l_2$ or $h_2 - l_1$).

However, this is not sufficient for negativity. Consider the example on the left-hand side of Figure (B.2). With a partial overlap between r_t and r_{t+1}, the high-low estimator is not necessarily negative. Particularly, it is straightforward to show that some combination of high and low prices will produce a positive HL estimate when $l_{t+1} \in (l_t, h_t)$ and $h_{t+1} > h_t$. When we decrease the spread S by half (right-hand side panel), the new relative position of the second-day low, $l_{t+1} > h_t$, implies that the HL estimate is negative.

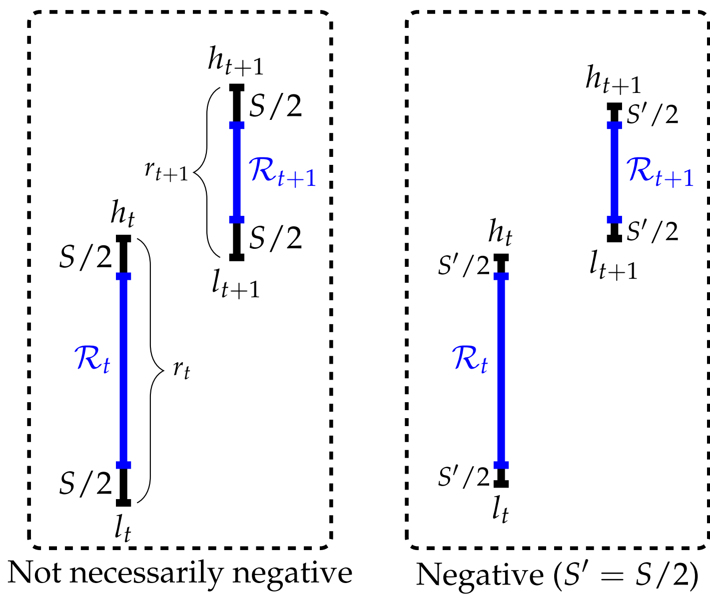


Figure B.2: Negativity Induced by the Spread Size in the HL Estimator

To organize these ideas, we derive two equivalent conditions that determine when the high-low spread is negative. While the first condition in Proposition B.1 is very intuitive and arises directly from the functional form of the implementable high-low measure, the second condition will be useful for us to prove the driving factors of negative estimates, in similar fashion to our treatment of the bias.

Proposition B.1. The high-low spread estimator is negative if and only if the following equivalent conditions hold:

$$\phi r_t^{min} < r_t^* \text{ or } |\eta_{t+1} - \eta_t| > \sqrt{r_t^2 + r_{t+1}^2} - \left(\frac{r_t + r_{t+1}}{2}\right)$$

where η_t is the (log) mid-range, $\eta_t \equiv \dfrac{h_t + l_t}{2}$, and r_t is the log range on day t.

Proof. The range of the high-low spread estimator $S(\alpha)$ only contains negative values when S is evaluated at negative values of α. Hence, negativity of α defines non-positive spread estimates. The parameter α is expressed as $\alpha = (1+\sqrt{2})(\sqrt{\beta} - \sqrt{\gamma})$, which implies that $\gamma > \beta$ determines when $\alpha < 0$. The parameter β always maps each day's range onto $\mathbb{R}_+$. Since γ includes extreme-valued functions, it may take the following values: $\gamma = r_t^2$, $\gamma = r_{t+1}^2$, $\gamma = (h_t - l_{t+1})^2$, or $\gamma = (h_{t+1} - l_t)^2$. In the first two cases, β is always greater than γ. For $\gamma = (h_{t+1} - l_t)^2$, we can write $\gamma > \beta$ as:

$$\underbrace{h_{t+1} - \frac{r_{t+1}}{2}}_{\eta_{t+1}} - \underbrace{\left(\frac{r_t}{2} + l_t\right)}_{\eta_t} + \frac{r_t}{2} + \frac{r_{t+1}}{2} > \sqrt{\beta} \tag{1}$$

and similarly, for $\gamma = (h_t - l_{t+1})^2$:

$$\underbrace{h_t - \frac{r_t}{2}}_{\eta_t} - \underbrace{\left(\frac{r_{t+1}}{2} + l_{t+1}\right)}_{\eta_{t+1}} + \frac{r_t}{2} + \frac{r_{t+1}}{2} > \sqrt{\beta} \tag{2}$$

so that (1) and (2) combined can be stated as:

$$|\eta_{t+1} - \eta_t| > \sqrt{\beta} - \left(\frac{r_t + r_{t+1}}{2}\right) \tag{3}$$

Note that the right-hand side of the negativity condition is always positive, since

$$\left(r_t^2 - r_{t+1}^2\right)^2 + 2\left(r_t^2 + r_{t+1}^2\right) > 0$$

which completes the proof of the proposition.

Generality. Throughout we assumed no consecutive-day price ties and well-defined daily ranges. Here, we relax this assumption and show that any price tie yields strictly positive spread values.

Identical consecutive-day low prices Let $h_t > h_{t+1}$ (w.l.o.g.) and $l_t = l_{t+1} \equiv a$. Given the definitions for β and γ, and the negativity condition slightly restated as $\gamma \geq \beta$, we have:

$$(h_t - a)^2 \geq (h_t - a)^2 + (h_{t+1} - a)^2 \tag{4}$$

which cannot hold since $h_{t+1} > a$.

Identical consecutive-day high prices Consider when $h_t = h_{t+1} = b$ and $l_t < l_{t+1}$ (w.l.o.g.). The negativity conditions becomes:

$$(b - l_t)^2 \geq (b - l_t)^2 + (b - l_{t+1})^2 \tag{5}$$

which is unfeasible since $b > l_{t+1}$.

Identical consecutive-day low and high prices Now, let $l_t = l_{t+1} = a$ and $h_t = h_{t+1} = b$. It is straightforward to see that $\gamma = 2\beta$, $\beta > 0$. Therefore, the spread is always strictly positive. ■

Using Proposition B.1, we now show that the negativity condition is more easily attained when the underlying spread is smaller or volatility is higher.

