The Contribution of Transaction Costs to Expected Stock Returns: A Novel Measure

Kazuhiro Hiraki, George Skiadopoulos

Working Paper No. 946 February 2023 ISSN 1473-0278

School of Economics and Finance

Queen Mary
University of London

The Contribution of Transaction Costs to Expected Stock Returns: A Novel Measure[*]

Kazuhiro Hiraki[†] George Skiadopoulos[‡]

December 17, 2022

Abstract

We document that a theoretically founded, real-time, and easy-to-implement option-based measure, termed *synthetic-stock difference* (SSD), accurately estimates the part of stock's expected return arising from stock's transaction costs. We calculate SSD for U.S. optionable stocks. SSD can be more than 10% per annum, it can fluctuate significantly over time and its cross-sectional dispersion widens over market crises periods. We confirm the accuracy of SSD by empirically verifying the predictions of a general asset pricing setting with transaction costs. First, we document its predicted type of connection with various proxies of stocks' transaction costs. Second, we conduct simple asset pricing tests which render further support. Our setting allows explaining the size of alphas reported by previous literature on the predictive ability of deviations from put-call parity.

JEL classification: C13, G10, G12, G13

Keywords: Transaction costs, Put-call parity, Return predictability, Informational content of options

[*]We are grateful to Ivo Welch for his valuable suggestions which improved the paper substantially. We thank Thang Ho for his excellent research assistance. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Bank of Japan. All remaining errors are the sole responsibility of the authors.

[†]Institute for Monetary and Economic Studies, Bank of Japan, 2-1-1 Nihonbashi-Hongokucho, Chuo-ku, Tokyo, 103-0021, Japan. kazuhiro.hiraki@boj.or.jp

[‡]School of Economics and Finance, Queen Mary University of London, Mile End Road, London, E1 4NS, UK, and Department of Banking and Financial Management, University of Piraeus, 80, Karaoli and Dimitriou Str, Piraeus, Greece. Also Senior Visiting Fellow at Bayes Business School (formerly Cass), City University of London. g.skiadopoulos@qmul.ac.uk and gskiado@unipi.gr

1 Introduction

Stock trading entails transaction costs.[1] A stock investor would require a compensation, in addition to the one for being exposed to risks, so that she also covers the incurred transaction costs. This simple rationale yields a decomposition of expected returns in two parts, the first reflecting the effect of transaction costs (termed *transaction costs alpha*, TC-alpha) and the second reflecting the effect of risk premia (the covariance risk premium term).[2]

Even though there is a large literature on the latter term, only few studies, including Amihud and Mendelson (1996), analyze how transaction costs contribute to expected stock returns.[3] This is because the estimation of the TC-alpha is not trivial; the size of the transaction costs does not suffice to determine the exact value of TC-alpha. For example, the sign of the TC-alpha depends on the investor's trading direction. In the case where the investor buys ((short-)sells) stocks, she may demand a greater (lower) expected return than the covariance risk premium to be profitable and compensate incurred transaction costs; this will yield a positive (negative) TC-alpha.

In this paper, we propose and empirically validate a novel measure to proxy the TC-alpha for any optionable stock. Our measure is based on properly scaled deviations from put-call parity, termed *synthetic-stock difference* (SSD hereafter).[4] SSD is a model-free, real-time, non-parametric and easy-to-implement measure. Its calculation requires only pairs of observable call and put option market prices with the same strike price and maturity and their underlying stock price. We calculate SSD for each optionable U.S. common stock from January 1996 to December 2020; these optionable stocks account on average for around 90% of the U.S. equity market capitalization. We find that SSD can become sizable when compared to the typical value of the U.S. equity risk premium; on some dates, SSD reaches more than 10% per annum. We also find that SSD is not stable and its value fluctuates month by month; this variation is significant for stocks which have big SSD (in absolute value). Moreover, the cross-sectional dispersion of SSD widens during periods of distress as expected, given that transaction costs magnify over crises.

[1] In this paper, transaction costs refer to typical trading costs (e.g., commissions, exchange fees, bid-ask spreads and market impact) that apply to buying and selling assets, as well as to short-sale costs.

[2] Our terminology of TC-alpha is analogous to that of Gârleanu and Pedersen (2011), who define the *margin alpha* as the expected excess return adjusted for risk (i.e., the covariance risk premium term) that arises due to the existence of margin constraints.

[3] Blume and Stambaugh (1983) and Asparouhova et al. (2010) study how microstructural noise affect the estimation of the equity risk premium.

[4] Intuitively, deviations from put-call parity stem from market frictions in the underlying stock and/or frictions in the option market. The source of the deviation (market and type of friction) can not be known in advance and it is a matter of empirical analysis. Our empirical result shows that SSD is an accurate proxy of TC-alpha, that is, SSD proxies effect of transaction costs in the underlying stock market.

We document that SSD is an accurate proxy of TC-alpha by providing two sets of results founded on the predictions of a general asset pricing setting with transaction costs. The first set documents that there is a particular type of connection between SSD and commonly used proxies of transaction costs for trading stocks in line with the theory. To establish this, as a prerequisite, we show that TC-alpha should lie in the interval $[-2\rho - \theta, 2\rho]$, where ρ is the relative size of the stock transaction cost and θ is the relative short-sale cost. This theoretically established TC-alpha interval shows that TC-alpha (and its empirical counterpart SSD) is not synonymous to ρ, yet it yields the following three predictions for how TC-alpha/SSD should be related to the transaction costs parameters. First, in the cross-section of stocks, the *value* of SSD should be non-linearly related to ρ; it would take more extreme values (either positive or negative) for stocks which are subject to greater transaction costs. Second, the *dispersion* of SSD should be positively and (approximately) linearly correlated with proxies of ρ. Third, the alpha of an SSD-sorted spread portfolio (big SDD stocks minus small SSD stocks) should be greater for stocks which face greater transaction costs. This is because a larger transaction cost parameter ρ and/or larger short-sale cost parameter θ translates to a wider range of TC-alpha.

Next, we confirm all three predictions empirically. We find that cross-sectionally, SSD has a U-shape relation with a list of low-frequency proxies of stocks' transaction costs, known to be highly correlated with their high-frequency counterparts. Regarding the second prediction, we employ a number of low-frequency proxies of transaction costs. We find that the average over time cross-sectional correlation between the monthly standard deviations of daily SSD and transaction costs is positive, for each proxy and it can reach a high value. For instance, the correlation with the VoV(%Spread) of Fong et al. (2017a) reaches 0.62.[5] With regard to the third prediction, we find that the alphas of the SSD-sorted spread portfolios increase as transaction costs increase.

The second set of results further verifies that SSD is an accurate proxy of TC-alpha. These are based on the predictions of a standard and general asset pricing formula which should hold for any asset pricing model in the presence of transaction costs, namely

$$\text{(Expected excess return)} = \alpha^{TC} + \text{(Covariance risk premium term)}, \tag{1}$$

where α^{TC} is TC-alpha. Based on equation (1), we test the validity of SSD as a proxy of α^{TC} in two ways. First, we examine whether SSD predicts stock returns cross-sectionally; if SSD measures TC-alpha accurately, stocks with a greater SSD should earn a greater

[5] VoV(%Spread) is a proxy of effective spreads. Fong et al. (2017a) report that it dominates other proxies of relative transaction costs for U.S. stocks, in terms of closely approximating high-frequency transaction costs proxies.

expected return. Second, we test whether equation (1) holds once we replace α^{TC} by its empirical counterpart SSD; again, this should be the case if SSD measures TC-alpha accurately. We confirm both predictions empirically. A long-short portfolio constructed by going long in stocks with high SSD and short in stocks with low SSD yields statistically significant alphas between 1.5–1.8% per month, depending on the model used to estimate alphas. In addition, we find that we cannot reject equation (1) when we employ standard factor models (e.g., Carhart, 1997 four-factor model, Hou et al., 2015 *q*-model).