PROPOSITION B.2. Let $f \equiv f(\eta_t, \eta_{t+1})$ and $g \equiv g(r_t, r_{t+1})$ be functions that define the left-hand and right-hand sides, respectively, of the negativity condition for the high-low spread estimator $\widehat{S}^{HL}$:

$$f(\eta_t, \eta_{t+1}) > g(r_t, r_{t+1})$$

For *ex-ante* changes in the bid-ask spread level S and variance σ^2, the following relationships hold:

$$\frac{\partial f}{\partial S} = 0, \quad \frac{\partial g}{\partial S} > 0, \quad \frac{\partial E[f]}{\partial \sigma^2} = (2 - 2\ln 2), \quad \frac{\partial E[g]}{\partial \sigma^2} > 0, \quad \text{and} \quad \frac{\partial^2 E[g]}{\partial \sigma^2 \partial S} > 0$$

Proof. Because the efficient mid-range is identical to the observed mid-range, the left-hand side of the inequality does not depend on the spread level. That is, volatility estimated with squared returns of mid-prices is independent of the spread level. Thus $\partial f / \partial S = 0$ follows. On the other hand, the right-hand side of the inequality is a function of the observed range, $g(r_t, r_{t+1})$, which depends on the spread. If we rewrite observed ranges in terms of efficient ranges, the modified function g depends on S in the following way:

$$g \equiv \sqrt{(\mathcal{R}_t + S)^2 + (\mathcal{R}_{t+1} + S)^2} - \left(\frac{\mathcal{R}_t + \mathcal{R}_{t+1} + 2S}{2}\right)$$

and therefore

$$\frac{\partial g}{\partial S} = \frac{(\mathcal{R}_t + S) + (\mathcal{R}_{t+1} + S)}{\sqrt{(\mathcal{R}_t + S)^2 + (\mathcal{R}_{t+1} + S)^2}} - 1 > 0$$

for well-defined efficient ranges. That implies that the right-hand side in the negativity condition increases in the spread; hence as the spread widens, $\sqrt{\beta}$ relatively exceeds $(r_t + r_{t+1})/2$, so that the inequality becomes more difficult to be attained. For relatively greater spreads, the HL estimator is more likely to be positive.

The second part of Proposition B.2 is also simple. Denote $r_{t+1} = \kappa r_t$, $\kappa > 0$, but not necessarily bounded by 1. After squaring both sides of the negativity condition with the proper substitutions for r_{t+1}, we have

$$(\eta_{t+1} - \eta_t)^2 > \frac{1}{4}\left(\kappa - 2\sqrt{\kappa^2 + 1} + 1\right)^2 r_t^2$$

for which we can replace the observed range with the efficient range. To simplify the calculations, assume $\overline{\kappa} = 1$. Under the maintained hypothesis of constant volatility, daily observed range values will be very close, and κ will differ from the unity when the spread is very small. The average value of κ in 210,000 days of simulated data when $\sigma = 3\%$ and with an infinitesimal spread is 1.09. The simplification is therefore empirically consistent with historical stock data. With $\overline{\kappa} = 1$, the term multiplying r_t^2 collapses to $(3 - 2\sqrt{2})$ and after taking expectations of both sides, we arrive at

$$E\left[(\eta_{t+1} - \eta_t)^2\right] > (3 - 2\sqrt{2})E\left[\mathcal{R}_t^2\right] + 2S(3 - 2\sqrt{2})E\left[\mathcal{R}_t\right] + (3 - 2\sqrt{2})S^2.$$

Direct substitutions for the moments above using (4) and (11) yield

$$\left(2 - \frac{k_2}{2}\right)\sigma^2 > (3 - 2\sqrt{2})k_2\sigma^2 + 2(3 - 2\sqrt{2})k_1 S\sqrt{\sigma^2} + (3 - 2\sqrt{2})S^2. \tag{6}$$

Since the left-hand side of (6) is now $E\left[f\right]$, we have $\partial E\left[f\right]/\partial\sigma^2 > 0$. Similarly, the right-hand side is given by $E[g]$; hence $\partial E[g]/\partial\sigma^2 > 0$ follows. The growth rate of $E[g]$ in σ^2 may exceed $2 - k_2/2$, which largely depends on the magnitude of S. Since $\partial^2 E[g]/\partial\sigma^2\partial S > 0$, smaller spreads will reduce $E[g]$ even while volatility increases, so that the net effect of volatility depends on the relative size between S and σ^2. This completes the proof of the proposition.

Varying κ. Without setting $\kappa = 1$, we can replace $r_t = \mathcal{R}_t + S$ in $(\eta_{t+1} - \eta_t)^2 > \frac{1}{4} r_t^2\left[(\kappa + 1) - 2\sqrt{1 + \kappa^2}\right]^2$ and manipulate it as:

$$(\eta_{t+1} - \eta_t)^2 > \mathcal{R}_t^2\Psi(\kappa) + 2\mathcal{R}_t S\Psi(\kappa) + S^2\Psi(\kappa) \tag{7}$$

where $\Psi(\kappa) \equiv \frac{1}{4}\left(\kappa - 2\sqrt{\kappa^2 + 1} + 1\right)^2$. We can then take expectations of both sides and substitute $E[(\eta_{t+1} - \eta_t)^2] \approx 0.61\sigma^2$, $E[\mathcal{R}_t^2] = 4\ln 2\sigma^2$ and $E[\mathcal{R}_t] = (\sqrt{8/\pi})\sigma$:

$$0.61\sigma^2 > \sigma^2\Psi(\kappa)4\ln 2 + 2S\Psi(\kappa)\left(\sqrt{\frac{8}{\pi}}\right)\sqrt{\sigma^2} + \Psi(\kappa)S^2. \tag{8}$$

We assume $\overline{\kappa} = 1$ to make (8) more tractable. For $\kappa \in [0.8, 1.2]$, the function $\Psi(\kappa)$ varies from 0.15 to 0.21, with $\Psi(1) = 0.17$. After plugging 0.17 into the expression above we have

$$0.61\sigma^2 > 0.17\left(2.77\sigma^2 + 3.19S\sqrt{\sigma^2} + S^2\right) \tag{9}$$

which clearly yields the first-order derivatives shown in the proposition. ■

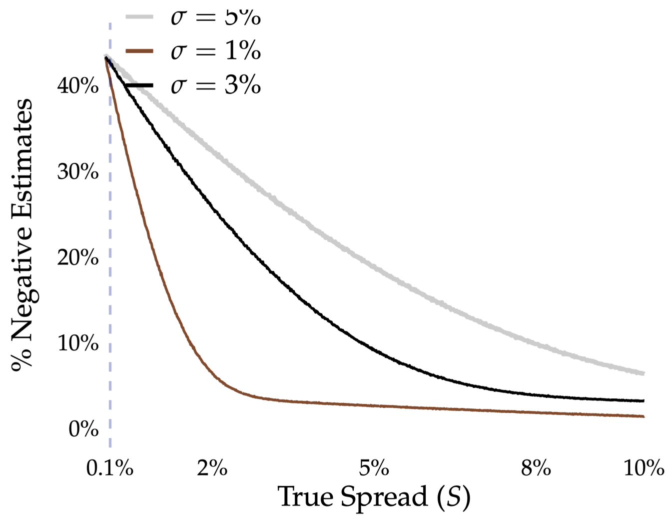
We let two arbitrary functions f and g represent the inequality condition for the HL estimator to make the proposition more intuitive. The left-hand side f contains only mid-range prices, η, while the right-hand side g contains only ranges r. Mid-range prices are independent of the spread, since $\eta_t \equiv (h_t + l_t)/2 = (H_t + L_t)/2$. The greater f is compared to g, the easier the negativity condition is attained. In turn, decreasing the spread makes it easier for the negativity condition to hold and therefore generate a negative estimate.