Our study contributes to three strands of literature. First, it contributes to the literature on the effect of market frictions to expected returns, including transaction costs (e.g., Amihud and Mendelson, 1996) and short-sale constraints (e.g., Chen et al., 2002; Ofek et al., 2004; Asquith et al., 2005; Drechsler and Drechsler, 2014).[6] Our contribution lies in proposing a novel way to measure TC-alpha. This differentiates us from existing studies that typically examine the relation between *overall* expected stock returns and the size of transaction costs. Furthermore, our study highlights the distinction between the size of transaction costs (ρ) and TC-alpha. Knowledge of the size of transaction costs does not suffice to estimate TC-alpha. The value of ρ determines the range in which TC-alpha lies, yet ρ is silent about the exact value that TC-alpha takes, as our theoretically derived TC-alpha interval reveals. This observation sheds light on the previously documented weak predictive power of frictions for the returns of large stocks (e.g., Hou et al., 2018). Grouping stocks in portfolios based on a proxy of the size of transaction costs places stocks with positive and negative TC-alpha in the same portfolio, thus it yields a portfolio with TC-alpha close to zero.

Our research is also related to studies which document that measures of deviations from put-call parity measure short-selling costs (e.g., Ofek et al., 2004; Muravyev et al., 2017; Cremers et al., 2019; Muravyev et al., 2022). We contribute to these studies by proposing a measure which quantifies *formally* the *effect*, rather than the size, of *broader* types of transaction costs to stock expected returns. We achieve this by developing our measure within a general asset pricing setting which highlights the importance of the choice of the scaling factor. Two remarks are in order here. First, the fact that our measure is related to transaction costs is not a take-down on alternative existing explanations for the predictive ability of deviations from put-call parity, such as informed

[6] More broadly, this literature includes the effect of liquidity risk (Acharya and Pedersen, 2005, Amihud, 2002), market and funding liquidity constraints (Brunnermeier and Pedersen, 2009), margin constraints (Chabakauri, 2013), margin and leverage constraints (Frazzini and Pedersen, 2014, Jylhä, 2018), uncertainty about future shorting costs (Engelberg et al., 2018), idiosyncratic volatility (Ang et al., 2006; Stambaugh et al., 2015), delay in the response of prices to information (Hou and Moskowitz, 2005), intermediaries' liquidity constraints (Nagel, 2012), exclusion of strategies with possible unlimited losses (Jarrow, 2016), and measures which are agnostic on the type of market frictions which affect expected returns (Brennan and Wang, 2010 and Hou et al., 2016).

trading (Bali and Hovakimian, 2009; Cremers and Weinbaum, 2010) and price pressure (Goncalves-Pinto et al., 2020); all these explanations can co-exist (Easley et al., 1998). Second, our paper should not be interpreted as one more paper on documenting that deviations from put-call parity predict future stock returns; in fact, there is already a voluminous literature on the ability of market option prices, including deviations from put-call parity, to predict future stock returns (see e.g., Gkionis et al., 2021 and references therein). Rather, our paper examines whether our proposed measure predicts future stock returns as part of *only one* of the tests to verify the testable implications of our theoretically founded measure; our SSD measure oughts to predict future stock returns as equation (1) shows.

Interestingly, as a by-product of our approach, we shed light on an important, yet previously overlooked finding; the predictive ability of deviations from put-call parity is large (alphas more than 1.0% per month, reported by above mentioned studies) given that the stock universe is confined to optionable stocks which tend to be large stocks with relatively low transaction costs. The foundation of our SSD measure within the TC-alpha augmented asset pricing setting enables us to explain the emergence of these sizable alphas. Our interval of TC-alpha predicts that the upper bound of TC-alpha of the spread long-short portfolios approximately equals $4\rho + \theta$. A commonly used value of the estimates of the relative transaction costs (around 0.4–0.5%) and short-sale cost (see e.g., Muravyev et al., 2017) justifies the size of the reported alphas of the SSD-sorted portfolios.[7]

Finally, more broadly speaking, our estimator shares a key insight with Gârleanu and Pedersen (2011) and Pasquariello (2014), that deviations from the law-of-one price occur due to limits to arbitrage caused by market frictions. The latter study constructs a composite measure of market frictions based on violations of arbitrage parities in stock, foreign exchange and bond markets. In this paper, we focus on transaction costs, and demonstrate that our scaled deviations from put-call parity measure, SSD, accurately proxies the effect of transaction costs for trading the underlying stock.

The rest of the paper is structured as follows. In Section 2, we introduce the synthetic-spot difference (SSD) measure and explain data sources and empirical procedures to calculate SSD for U.S. individual stocks. Section 3 documents the empirical characteristics of SSD. Then, we provide the two sets of results which confirm that SSD is a valid proxy of TC-alpha. Section 4 concludes.

[7]This value of the relative transaction costs is in line with the estimates provided by studies which use low-frequency transaction costs estimates (e.g., Hasbrouck, 2009; Goyenko et al., 2009; Fong et al., 2017b). These estimates are obtained from the TAQ database, which contains information on the trades executed by retail investors and liquidity traders. We cite these low-frequency transaction costs estimates because our database also includes trades executed by the same type of investors. See Appendix D for further discussion on this.

2 Synthetic-stock difference and transaction costs

2.1 Motivation

Let $R_{t,T}$ be the return of the stock from time t to T ($t < T$) and $m_{t,T}$ be the pricing kernel. In a frictionless market, $1 = \mathbb{E}_t^{\mathbb{P}}[m_{t,T}R_{t,T}]$ holds (see e.g., Cochrane, 2005), where $\mathbb{E}_t^{\mathbb{P}}[\cdot]$ denotes the conditional expectation operator under the physical measure $\mathbb{P}$, given the information up to time t. However, in the presence of transaction costs, He and Modest (1995) show that the following inequality relation holds instead,

$$\frac{1-\rho}{1+\rho} \leq \mathbb{E}_t^{\mathbb{P}}\left[m_{t,T}R_{t,T}\right] \leq \frac{1+\rho}{1-\rho}, \tag{2}$$

where ρ denotes the one-way relative transaction cost. Furthermore, by also considering short-sale cost of θ on top of the symmetric transation costs ρ into account, the inequality (2) becomes

$$\frac{1-\rho-\theta}{1+\rho} \leq \mathbb{E}_t^{\mathbb{P}}\left[m_{t,T}R_{t,T}\right] \leq \frac{1+\rho}{1-\rho}. \tag{3}$$

Equation (3) yields the following expression for the expected excess return,[8]

$$\mathbb{E}_t^{\mathbb{P}}[R_{t,T}] - R_{t,T}^0 = \alpha_{t,T}^{TC} - R_{t,T}^0 Cov_t^{\mathbb{P}}(m_{t,T}, R_{t,T}), \tag{4}$$

where $R_{t,T}^0$ is the gross risk-free rate from t to T, and $\alpha_{t,T}^{TC}$ satisfies the following inequality,

$$-R_{t,T}^0 \frac{2\rho+\theta}{1+\rho} \leq \alpha_{t,T}^{TC} \leq R_{t,T}^0 \frac{2\rho}{1-\rho}. \tag{5}$$

The second term on the right-hand side of equation (4) represents the standard covariance risk premium term (i.e., the covariance between the stock return and the pricing kernel). Equation (4) shows that in the presence of transaction costs (including short-sale costs), the expected excess return is the sum of the covariance risk premium and the alpha term, $\alpha_{t,T}^{TC}$. We call $\alpha_{t,T}^{TC}$ the *transaction costs alpha* (TC-alpha in short), because it is the part of the expected excess return arising due to the inclusion of transaction costs that cannot be explained by the covariance risk premium.