While greater volatility also increases the frequency of negative estimates, even if volatility approached zero, a small spread would still generate a large number of non-positive spreads. That is because increasing volatility has an ambiguous effect on the proportion of negative estimates, as $E[g]$ may or may not exceed $E[f]$. Thus, S regulates the relative contribution of σ^2 in $E[g]$. This implies that ultimately the spread is more important than volatility to determine when the high-low estimator is negative.

To see this point formally, because S acts as a limiting factor to the growth of $E[g]$ with respect to σ^2, we should expect a high proportion of negative spreads for a narrow spread, regardless of the volatility level. In the limit, as $S \to 0$, the right-hand side of (6) converges to $(3 - 2\sqrt{2})k_2\sigma^2$, which is the minimum of $E[g]$ with respect to S and smaller than $2 - k_2/2$. In the theoretical case, the expected frequency of negative spreads is 100%. In practice, either with simulated or actual data that can only be observed at a discrete set of times, the proportion of negative estimates is maximized when the underlying asset is very liquid. In Figure (B.3), when the true spread is 0.1%, increasing volatility from 1% to 5% produces almost no impact in the proportion of negative HL estimates.

Observable differences in the frequency of negative estimates resulting from varying σ^2 start to appear as the spread grows. Locally widening the spread enables volatility to contribute toward the growth rate of $E[g]$. The difference $E[f] - E[g]$ increases in σ^2 up to a spread level where $E[g]$ is greater than $E[f]$. At this point, the frequency of negative estimates is minimized, and different levels of volatility are again redundant to (6).

Figure (B.3) illustrates the proposition's prediction, where the frequency of negative estimates in the ideal simulated setting for each spread level is plotted using three different price volatility levels. The monotonicity given in Proposition C2 clearly holds up. While the negative share with a daily volatility of 1% decays faster than shares at higher volatility levels, when spreads are around the median value of US stocks, the frequency of negatives is about the same, around 40%.



For very small spreads (below 0.1%), the frequency of negative spread estimates is always greater than 40%, regardless of price volatility. For reference, the median effective spread in US stocks during 2003-2015 is 0.3%.

Notes: This figure computes the average frequency of negative high-low estimates over 10,000 trading months for each 0.1-percentage-point increment in the spread between 0.1% and 10% from simulated data. The data generating process is described in detail in the text and is designed to maintain all model assumptions, including constant true spreads in each sample. We report the relationship with three different volatility values.

Figure B.3: Relationship Between Spread Size, % of Negative Estimates Under Ideal Conditions

C Supplementary Analysis: Close-High-Low Estimator

This section replicates most of our main results for the high-low and roll estimators using the close-high-low (CHL) measure from Abdi and Ranaldo (2017). We also consider this effective spread proxy because in the original paper, Abdi and Ranaldo (2017) show results, especially in the cross-section, that suggest a slight edge of the measure over the high-low estimator. Unfortunately, Figures (D.1) and (D.4) show that the estimator suffers the same significant decreases in performance in the cross-section of monthly, annual, and changes in effective spreads as the high-low proxy. Performance in the entire cross-section is significantly driven by few large-spread stocks, with correlation between the measure and the benchmark spread decreasing as the effective spread narrows. Indeed, the patters are remarkably similar to the Roll and high-low estimators.

Although we do not offer formal proofs for the systematic biases we study in this paper - the moment and small sample bias - we show that the empirical and simulated results implied by our previous derivations apply almost identically to the close-high-low estimator. Furthermore, we present a complete analysis of the estimator's negativity determinants, which reveal the same driving forces behind degenerate estimates in this case as with the high-low proxy. Those are smaller spreads and higher volatility, with spreads contributing more significantly to generating negative spreads. Simulation results show that the bias behaves in the same fashion as with the high-low estimator. All in all, our results confirm the incidence of the same misbehavior patterns and driving forces for the CHL measure as in the HL and Roll estimators.

The estimator. The close-high-low spread estimator from Abdi and Ranaldo (2017) combines the use of high and low prices from the HL measure with daily closing prices from Roll (1984). In a sense, the measure uses the Parkinson-Garman-Klauss framework more broadly by integrating the range and daily returns to estimate bid-ask spreads along the lines of HL. The intuition for augmenting the information set is that closing prices are more "contaminated" by the bid-ask bounce than the range (Alizadeh et al. (2002)). Let the mid-range, or average price on day t, be defined as

$$\eta_t \equiv \frac{h_t + l_t}{2} = \frac{r_t}{2} + l_t \tag{10}$$

which clearly coincides with the mid-range of daily efficient extreme prices, $\eta_t = (\mathcal{H}_t + \mathcal{L}_t)/2$. A crucial result in Abdi and Ranaldo (2017) establishes how the variance of the mid-range returns relates to the efficient price volatility σ^2:

$$E\left[(\eta_{t+1} - \eta_t)^2\right] = \left(2 - \frac{k_2}{2}\right)\sigma^2 \tag{11}$$

where the variance may be replaced with $\widehat{\sigma}_{\eta}^{2}$ for estimated squared returns of consecutive-day mid-range prices. Under the validity of (11), and because the average of consecutive mid-ranges,

$$S^{CHL} = 2\sqrt{\left(c_t - \eta_t\right)\left(c_t - \eta_{t+1}\right)}. \tag{12}$$

Negativity of the close-high-low estimator. We follow the same steps for the close-high-low estimator to prove the drivers of negative estimates as we did with the high-low proxy. Even though the CHL measure does not directly depend on the range, it does depend on the mid-range. Moreover, our simulation results show a behavior dynamic of CHL similar to HL with respect to negative spreads, true spread size, and bias Figure (C.2). This suggests a common channel through which the spread-to-volatility ratio affects both proxies. From the expression for CHL_t, the straightforward negativity condition of the close-high-low estimator is:

$$c_t \in \left(\eta_t, \eta_{t+1}\right) \tag{13}$$

when the second-day mid-range is above the first-day mid-range, and $c_t \in \left(\eta_{t+1}, \eta_t\right)$ in the symmetric case. When the price variation across two days is large enough (measured by the difference in mid-prices), a wide range of values for c_t yields negative CHL spread estimates. We can rewrite (13) as (by analogy the symmetric case follows)

$$\eta_{t+1} - \eta_t > c_t - \eta_t \tag{14}$$

and similarly to HL, define an inequality condition to investigate the effects of spread and volatility levels in generating negative spread estimates.

Proposition C.1 Let $f \equiv f\left(\eta_t, \eta_{t+1}\right)$ and $v \equiv v\left(c_t, \eta_t\right)$ be functions that define the left-hand and right-hand sides, respectively, of the negativity condition for the close-high-low spread estimator:

$$f\left(\eta_t, \eta_{t+1}\right) > v\left(c_t, \eta_t\right) \tag{15}$$

For *ex-ante* changes in the bid-ask spread level S and variance σ^2, the following relationships hold:

$$\frac{\partial f}{\partial S} = 0, \quad \frac{\partial E[v]}{\partial S} > 0, \quad \frac{\partial E[f]}{\partial \sigma^2} = k_3, \quad \frac{\partial E[v]}{\partial \sigma^2} = \frac{k_3}{2} \quad \text{and} \quad \frac{\partial^2 E[v]}{\partial \sigma^2 \partial S} = 0. \tag{16}$$

Proof. The close-high-low spread estimator is negatively defined if the following holds:

$$\eta_{t+1} > c_t > \eta_t \tag{17}$$

or, equivalently,

$$\eta_{t+1} - \eta_t > c_t - \eta_t > 0. \tag{18}$$

We can replace observed mid-ranges and close prices with true values:

$$\eta_{t+1} - \eta_t > \mathcal{C}_t + q_t \frac{S}{2} - \left(\frac{\mathcal{H}_t + \mathcal{L}_t}{2} \right) \tag{19}$$

and square both sides such that

$$(\eta_{t+1} - \eta_t)^2 > \left[\left(q_t \frac{S}{2} + \frac{\mathcal{C}_t}{2} - \frac{\mathcal{H}_t}{2} \right) + \left(\frac{\mathcal{C}_t}{2} - \frac{\mathcal{L}_t}{2} \right) \right]^2 \tag{20}$$

and further simplify it as

$$\begin{aligned}(\eta_{t+1} - \eta_t)^2 >& \frac{1}{4} \left[q_t^2 S^2 + 2 q_t S \left(\mathcal{C}_t - \mathcal{H}_t \right) + \left(\mathcal{C}_t - \mathcal{H}_t \right)^2 \right] + \\ &+ \frac{1}{2} \left[q_t S \mathcal{C}_t + \mathcal{C}_t^2 - q_t S \mathcal{L}_t + \mathcal{C}_t \mathcal{L}_t - \mathcal{C}_t \mathcal{H}_t + \mathcal{H}_t \mathcal{L}_t \right] + \frac{1}{4} \left(\mathcal{C}_t - \mathcal{L}_t \right)^2 .\end{aligned}$$

After taking expectations of the above, we arrive at the intermediate expression

$$E\left[(\eta_{t+1} - \eta_t)^2 \right] > \frac{S^2}{4} + \frac{1}{4} E\left[\left(\mathcal{C}_t - \mathcal{H}_t \right)^2 \right] + \frac{1}{2} \left(E\left[\mathcal{C}_t \mathcal{L}_t \right] - E\left[\mathcal{C}_t \mathcal{H}_t \right] + E\left[\mathcal{H}_t \mathcal{L}_t \right] \right) + \frac{1}{4} E\left[\left(\mathcal{C}_t - \mathcal{L}_t \right)^2 \right] \tag{21}$$

which can be easily solved for with the moments given in Garman and Klass (1980). The final inequality is

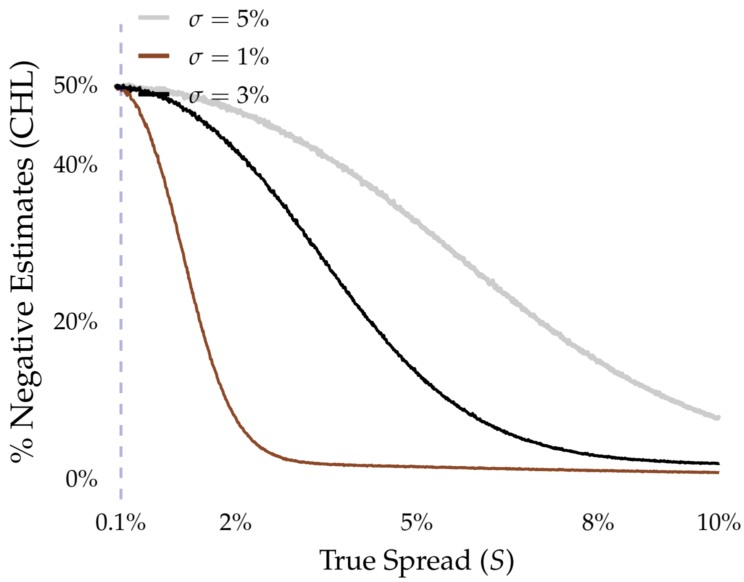
$$(2 - 2 \ln 2)\, \sigma^2 > \frac{S^2}{4} + \frac{\sigma^2}{4} + \frac{1}{2} (1 - 2 \ln 2)\, \sigma^2 + \frac{\sigma^2}{4} \tag{22}$$

and therefore:

$$(2 - 2 \ln 2)\, \sigma^2 > \frac{S^2}{4} + (1 - \ln 2)\, \sigma^2. \tag{23}$$

■

Proposition C.1 is intended to mimic the structure used in Proposition B.2. The generic function f is intentionally identical to f in the HL estimator. The proof, however, is slightly different from the proposition in the high-low case. Since we do not assume the value of the trade indicator q_t in $c_t = \mathcal{C}_t + q_t S/2$, we can only work with the expected value of v, without assuming a deterministic q_t. We also use moments derived for $E\left[\mathcal{C}^a \mathcal{H}^b \mathcal{L}^c\right]$ in Garman and Klass (1980).