Inequality (5) yields two important implications. First, it shows that TC-alpha can be either positive or negative. A positive (negative) TC-alpha occurs when agents buy ((short-)sell) the stock. In this case, the expected return should be greater (less) than the covariance risk premium so that transaction costs (including short-sale costs, if applicable) incurred by the agent are covered. Second, the quantitative impact of the $\alpha_{t,T}^{TC}$ term on

[8] For simplicity, we assume that the risk-free bond market is frictionless, that is, $\mathbb{E}_t^{\mathbb{P}}[m_{t,T}] = 1/R_{t,T}^0$ holds. Relaxing this assumption does not affect the subsequent empirical analysis qualitatively.

the expected stock return can be large. Even if we ignore the short-sale cost parameter θ and assume a typical value of $\rho = 0.5\%$ (e.g., Hasbrouck, 2009; Novy-Marx and Velikov, 2016), inequality (5) suggests that $\alpha_{t,T}^{TC}$ can be as small as -1% or as big as $+1\%$. When considering short-sale costs θ, more negative values can occur as the lower bound further decreases approximately by the value of θ. We comment further on the values of ρ and θ in Appendix D.

The estimation of $\alpha_{t,T}^{TC}$ is not trivial. Knowledge of the relative transaction cost ρ (and θ) of the stock does not suffice to estimate $\alpha_{t,T}^{TC}$; the value of ρ determines the *range* in which $\alpha_{t,T}^{TC}$ lies, yet ρ is silent about the *exact* value that $\alpha_{t,T}^{TC}$ takes. We will infer $\alpha_{t,T}^{TC}$ from market option prices.

2.2 Synthetic-stock difference

We define the *synthetic stock price* $\widetilde{S}_t(K,T)$ as

$$\widetilde{S}_t(K,T) = C_t(K,T) - P_t(K,T) + \frac{K + D_T}{R_{t,T}^0}, \tag{6}$$

where $C_t(K,T)$ and $P_t(K,T)$ are the time t price of European call and put option with strike price K and maturity T, respectively, and D_T is the dividend payment at time T. For simplicity, we assume that the stock pays dividends only at time T and its amount is known at time t. The deterministic dividend payment assumption is plausible for short-term maturity options as the ones that we will use for our empirical analysis; near-future dividends are usually pre-announced.

Next, we define the *synthetic-stock difference* (SSD) as

$$SSD_t(K,T) = \frac{R_{t,T}^0}{T-t} \frac{\widetilde{S}_t(K,T) - S_t}{S_t}. \tag{7}$$

SSD is a relative deviation from put-call parity scaled by the ratio of the gross risk-free rate to the options' time-to-maturity. Appendix B explains that the definition of SSD as in equation (7) is theoretically founded. Notably, the numerator of the scaling factor ensures that SSD is related to expected stock returns. We scale SSD by time-to-maturity so that it is comparable across different options' maturities T. This is because, as we will argue shortly, SSD proxies a part of expected *returns* (i.e., TC-alpha). Normalizing it to, say, annualized returns, is also necessary since our empirical SSD measure will aggregate deviations from put-call parity of pairs with different strikes and different times-to-maturity.

Since SSD is scaled deviations from put-call parity, it should be zero when the stock market and option market are frictionless and arbitrage-free. On the contrary, a non-zero SSD indicates the existence of market frictions that limit arbitrage either in the underlying market and/or in the option market. The detection of the source of deviation is an empirical issue. In Section 3, we provide empirical evidence that SSD accurately proxies the underlying stock's TC-alpha.

Equation (7) shows that the value of SSD depends on the choice of strike K and maturity T. To construct our empirical SSD measure, at any point in time, we aggregate the SSDs calculated from pairs of options with different stirke prices and maturities to obtain a single SSD value for each stock-date pair. The construction of a composite SSD measure is desirable because it may cancel out any noise in individual SSDs caused by measurement errors in option prices. This will improve the accuracy in the estimation of TC-alpha.

We aggregate SSDs as follows. First, for each traded maturity T_i, we take the weighted average of the SSDs across strikes for which the call and put implied volatilities (IVs) are available. That is,

$$SSD_t(T_i) = \sum_{K_j} w_j SSD_t(K_j, T_i), \tag{8}$$

where the weight w_j is the relative open interest of the corresponding options, in line with Cremers and Weinbaum (2010). Then, we linearly interpolate $SSD_t(T_i)$ obtained from the two traded maturities surrounding the 30-day maturity, to construct the 30-day constant maturity measure. On a particular date, we treat the aggregated SSD as missing, if the 30-day maturity is not bracketed by two traded maturities. Hereafter, we denote the SSD measure constructed in this way by SSD_t.[9]

2.3 Data

To calculate SSDs, we obtain U.S. equity option prices and IVs from the Option Price file of the OptionMetrics Ivy DB database (OM) via the Wharton Research Data Services (WRDS). Our dataset spans January 1996 to December 2020. Options written on the U.S. individual equities are American style. OM calculates IVs via the Cox et al. (1979) binomial tree model, which takes the early exercise premium of American options into account. To calculate synthetic stock prices, we convert OM-IVs to the corresponding European option prices via the Black and Scholes (1973) option pricing formula, in line with Martin and Wagner (2019). We obtain the risk-free rate and dividend payment

[9] As an unreported robustness check, we also construct an alternative SSD measure obtained from the put-call option pair whose strike price is closest to the spot price and whose maturity is closest to 30-day. This alternative SSD measure is a noisier version of our baseline SSD. For example, its ability to forecast future stock returns cross-sectionally, as equation (4) predicts, is weaker than that of our baseline SSD, SSD_t. This is expected as it utilizes information from only one pair of call and puts and hence it may be impacted from any measurement errors in market option prices.

data from the OM database to calculate the present value of dividend payments over the option's life time. We remove IVs if the recorded corresponding option bid price is non-positive, the IV is missing, or the option's open interest is non-positive. We discard data with time to maturity shorter than 8 days or longer than 270 days. We keep option data only when the moneyness K/S_t is between 0.9 and 1.1 to ensure that the most liquid option contracts are considered.

For the stock return data, we obtain stock returns from the Center for Research in Security Prices (CRSP). In line with the literature, our stock universe consists of all U.S. common stocks (CRSP share codes 10 and 11).[10] For our subsequent tests, we also obtain the time-series of risk factors in the Carhart (1997) 4-factor model (FFC) from Kenneth French's website, the Stambaugh and Yuan (2017) mispricing factor model (SY) from Yu Yuan's website, and the Hou et al. (2015) q-factor (HXZq) from WRDS. We construct various firms' and stocks' characteristics variables and proxies of transaction costs based on the CRSP and Compustat database. Appendix C provides the definition and data source of the various variables.