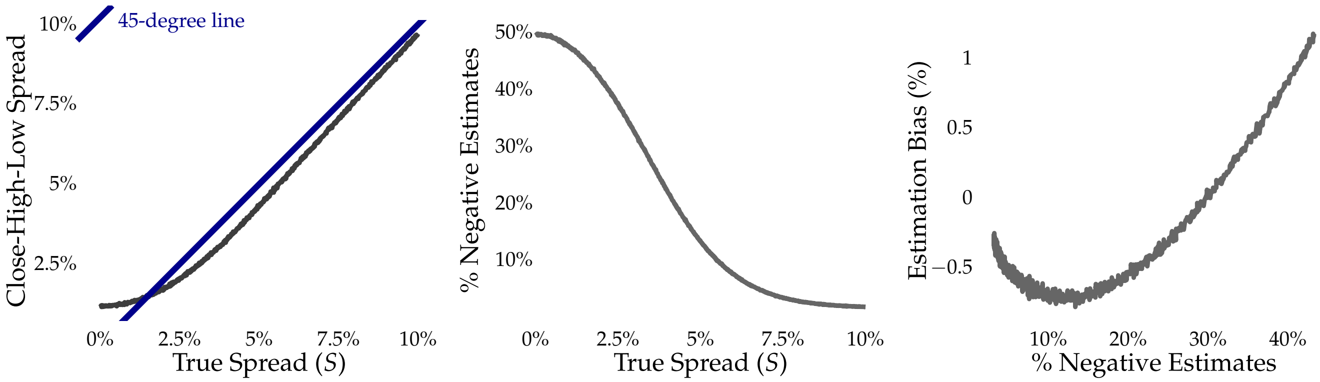
For very small spreads (below 0.1%), the frequency of negative close-high-low spread estimates is always greater than 50%, regardless of price volatility. For reference, the median effective spread in US stocks during 2003-2015 is 0.3%.

Notes: This figure computes the average frequency of negative close-high-low estimates over 10,000 trading months for each 0.1-percentage-point increment in the spread between 0.1% and 10% from simulated data. The data generating process is described in detail in the text, being identical to the DGP used in the exercises for the high-low estimator. Moreover, it is also designed to maintain all model assumptions, including constant true spreads in each sample. We report the relationship with three different volatility values.

Figure C.1: Relationship Between Spread Size, % of Negative Estimates Under Ideal Conditions

The interpretation of Proposition C.1 is also very similar to Proposition B.2, although simpler. Decreasing the spread size lowers the expected value of the right-hand side of the negativity condition, contributing to more negative estimates. When the spread is small, increases in σ^2 cannot induce $E[f] < E[v]$, which explains why in Figure (C.1) the proportion of negative CHL estimates is around 50% when $S = 0.1\%$, regardless of volatility size. Note that the figure shows the same monotonic patterns as with the high-low measure, revealing an even higher proportion of negative estimates in the close-high-low case for most spread sizes.

Moreover, for small spreads, any value of σ^2 increases $E[f]$ faster than $E[v]$ – a pattern that is eventually reversed when the spread is large enough so that the right-hand side becomes greater than $E[f]$. For completeness, we also provide the following alternative result that shows why smaller spreads induce negative CHL estimates more often.



For small spreads (below 1%), the close-high-low estimator suffers from an upward bias and a higher fraction of negative estimates. As the spread becomes larger, the bias decreases and fewer estimates turn out negative. For reference, the median effective spread in US stocks during 2003-2015 is 0.3%.

Notes: This figure computes 10,000 monthly averages of the close-high-low spread estimator (with zeros adjustment) and share of negative daily estimates for each 0.1-percentage-point increment in the spread between 0.1% and 10% from simulated data. The data generating process is described in detail in the text and is designed to maintain all model assumptions, including constant true spreads in each sample. Bias is defined by the difference between the estimated spread and the true spread.

Figure C.2: Bias, Negative Estimates, and True Spread Under Ideal Conditions

PROPOSITION C.2. The probability that c_t falls in the interval (η_t, η_{t+1}), and therefore the negativity condition for CHL is attained, increases as the spread size decreases.

Proof. The probability that the negativity condition of CHL is attained, $\Pr\left[\eta_t < c_t < \eta_{t+1}\right]$, is given by

$$\int_{\eta_t}^{\eta_{t+1}} f\left(c_t\right) dc_t = \int_{l_t}^{h_t} f\left(c_t\right) dc_t - \underbrace{\left(\int_{\eta_{t+1}}^{h_t} f\left(c_t\right) dc_t + \int_{l_t}^{\eta_t} f\left(c_t\right) dc_t\right)}_{\mathcal{K}}.$$

Let $h_t^* = h_t - \delta$ and $l_t^* = l_t + \delta$ represent modified daily high and low prices from a decrease of 2δ in the spread S. Let $\Pr\left[\eta_t^* < c_t < \eta_{t+1}^*\right]$ denote the probability that c_t falls within the negativity interval *ex post* the spread decrease. Then, it follows that

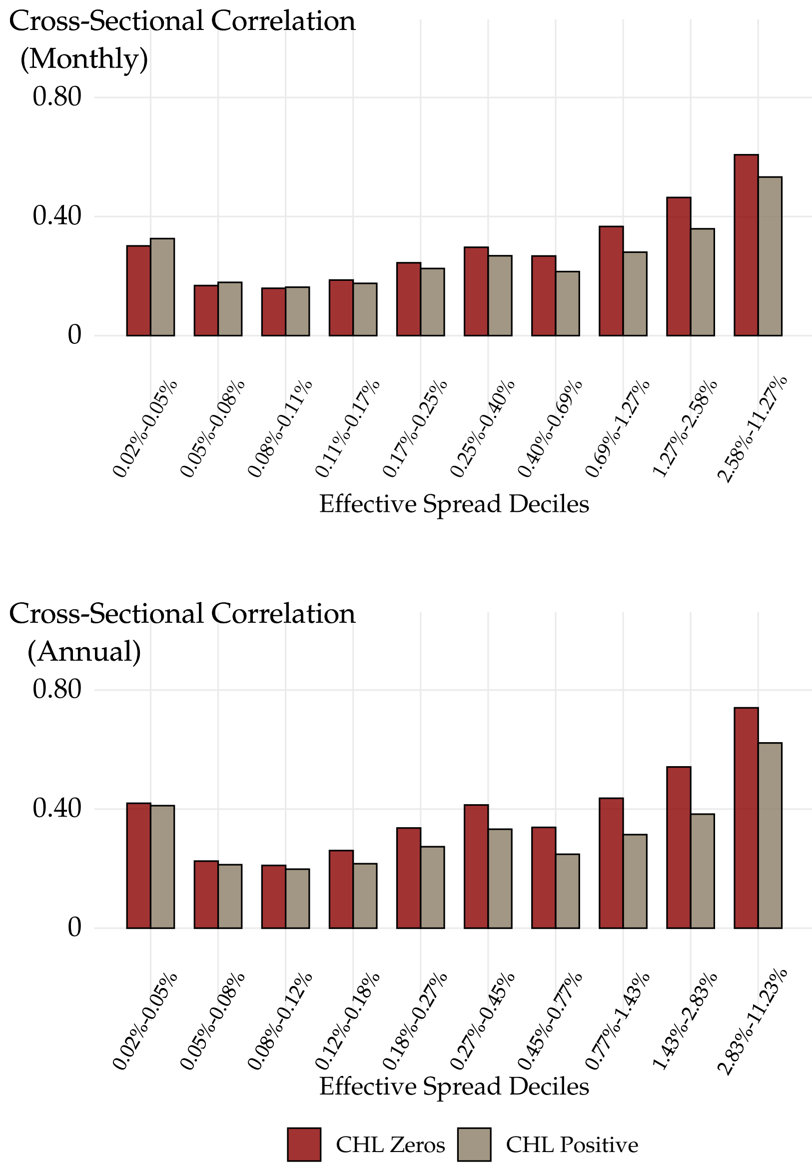
$$\left(\int_{\eta_{t+1}}^{h_t} f\left(c_t\right) dc_t + \int_{l_t}^{\eta_t} f\left(c_t\right) dc_t\right) \geq \left(\int_{\eta_{t+1}}^{h_t^*} f\left(c_t\right) dc_t + \int_{l_t^*}^{\eta_t} f\left(c_t\right) dc_t\right)$$