3 Empirical results

In Section 3.1, we report the summary statistics of SSD. Then, we document that SSD proxies TC-alpha well by verifying the testable implications of our theoretical setting outlined in Section 3.2. Specifically, Sections 3.3 and 3.4 document that there is a connection between SSD and transaction costs proxies, in a theoretically expected way. In Section 3.5, we document that SSD accurately proxies TC-alpha, based on asset pricing tests implied by the TC-alpha-augmented asset pricing setting (equation (4)).

3.1 Summary statistics

Table 1 reports the summary statistics of SSD_t calculated at the end of each month; both measures are normalized corresponding to a 30-day stock return. We can see that there are about 433,000 stock-month observations. This yields on average about 1,440 stocks in each month which ensures the formation of well-diversified decile portfolios in the subsequent analysis.

The percentage of the negative observations (the neg column) is about 53%, that is, there are slightly more cases where the underlying stock price is higher than the synthetic stock price. The distribution of SSD is skewed to the left and has a longer left tail; mean

[10]The stock identifiers differ in the CRSP and OptionMetrics databases (PERMNO and SECID, respectively). We link CRSP and OptionMetrics by the macro provided by the WRDS, which matches PERMNO and SECID.

and median are slightly negative (although they are virtually equal to zero), the absolute value of the fifth percentile is greater than the 95th percentile. These empirical patterns are consistent with equation (5), as they suggest an asymmetrically longer left tail due to the existence of short-sale costs. SSD can become sizable compared with average equity risk premia (e.g., Mehra, 2012 reports monthly equity risk premium of about 0.5% per month); it takes both positive and negative values, ranging from -1.24% to 0.92% per 30-day in a 5th to 95th percentile range. They also have fairly large cross-sectional variations; the standard deviation is about 0.9% and the interquartile range (IQR, the difference between 75th and 25th percentile points) is about 0.5% on average across stocks.

[Table 1 about here.]

Next, we examine the evolution of the cross-sectional dispersions of SSD over time. Figure 1 shows the time-series evolution of the monthly IQR of SSD_t. We can see that most of the spikes in the IQR correspond to periods of market turmoil, such as the Russian default and LTCM crisis (August to September 1998), the collapse of Lehman Brothers and ensuing market meltdown (September to November 2008), the European debt crisis (November 2011 associated with the uncertainty with the political situation in Greece), the Chinese stock market turmoil (June 2015 to January 2016), and the early stage of the Covid-19 pandemic (March 2020). In addition, with the exception of these distressed periods, we can observe a secular decline in the IQR. These empirical patterns are consistent with the fact that transaction costs intensify over market crises (Nagel, 2012; Hou et al., 2016), and the evidence that transaction costs have decreased over the recent years partly due to the automation of markets (Green et al., 2017).

[Figure 1 about here.]

Then, we examine the persistence of SSD values over time in the cross-section of stocks. This will enable us to understand how TC-alpha changes over time cross-sectionally. To this end, at the end of each month, we classify stocks in tercile portfolios; tercile 1 (3) contains stocks with very negative (positive) SSD_t and tercile 2 contains stocks whose SSD_t is close to zero. Then, we calculate the empirical transition probability matrix between the SSD_t tercile portfolios. That is, p_{ij} denotes the probability of stocks moving from the i-th tercile in month $t-1$ to the j-th tercile in month t. Table 2 reports the transition probabilities.

The first and the third row of Table 2 suggest that SSD fluctuates randomly in the succeeding month for stocks whose value of SSD is either very high or low in the current month; these two rows show that the transition probabilities are all roughly close to one third regardless of destination bins. On the other hand, the SSD of stocks in tercile 2 (i.e., SSDs which do not take extreme values in the current month) are relatively more likely to stay in tercile 2 and thus exhibiting a non-extreme SSD value in the next period. This pattern is consistent with our theoretically derived interval of TC-alpha, that is, stocks with lower ρ value have a narrower SSD range. Conditional on observing SSD close to zero (i.e., observing a stock in tercile 2), the stock's ρ parameter is expected to be small, and hence the possible range of SSD value is narrower. Therefore, the probability of moving from tercile 2 to tercile 1 or 3 in the following month is smaller.

[Table 2 about here.]

3.2 Testable implications of the range of TC-alpha

We test whether SSD accurately proxies TC-alpha by examining whether the empirical patterns of SSD are in line with the following three testable implications of the theoretically derived interval of TC-alpha (equation (5)).

First, in the cross-section of stocks, the *value* of SSD should be non-linearly related to ρ; it would take more extreme values (either positive or negative) for stocks which are subject to greater transaction costs. Second, the *dispersion* (measured either as the range or as the standard deviation) of SSD should be positively and (approximately) linearly correlated to proxies of ρ. The range $4\rho + \theta$ is an affine function of ρ and the standard deviation is proportional to the range at least approximately.

Third, equation (5) suggests that the alpha of an SSD-sorted spread portfolio (big SDD stocks minus small SSD stocks) should be greater for stocks which face greater transaction costs. This is because a greater value for the transaction cost parameter ρ and/or for the short-sale cost parameter θ translates to a wider range of TC-alpha, and hence a bigger difference between the TC-alphas of the SSD-sorted long-leg and short-leg portfolios. Note that the short-sale constraints parameter θ is asymmetric in the sense that a larger θ signifies wider negative SSD regions whereas it does not affect the positive SSD region. This suggests that larger θ ties to more stock underperformance (i.e., negative alpha). This is consistent with large body of literature that documents that severer short-sale constraints are asymmetrically associated with stock underperformance (see e.g., Jones and Lamont, 2002; Boehme et al., 2006; Stambaugh et al., 2015, among others).

3.3 Relation between SSD and transaction costs proxies

To test the first and second implications mentioned in Section 3.2, we examine the relation of SSD with a number of popular low-frequency proxies of transaction costs which have also been documented to be highly correlated with their high-frequency proxies counterparts (e.g., Goyenko et al., 2009; Fong et al., 2017b). We employ both *percent-cost* liquidity proxies and *cost-per-dollar-volume* liquidity proxies to proxy our ρ parameter. The former type of proxies measure the percentage transaction cost and the latter assess the marginal cost of trading an additional dollar amount of a large trade (Fong et al., 2017b, p.1357).[11] Regarding the percent-cost liquidity proxies, we employ the relative bid-ask spread (Chung and Zhang, 2014), Roll's (1984) measure, High-Low measure of Corwin and Schultz (2012), FHT of Fong et al. (2017b), VoV(%Spread) of Fong et al. (2017a). For the cost-per-dollar volume liquidity proxies, we examine Amihud's (2002) measure, and four types of extended Amihud measures, namely, the relative bid-ask impact, FHT impact, High-Low impact, and VoV impact (see Goyenko et al., 2009 for the definition of extended Amihud measures). Appendix C provides details on the calculation of these measures.

[Table 3 about here.]

Table 3 reports the average value of the above mentioned proxies of liquidity and standard stock characteristics for each one of the value-weighted decile portfolios sorted by SSD_t. Portfolio 1 (10) contains stocks with the most negative (positive) SSD_t, whereas the middle portfolio (Portfolio 6) contains stocks whose SSD_t is close to zero. Consistent with the first implication, we can see a U-shape relation between the cross-sectional SSD_t and each one of the employed proxies of transaction costs. This manifests that a large value of SSD (either very positive and nevative) is associated with greater transaction costs. This result is consistent with the first theoretical prediction of equation (5); TC-alpha (or its empirical counterpart, SSD) can take an extreme value only when the transaction costs parameter ρ is large. Moreover, we can see that the trough of the U-shape relation occurs at Portfolio 6. This is also in line with the theoretical prediction of equation (5); stocks in Portfolio 6 are expected to have a small value of ρ, conditional on observing SSD close to zero. Regarding the relation between SSD and other standard stocks' characteristics, we observe an inverse U-shape relation for the SIZE (the logarithm of the market equity) and the stock price level (S_t). Again, this is expected as larger (smaller) these two variables are known to be related with smaller (larger) stock market transaction costs (see e.g., Jegadeesh and Titman, 2001; Novy-Marx and Velikov, 2016).

Next, we examine the second implication that there should be a linear and positive correlation between transaction costs and the dispersion of SSD. We use two alternative measures of dispersion of SSD, that is its range and its standard deviation. For each

[11] It is possible to extend our model to incorporate trading volume-dependent transaction costs. We do not take this approach because it does not change the essence of the model and it complicates its exposition.

month and for each stock, we calculate the range and the standard deviation of daily SSD_t observed over the month. We treat these two statistics missing if no more than ten observations are available in the month. Table 4 reports the summary statistics of the two dispersion statistics of SSD and of the transaction costs proxies over the stock-month observations with valid SSD_t.

[Table 4 about here.]

Table 5 shows the time-series average of monthly cross-sectional correlations between the two dispersion statistics of SSD and the transaction costs proxies. We can see that the dispersion statistics of SSD are positively correlated with each transaction costs proxy contained in our menu of transaction costs measures. These correlations can be quite high for some proxies. For instance, the standard deviation of SSD exhibits a high correlation of 0.62 with VoV(%Spread), which has been found to dominate other proxies of relative transaction costs for U.S. stocks, in terms of closely approximating high-frequency transaction costs proxies (Fong et al., 2017a). These results validate the second implication in Section 3.2 and render further empirical support to the accuracy of SSD as a measure of TC-alpha.

[Table 5 about here.]

3.4 Predictive ability of SSD and transaction costs

We test the third implication of equation (5) shown in Section 3.2 which dictates that the return predictive ability of SSD should be greater for stocks which face gretater transaction costs. We conduct dependent bivariate portfolio sort analysis. First, we sort stocks in portfolios by their respective proxy for ρ or θ. Then, within any given portfolio, we sort stocks in value-weighted portfolios based on their respective SSD, and calculate the spread portfolios' average returns and alphas. We report results based on the relative bid-ask spread as a proxy of ρ (Chung and Zhang, 2014) due to space constraints; results are qualitatively similar when alternative proxies of ρ in this paper are used as a sorting variable. We use the relative short interest (RSI) as a proxy of short-sale constraints; a greater RSI signifies greater short-sale costs (Asquith et al., 2005).

Table 6, Panel A (B) reports results for the bivariate dependent sort, first by the relative bid-ask spread (RSI), and then by SSD_t. We can see that the average return and α_{FFC} of the SSD-sorted spread portfolios increase as the bid-ask spread and RSI increase. Therefore, our results confirm the prediction that the dispersion of SSD and hence the predictive power of SSD is greater among stocks which are subject to bigger transaction costs and short-sale costs. Moreover, the RSI-SSD double sort result shows that only sigificantly negative (positive) alphas can be observed in the largest (smallest) RSI bin. Since larger RSI implies severer short-sale constraints (i.e., greater value of θ), this result corroborates the implication that stocks with larger θ underperform.

[Table 6 about here.]

A remark is in order about our result that the portfolio sort analysis using friction-related variables (transaction costs proxies in our case) yield large spread portfolio alphas. This finding may seem puzzling because the previous literature finds a weak return predictive power of friction-related variables among large stocks (e.g., Hou et al., 2018). We reconcile our findings with these of the previous research as follows. Our theoretical interval in equation (5) implies that grouping stocks by a proxy of ρ is not an effective way to trace which stocks will outperform and which will underperform (in terms of the sign of SSD). This is because the relation between ρ and SSD is non-linear; a portfolio consisting of stocks with large ρ contains stocks with big negative and positive SSD which may offset each other and will yield a close to zero SSD (TC-alpha), thus showing no outperformance. In fact, our bivariate portfolio sort result in Table 6, Panel A, shows that the widest relative bid-ask spread bin contains sub-portfolios with negative average return and positive average return.

3.5 Accuracy of SSD measure: Further asset pricing tests

We provide two further tests to confirm the accuracy of SSD to estimate TC-alpha. First, equation (4) shows that positive (negative) TC-alphas predict positive (negative) abnormal return. Similarly, stocks with higher TC-alphas should earn a greater average return than stocks with lower TC-alphas, implying that a TC-alpha-sorted long-short spread portfolio will earn positive abnormal return. These relations should also hold for the SSD-sorted stocks, if SSD accurately measures TC-alpha. To test this, every month, we sort stocks by SSD_t, where Portfolio 1 (10) contains the stocks with the lowest (highest) SSD_t and we compute the value-weighted monthly portfolio post-ranking return. We also construct a zero-cost long-short spread portfolio, where we go long in Portfolio 10 and short in Portfolio 1. We estimate the alpha of each portfolio with respect to the FFC, SY, and HXZq models.

Table 7 reports the results. Decile portfolios with low and negative SSD_t (Portfolios 1, 2) earn negative and significant risk-adjusted returns, whereas those with high and positive SSD_t (Portfolio 10) earn positive and significant risk-adjusted returns. This implies that a positive (negative) SSD predicts a positive (negative) abnormal return.

[12] The alpha of the SDD-sorted spread portfolio is sizable, more than 1.6% per month. These results showcase the predictive ability of SSD in line with theory and support the accuracy of SSD as a measure of TC-alpha.[13]

[Table 7 about here.]

The obtained sizable alphas are not surprising. Previous studies also report that the alphas of spread portfolios formed based on deviations from put-call parity exceed 1% per month (Bali and Hovakimian, 2009, Table 3; Cremers and Weinbaum, 2010, Table 2; Goncalves-Pinto et al., 2020, Tables 2 and 3, among others). Nevertheless, the literature does not provide an explanation on why such sizable alphas emerge within the universe of optionable stocks, which tend to be large stocks. In Appendix D, we provide an explanation on this important, yet overlooked empirical finding, based on the TC-alpha-augmented general asset pricing setting and our SSD measure.

Next, we provide additional evidence that SSD is an accurate proxy of TC-alpha based on a regression-based asset pricing test. Specifically, we test the following testable hypothesis of equation (4) formulated in a regression setting as

$$R_{t,T} - R^0_{t,T} = \alpha + \beta' f_{t,T} + \gamma SSD_t + \varepsilon_T, \tag{9}$$

where $f_{t,T}$ is a vector of risk-factors (i.e., $\beta' f_{t,T}$ approximates the covariance risk premium term). Under equation (4) and assuming that SSD proxies TC-alpha, then $\alpha = 0$ and $\gamma = 1$ should hold. We test this hypothesis by a pooled regression of the value-weighted decile portfolio returns constructed by sorting stocks by SSD_t. We employ FFC, SY, and HXZq models as alternative sets of risk factors $f_{t,T}$. Table 8 reports the results. We can see that the point estimates of α are not significant regardless of the risk-factor models. Moreover, the results from the Wald test show that the joint null hypothesis of $H_0 : \alpha = 0, \gamma = 1$ cannot be rejected at a 10% significance level for SY and HXZq models and at a 1% significance level for FFC model. Therefore, the asset pricing equation (4) cannot be rejected when SSD_t is used as a proxy of TC-alpha. This result supports (in the sense that our test did not reject it) the transaction cost-augmented asset pricing model of equation (4) and also suggests that SSD_t estimates TC-alpha accurately.