since

$$\left(\int_{\eta_{t+1}}^{h_t^*} f\left(c_t\right) dc_t + \int_{l_t^*}^{\eta_t} f\left(c_t\right) dc_t\right) \leq \left(\int_{h_t-\delta}^{h_t} f\left(c_t\right) dc_t + \int_{\eta_{t+1}}^{h_t-\delta} f\left(c_t\right) dc_t + \int_{l_t+\delta}^{\eta_t} f\left(c_t\right) dc_t + \int_{l_t}^{l_t+\delta} f\left(c_t\right) dc_t\right)$$

and therefore $\int_{\eta_t}^{\eta_{t+1}} f\left(c_t\right) dc_t \leq \int_{\eta_t^*}^{\eta_{t+1}^*} f\left(c_t\right) dc_t$. The equality holds if and only if $\mathcal{K} = 0$. This can be the case only when $\eta_t = \eta_{t+1}$. Hence, $\Pr\left[\eta_t^* < c_t < \eta_{t+1}^*\right] > \Pr\left[\eta_t < c_t < \eta_{t+1}\right]$ and we conclude the proof. ■

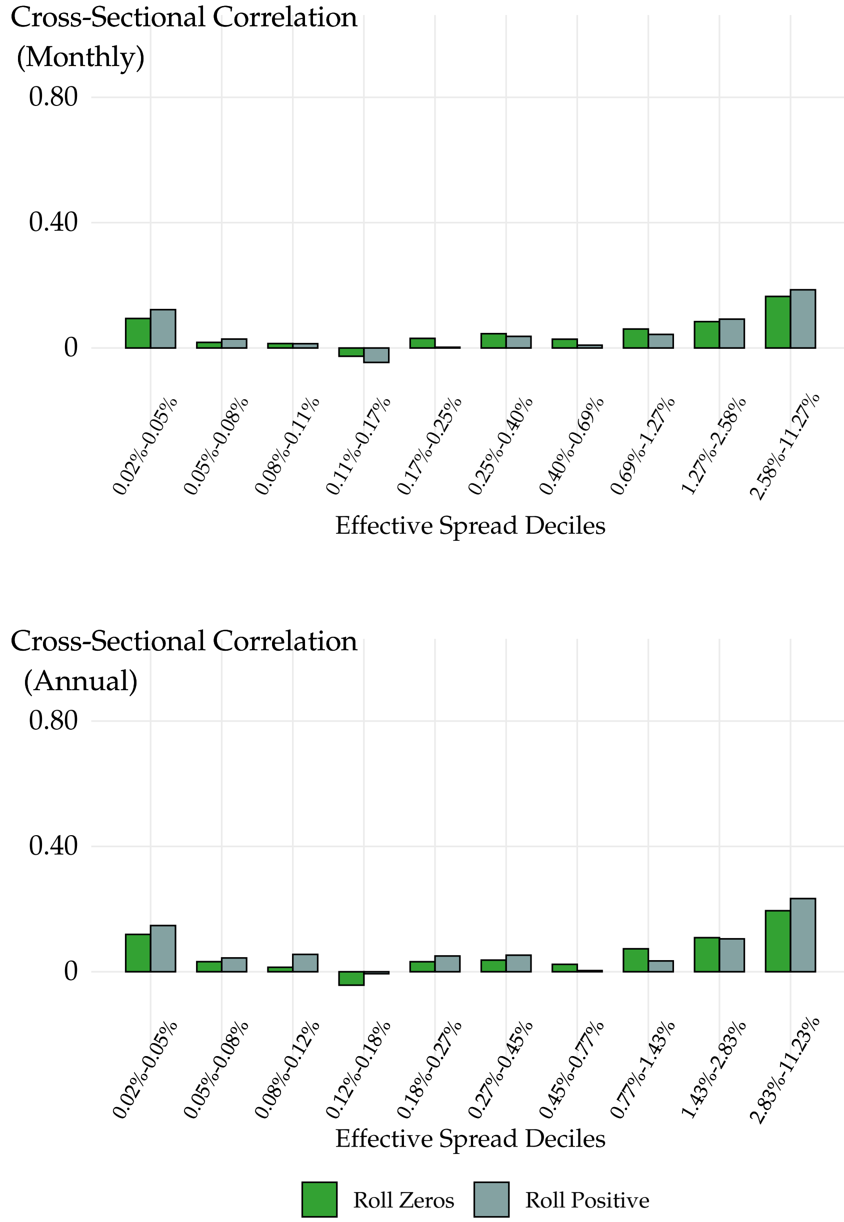
D Appendix Figures



In all effective spread size deciles, monthly averages of close-high-low estimates without ad hoc adjustments correlate poorly with monthly average effective spreads. Correlations from the estimates with the positives and zeros adjustments behave similarly, with the best performance in the least liquid stocks (above 50% correlation with effective spreads). The frequency of daily estimated spreads that turn out negative is above 50% in the most liquid stocks, decreasing as effective spread bands increase.

Notes: This figure shows cross-sectional correlations between three versions of the close-high-low spread estimator and effective spreads from DTAQ US stock data. There are 3,398 stocks in total and sample construction details are described in Appendix (I).