[Table 8 about here.]

[12] We also confirm a monotonic relation between the portfolios' average returns and SSD_t. We conduct the Patton and Timmermann (2010) monotonicity test. The null of no systematic relation between SSD_t and the post ranking decile portfolio returns is rejected in favour of an increasing relation with $p = 0.017$.

[13] Our unreported robustness results also confirm that the predictive results are not driven by short-term reversals, non-synchronous trading in the option and underlying market, low prices, and alternative breakpoints to construct stock portfolios.

4 Conclusions

In the presence of transaction costs in trading stocks (including short-salling costs), expected stock returns contain the *transaction costs alpha* (TC-alpha) term in addition to the conventional risk premium term. We propose a model-free, real-time and easy-to-implement measure of TC-alpha, namely the *synthetic-spot difference* (SSD). SSD is properly scaled price differences between the synthetic stock price, replicated at any point in time from pairs of European call and put option prices with the same strike price and maturity, and the quoted underlying stock price.

We calculate SSD for each optionable U.S. common stock over January 1996 to December 2020. We find that extreme negative SSD values are more likely to occur than extreme positive ones, due to the presence of short-sale costs. SSD can become sizable, especially during a market turmoil, and it can be highly volatile over time for stocks with a very low or a very high SSD. We document that SSD accurately proxies TC-alpha by providing two sets of results. First, we confirm that there is a significant connection between SSD and several commonly used proxies of transaction costs. Second, we find that SSD predicts stock returns cross-sectionally, and it is consistent with the cross-section of observed stock returns regardless of the chosen asset pricing model. Both sets of results are in line with the predictions of a general and standard asset pricing setting in the presence of transaction costs.

Notably, as a by-product of our analysis, our setting shows that the strength of the documented predictability (alpha) of SSD is in line with standard empirical values of a host of transaction costs proxies. Thus, our findings address an issue of importance overlooked by previous literature, that is, the large alphas obtained from sorting underlying stocks based on deviations from put-call parity. From a practitioner's perspective, the degree to which this alpha is exploitable in practice depends crucially on the type of investors and the size of transaction costs they face.

Our findings imply that transaction costs play an important role in determining expected stock returns, even for big stocks. Hence, our measure can be used by future research to address questions in asset pricing and portfolio construction which would require knowledge of the impact of transaction costs to expected returns.

A Proof of equations in Section 2.1

First, we prove the right inequality in equation (3) as in He and Modest (1995). They assume a utility-based asset pricing model; the pricing kernel $m_{t,T}$ equals the intertemporal marginal rate of substitution $\beta u'(c_T)/u'(c_t)$, where β is the discount factor, $u(c_t)$ is a time-separable utility function and c_t is consumption at time t. Consider shifting one dollar from current consumption to time T consumption. In this case, the investor can buy $\frac{1}{S_t(1+\rho)}$ units of the stock at time t and can obtain $\frac{1}{S_t(1+\rho)}S_T(1-\rho)$ dollars at time T. At the optimum point, we obtain the following inequality because deviating from the optimum (weakly) reduces the expected utility.

$$-u'(c_t) + \mathbb{E}_t^{\mathbb{P}}\left[\beta u'(c_T)\left(\frac{1-\rho}{1+\rho}\right)R_{t,T}\right] \leq 0.$$

This proves the right inequality of equation (3).

The proof of the left inequality of equation (3) is similar. Consider increasing one dollar of consumption by short-selling stock at time t. In this case, the investor needs to short-sell $\frac{1}{(1-\rho-\theta)S_t}$ units of the stock. To cover this short-sale position at time $t+1$, the investor needs to pay $\frac{1}{(1-\rho-\theta)S_t}(1+\rho)S_T$ dollars. The optimality condition similar to the above inequality yields

$$u'(c_t) - \mathbb{E}_t^{\mathbb{P}}\left[\beta u'(c_T)\left(\frac{1+\rho}{1-\rho-\theta}\right)R_{t,T}\right] \leq 0.$$

This proves the left inequality of equation (3).

Replacing the middle term of equation (3) with $\mathbb{E}_t^{\mathbb{P}}[m_{t,T}R_{t,T}] = Cov_t^{\mathbb{P}}(m_{t,T}, R_{t,T}) + \mathbb{E}_t^{\mathbb{P}}[R_{t,T}]/R_{t,T}^0$ and doing some algebra yields that the expected excess return satisfies equation (4) with the range condition for TC-alpha, equation (5).

B Definition of SSD: Theoretical foundation

Let $m_{t,T}^*$ be the pricing kernel of a marginal investor who trades in the stock and option market. In a frictionless market, the market price of an asset should equal its expected payoff discounted by the pricing kernel. Since the underlying stock and the synthetic stock have identical cash flows at time T, their prices should be identical. This leads to the following well-known *put-call parity* relation.

$$S_t = \mathbb{E}_t^{\mathbb{P}}[m_{t,T}^*(S_T + D_T)] = \widetilde{S}_t(K,T). \tag{B.1}$$

On the other hand, in real markets where market frictions exist, put-call parity does not hold, and deviations from put-call parity frequently become sizeable. This means that either or both the underlying stock price and the synthetic stock price deviate from their expected discounted payoff value $\mathbb{E}_t^{\mathbb{P}}[m_{t,T}^*(S_T + D_T)]$. We denote the wedge between the underlying stock price (the synthetic stock price) and its expected discounted payoff value by $\omega_t(T)$ ($\widetilde{\omega}_t(K,T)$), that is,

$$S_t = \omega_t(T) + \mathbb{E}_t^{\mathbb{P}}[m_{t,T}^*(S_T + D_T)], \tag{B.2}$$
$$\widetilde{S}_t(K,T) = \widetilde{\omega}_t(K,T) + \mathbb{E}_t^{\mathbb{P}}[m_{t,T}^*(S_T + D_T)]. \tag{B.3}$$

A non-zero wedge term of the underlying stock, ω_t, may arise from frictions in trading the underlying stock including transaction costs (He and Modest, 1995), margin constraints (Gârleanu and Pedersen, 2011), short-sale constraints (Ofek et al., 2004), among others. For the synthetic stock, a non-zero wedge term $\widetilde{\omega}_t$ may arise from frictions in trading options (e.g., Frazzini and Pedersen, 2022; Santa-Clara and Saretto, 2009; Hitzemann et al., 2018). Note that ω_t and $\widetilde{\omega}_t$ may take positive or negative values.

By transforming equation (B.2), we obtain

$$\mathbb{E}_t^{\mathbb{P}}[R_{t,T}] - R_{t,T}^0 = -\frac{R_{t,T}^0 \omega_t(T)}{S_t} - R_{t,T}^0 Cov_t^{\mathbb{P}}(R_{t,T}, m_{t,T}). \tag{B.4}$$

Comparing equations (4) and (B.2) yields that $\alpha_t^{TC} = -R_{t,T}^0 \omega_t(T)/S_t$. Equations (B.2) and (B.3) yield

$$R_{t,T}^0 \frac{\widetilde{S}_t(K,T) - S_t}{S_t} = R_{t,T}^0 \frac{\widetilde{\omega}_t(K,T) - \omega_t(T)}{S_t}. \tag{B.5}$$

Note that, if $\widetilde{\omega}_t(K,T)$ is negligible, then it follows that the left-hand side of equation (B.5) proxies TC-alpha as equations (B.4) and (B.5) show. This is the basis for our definition of SSD in equation (7).

One cannot directly prove that $\widetilde{\omega}_t(K,T)$ is negligible. Instead, we examine whether SSD proxies TC-alpha accurately by performing two types of tests. First, we examine its connection with well-known proxies of transaction costs. Second, we test whether testable asset pricing hypotheses regarding TC-alpha hold if we use SSD as a proxy of TC-alpha (see Section 3).

C Description of variables

Relative bid-ask spread (BAS): We calculate the daily relative bid-ask spread as $BAS_d^i = (S_d^{ask,i} - S_d^{bid,i})/(0.5(S^{ask,i} + S^{bid,i}))$. Then, we average the daily bid-ask spread within each month. We require there are at least 15 non-missing observations. Data are obtained from the CRSP database.

FHT: We follow Fong et al. (2017b) and calculate FHT within each stock-month by

$$FHT = 2\sigma \cdot N^{-1}\left(\frac{1+z}{2}\right),$$

where σ is the standard deviation of non-zero stock returns within the month and N^{-1} is the inverse of the cumulative standard normal function. $z = ZRD/(TD + NTD)$ is the zeros, where ZRD is the number of zero returns days, TD is the number of non-zero trading volume days, and NTD is the number of zero trading volume days. We treat FHT missing if $TD < 5$ or $NZRD = TD + NTD - ZRD < 11$).

Roll: We follow Roll (1984) and calculate the following equation for each stock-month,

$$Roll = \begin{cases} \frac{2\sqrt{-Cov(\Delta S_t, \Delta S_{t-1})}}{\overline{S}}, & \text{if } Cov(\Delta S_t, \Delta S_{t-1}) > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where ΔS_t is the daily price change and $\overline{S}$ is the average stock price within the month. We treat Roll missing if $TD < 5$ or $NZRD < 11$). We normalize the Roll measure by the average stock price so that this measure corresponds to the relative spread.

VoV(%Spread): We follow Fong et al. (2017a) and calculate $VoV(\%Spread)$ for each stock-month by

$$VoV(\%Spread) = \frac{a \cdot \sigma^{2/3}}{V^{1/3}},$$

where $a = 8$, σ is the standard deviation of daily returns within the month, and V is the average daily dollar trading volume within the month. The trading volume is deflated by the consumer price index (CPI-U nromalized as 1.0 in January 2000). We treat VoV(%Spread) missing if $TD < 5$ or $NZRD = TD + NTD - ZRD < 11$).

High-Low: To calculate Corwin and Schultz's (2012) High-Low measure, first we calculate X_t for each trading day by

$$X_t = \frac{2(e^{a_t} - 1)}{1 + e^{a_t}}, \quad where \quad a_t = \frac{\sqrt{2\beta_t} - \sqrt{\beta_t}}{3 - 2\sqrt{2}} - \frac{\gamma_t}{3 - 2\sqrt{2}}$$

and γ_t is the squared log(High/Low) with High and Low are those over the previous two days. If X_t is negative, we replace it with $X_t = 0$. Then, we calculate High-Low measure for each stock-month by averaging X_t within the month. We treat High-Low missing if $TD < 5$ or $NZRD = TD + NTD - ZRD < 11$).

Amihud: We calculate daily Amihud's (2002) illiquidity measure as the ratio of the absolute daily return to the dollar trading volume, $Illiq_d^i = |R_d^i|/(S_d^i Vol_d^i)$, where R_d^i and Vol_d^i are the daily return and the trading volume of i-th stock on day d. Then, we average daily illiquidity measure within the month. We require there are at least 15 non-missing observations. The stock returns, stock prices, and trading volumes are obtained from the CRSP database. The trading volume of the NASDAQ equities is adjusted by following Gao and Ritter (2010).

Extended Amihud (FHT, Roll, VoV, High-Low): We follow Goyenko et al. (2009) and calculate the four types of Extended Amihud proxies, defined as

$$\text{Extended Amihud} = \frac{\text{Spread Proxy}}{\text{Average Daily Dollar Volume}}.$$

As a spread proxy, we use the FHT, Roll, VoV(%Spread), and High-Low, separately. The trading volume of the NASDAQ equities is adjusted by following Gao and Ritter (2010).

SIZE: Size is the natural logarithm of the market equity. The market equity is calculated as the product of the number of outstanding share with the price of the stock at the end of each month. Data are obtained from the CRSP database.

Relative short interest (RSI): The relative short interest (RSI) is calculated as the ratio of the number of short interest to the number of outstanding share. The short interest data is obtained from the Compustat North America, Supplemental Short Interest File via the WRDS. Until the end of 2006, the Compustat records the short interest at the middle of any given month (typically 15th day of each month). Since 2007, the short interest file contains the short interest at the middle of months and the end of months. We use the end-of-month short interest data since 2007 because we sort stocks in portfolios at the end-of-each month in our analysis. The number of outstanding share is obtained from the CRSP database.

Book-to-Market equity (B/M): We follow Davis et al. (2000) to measure book equity as stockholders' book equity, plus balance sheet deferred taxes and investment tax credit (Compustat annual item TXDITC) if available, minus the book value of preferred stock. Stockholders' equity is the value reported by Compustat (item SEQ), if it is available. If not, we measure stockholders' equity as the book value of common equity (item CEQ) plus the par value of preferred stock (item PSTK), or the book value of assets (item AT) minus total liabilities (item LT). Depending on availability, we use redemption (item PSTKRV), liquidating (item PSTKL), or par value (item PSTK) for the book value of preferred stock. From June of each year t to May of $t+1$, the book-to-market equity (B/M) is calculated as the ratio of the book equity for the fiscal year ending in calendar year $t-1$ to the market equity at the end of December of year $t-1$. We treat non-positive B/M data as missing.

D Size of SSD-portfolio alphas: An explanation

The previous literature on the predictive ability of measures of deviations from put-call parity also reports sizable alphas (Ofek et al., 2004; Bali and Hovakimian, 2009; Cremers and Weinbaum, 2010; Muravyev et al., 2017; Cremers et al., 2019; Goncalves-Pinto et al., 2020). The theoretical setting of our proposed SSD measure can provide an explanation to the size of reported alphas. Our framework can express the maximum attainable TC-alpha of a long-short spread portfolio as a function of the transaction costs parameter ρ and short-sale costs parameter θ; it is approximately equal to $4\rho + \theta$. Therefore, the question is whether ρ and θ are empirically large enough to accommodate the magnitude of the reported SSD-sorted portfolio alphas.