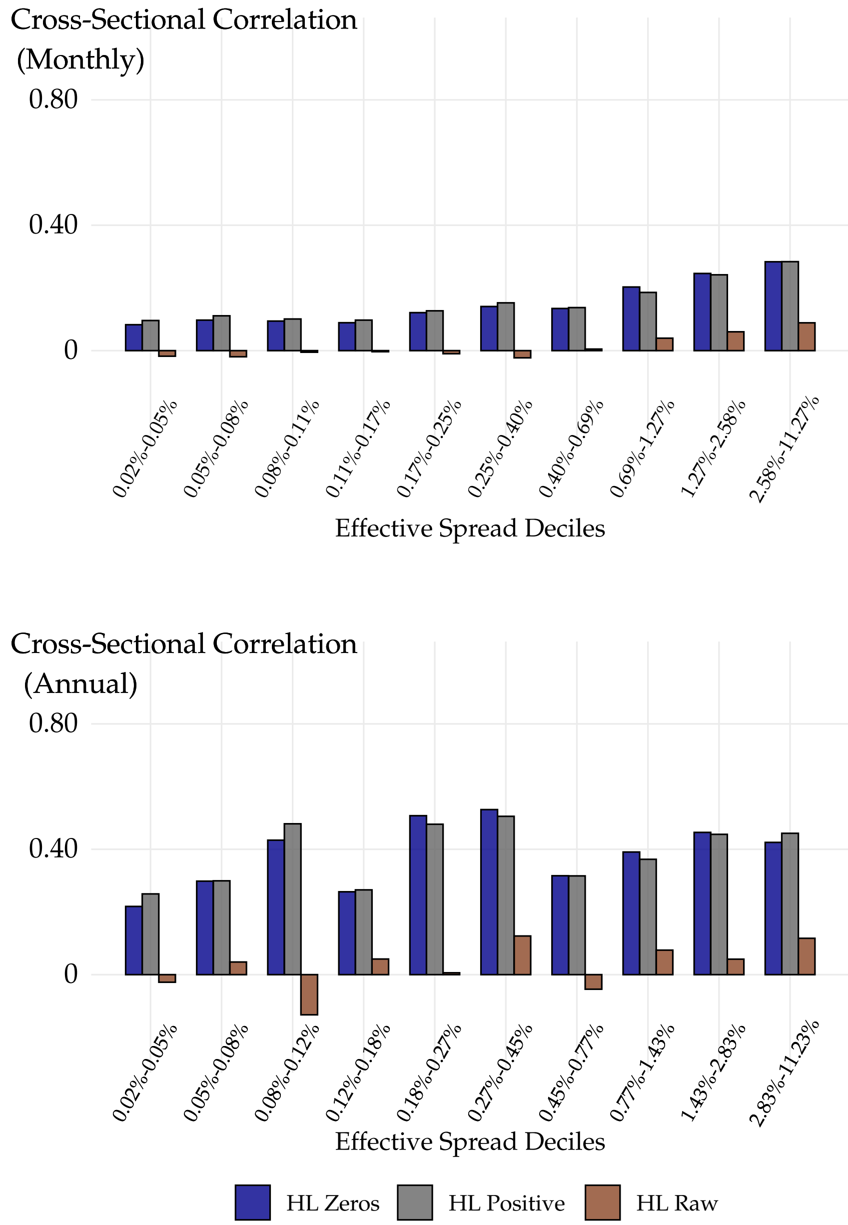
Figure D.1: Cross-Sectional Correlation: CHL Estimates and Effective Spread By Effective Spread Decile Size



In all effective spread size deciles, monthly averages of Roll estimates without ad hoc adjustments correlate poorly with monthly average effective spreads. Correlations from the estimates with the positives and zeros adjustments behave similarly, with the best performance in the least liquid stocks (above 15% correlation with effective spreads). The frequency of daily estimated spreads that turn out negative is above 40% in the most liquid stocks, decreasing as effective spread bands increase.

Notes: This figure shows cross-sectional correlations between three versions of the Roll spread estimator and effective spreads from DTAQ US stock data. There are 3,398 stocks in total and sample construction details are described in Appendix (I).

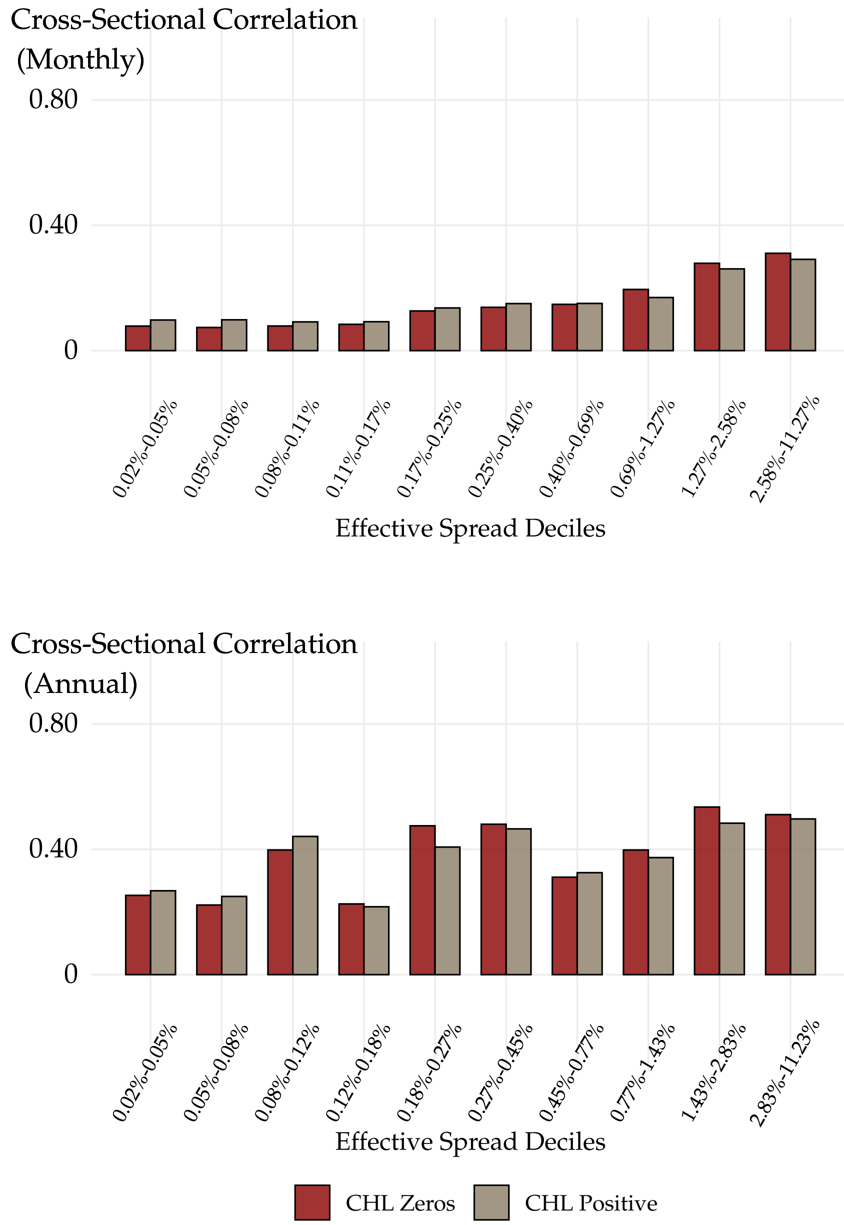
Figure D.2: Cross-Sectional Correlation: Roll Estimates and Effective Spread By Effective Spread Decile Size



In all effective spread size deciles, monthly averages of changes in high-low estimates without ad hoc adjustments correlate poorly with monthly average change in effective spreads. Correlations from the estimates with the positives and zeros adjustments behave similarly, with the best performance in the least liquid stocks (above 30% correlation with effective spreads). The frequency of daily estimated spreads that turn out negative is above 40% in the most liquid stocks, decreasing as effective spread bands increase.

Notes: This figure shows cross-sectional correlations between three versions of the close-high-low spread estimator and effective spreads from DTAQ US stock data. There are 3,398 stocks in total and sample construction details are described in Appendix (I).

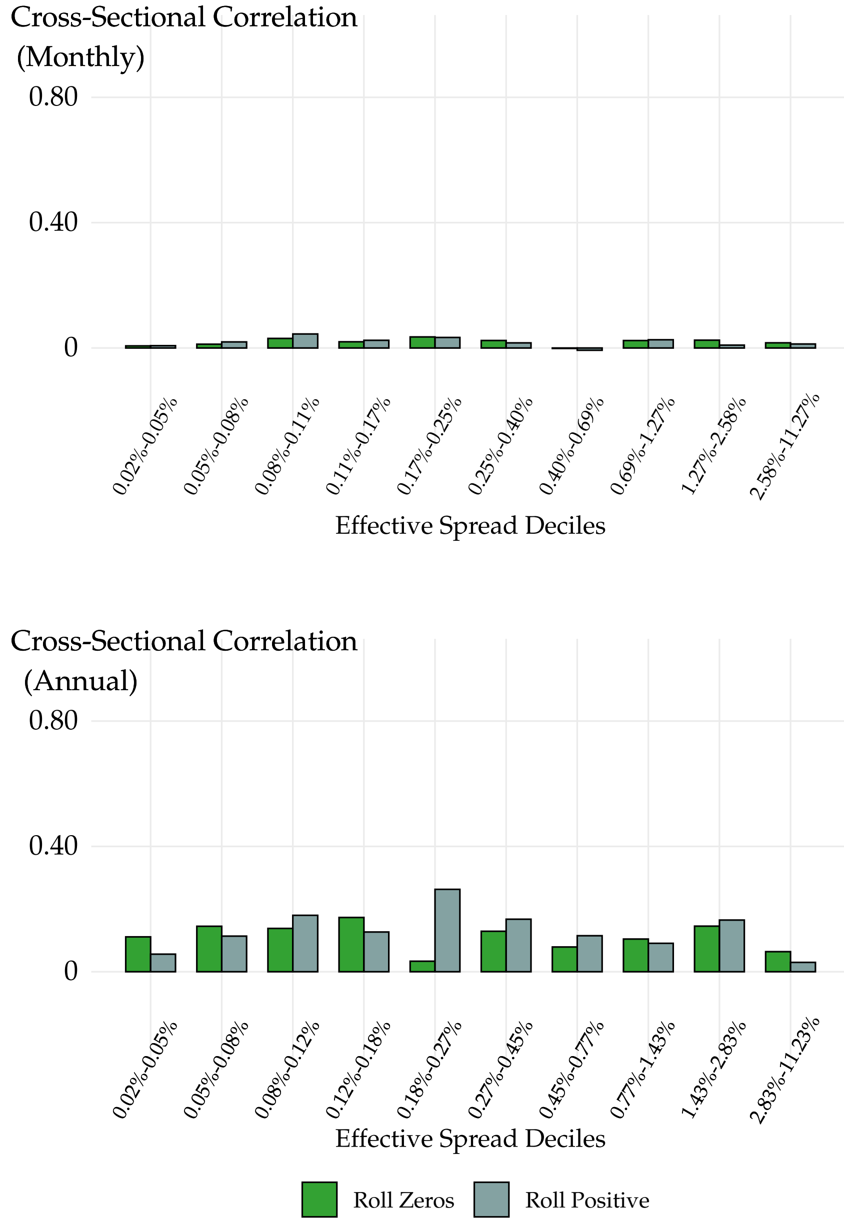
Figure D.3: Cross-Sectional Correlation: Changes in HL Estimates and Changes in Effective Spreads By Effective Spread Decile Size



In all effective spread size deciles, monthly averages of changes in close-high-low estimates without ad hoc adjustments correlate poorly with monthly changes of average effective spreads. Correlations from the estimates with the positives and zeros adjustments behave similarly, with the best performance in the least liquid stocks (above 30% correlation with effective spreads). The frequency of daily estimated spreads that turn out negative is above 50% in the most liquid stocks, decreasing as effective spread bands increase.

Notes: This figure shows cross-sectional correlations between three versions of the close-high-low spread estimator and effective spreads from DTAQ US stock data. There are 3,398 stocks in total and sample construction details are described in Appendix (I).

Figure D.4: Cross-Sectional Correlation: Changes in CHL Estimates and Effective Spreads By Effective Spread Decile Size



In all effective spread size deciles, monthly averages of changes in Roll estimates without ad hoc adjustments correlate poorly with monthly changes of average effective spreads. Correlations from the estimates with the positives and zeros adjustments behave similarly, with the best performance in the least liquid stocks (above 1% correlation with effective spreads). The frequency of daily estimated spreads that turn out negative is above 40% in the most liquid stocks, decreasing as effective spread bands increase.

Notes: This figure shows cross-sectional correlations between three versions of the Roll spread estimator and effective spreads from DTAQ US stock data. There are 3,398 stocks in total and sample construction details are described in Appendix (I).

Figure D.5: Cross-Sectional Correlation: Changes in Roll Estimates and Changes in Effective Spreads By Effective Spread Decile Size