Regarding ρ, the 95th percentile value of the six proxies of effective spreads is at least 0.76 (for FHT) (see Table 4). This suggests that 4ρ is at least in the order of 1.5% for stocks in the spread portfolio; effective spreads conceptually measure 2ρ and stocks in the spread portfolio (i.e., stocks in Portfolios 1 or 10) are subject to large ρ as documented by the U-shape relation reported in Table 3. With regard to the short-sale costs parameter θ, Table 1, Panel B in Muravyev et al. (2017) documents that the indicative borrowing fee at the higher end of the distribution of short-sale fees across optionable stocks, varies between 7 bps (90th percentile point) and 1.3% (99th percentile point) per month. Therefore, the $4\rho + \theta$ is likely larger than the order of 1.6%, that is, the reported alpha of the SSD-sorted spread portfolios is within the upper bound of the alpha of TC-alpha-sorted portfolios.[A.1]

A final remark is in order. Frazzini et al. (2014) report lower estimates of effective transaction costs based on a proprietary dataset. Hence, one may argue that the above justification of the reported SSD-alphas relies on an overestimated ρ. However, a caveat should be pointed out in this argument. As Frazzini et al. (2014) argue, their database yields much smaller estimated transaction costs because proprietary datasets contain exclusively trading records of large sophisticated institutional investors; this is a small subset of the information the TAQ database cover. Instead, our empirical analysis is based

[A.1] Our employed estimates of transaction costs are also largely in line with Novy-Marx and Velikov (2016). They report effective transaction costs estimated based on Hasbrouck's (2009) Gibbs sampling method by decades for the 2,000 largest firms as a function of the firm size ranking (Figure 1, p.109). WE find that the average size ranking of the stocks in the SSD-sorted decile spread portfolio is about 1,500. Hence, Novy-Marx and Velikov (2016) effective transaction cost values for the 1,500th largest stock over time is a reasonable benchmark to compare with the average effective transaction costs proxies of the two extreme SSD-sorted portfolios (i.e., Portfolios 1 and 10). They find that the ρ of the 1,500th largest stock is about 0.6% in 1990 and it reduces to about 0.4% in 2000s. Therefore, on average across our sample period, $4\rho = 1.5\%$ will be a plausible assumption. This value is also in line with the stylized effective transaction costs function adopted in Brandt et al. (2009) and DeMiguel et al. (2020). The stylized function in these studies determines transaction costs of individual stocks as a function of the stock size and calendar year, and it is calibrated to estimated effective transaction costs based on Hasbrouck's (2009) method.

on the CRSP and OptionMetrics databases. These reflect the actions of multiple classes of investors, including retail investors and liquidity traders, as the TAQ database does. Therefore, to ensure consistency in our discussion of the size of SSD-sorted portfolio alphas, we need to rely on the transaction costs estimates reported in the literature obtained from low-frequency transaction costs measures that have been found to proxy TAQ-based high-frequency transaction costs measures.[A.2]

References

Acharya, V. V. and L. H. Pedersen (2005) "Asset Pricing with Liquidity Risk," *Journal of Financial Economics*, 77, 375–410.

Amihud, Y. (2002) "Illiquidity and Stock Returns: Cross-section and Time-Series Effects," *Journal of Financial Markets*, 5, 31–56.

Amihud, Y. and H. Mendelson (1996) "Asset Pricing and the Bid-Ask Spread," *Journal of Financial Economics*, 17, 223–249.

Ang, A., J. Chen, and Y. Xing (2006) "Downside Risk," *Review of Financial Studies*, 19, 1191–1239.

Asparouhova, E., H. Bessembinder, and I. Kalcheva (2010) "Liquidity Biases in Asset Pricing Tests," *Journal of Financial Economics*, 96, 215–237.

Asquith, P., P. A. Pathak, and J. R. Ritter (2005) "Short Interest, Institutional Ownership, and Stock Returns," *Journal of Financial Economics*, 78, 243–276.

Bali, T. G. and A. Hovakimian (2009) "Volatility Spreads and Expected Stock Returns," *Management Science*, 55, 1797–1812.

Black, F. and M. Scholes (1973) "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637–654.

Blume, M. E. and R. F. Stambaugh (1983) "Biases in Computed Returns," *Journal of Financial Economics*, 12, 387–404.

[A.2]Our discussion suggests that large alphas associated with SSD can be explained by the existence of transaction costs. Nevertheless, this does not necessarily mean that no one can exploit these large alphas. Sophisticated investors who are subject to smaller transaction costs (like reported in Frazzini et al., 2014) may exploit the informational content of SSD to earn abnormal returns. As long as such investors are relatively small in the overall market trading flow, this does not contradict the case that seemingly unexploited SSD-alphas remain at the aggregate level.

Boehme, R. D., B. R. Danielsen, and S. M. Sorescu (2006) "Short-Sale Constraints, Differences of Opinion, and Overvaluation," *Journal of Financial and Quantitative Analysis*, 41, 455–487.

Brandt, M. W., P. Santa-Clara, and R. Valkanov (2009) "Parametric Portfolio Policies: Exploiting Characteristics in the Cross-Section of Equity Returns," *Review of Financial Studies*, 22, 3411–3447.

Brennan, M. J. and A. W. Wang (2010) "The Mispricing Return Premium," *Review of Financial Studies*, 23, 3437–3468.

Brunnermeier, M. K. and L. H. Pedersen (2009) "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, 22, 2201–2238.

Carhart, M. M. (1997) "On Persistence in Mutual Fund Performance," *Journal of Finance*, 52, 57–82.

Chabakauri, G. (2013) "Dynamic Equilibrium with Two Stocks, Heterogeneous Investors, and Portfolio Constraints," *Review of Financial Studies*, 26, 3104–3141.

Chen, J., H. Hong, and J. C. Stein (2002) "Breadth of Ownership and Stock Returns," *Journal of Financial Economics*, 66, 171–205.

Chung, K. H. and H. Zhang (2014) "A Simple Approximation of Intraday Spreads Using Daily Data," *Journal of Financial Markets*, 17, 94–120.

Cochrane, J. H. (2005) *Asset Pricing*, revised edition ed.

Corwin, S. A. and P. Schultz (2012) "A Simple Way to Estimate Bid-Ask Spreads from Daily High and Low Prices," *Journal of Finance*, 67, 719–759.

Cox, J. C., S. A. Ross, and M. Rubinstein (1979) "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, 7, 229–263.

Cremers, M., R. Goyenko, and P. Schultz (2019) "Do Option-Based Measures of Stock Mispricing Find Investment Opportunities or Market Frictions?" SSRN Working Paper.

Cremers, M. and D. Weinbaum (2010) "Deviations from Put-Call Parity and Stock Return Predictability," *Journal of Financial and Quantitative Analysis*, 45, 335–367.

Davis, J. L., E. F. Fama, and K. R. French (2000) "Characteristics, Covariances, and Average Returns: 1929 to 1997," *Journal of Finance*, 55, 389–406.

DeMiguel, V., A. Martın-Utrera, F. J. Nogales, and R. Uppal (2020) "A Transaction-Cost Perspective on the Multitude of Firm Characteristics," *Review of Financial Studies*, 33, 2180–2222.

Drechsler, I. and Q. F. Drechsler (2014) "The Shorting Premium and Asset Pricing Anomalies," NBER Working Paper 20282, National Bureau of Economic Research.

Easley, D., M. O'Hara, and P. Srinivas (1998) "Option Volume and Stock Prices: Evidence on Where Informed Traders Trade," *Journal of Finance*, 53, 431–465.

Engelberg, J., A. V. Reed, and M. Ringgenberg (2018) "Short Selling Risk," *Journal of Finance*, 73, 755–786.

Fong, K. Y. L., C. W. Holden, and O. Tobek (2017a) "Are Volatility Over Volume Liquidity Proxies Useful For Global Or US Research?" SSRN Working Paper.

Fong, K. Y. L., C. W. Holden, and C. A. Trzcinka (2017b) "What Are the Best Liquidity Proxies for Global Research?" *Review of Finance*, 21, 1355–1401.

Frazzini, A., R. Israel, and T. J. Moskowitz (2014) "Trading Costs of Asset Pricing Anomalies," Fama-Miller Working Paper No. 14-05, Chicago Booth.

Frazzini, A. and L. Pedersen (2022) "Embedded Leverage," *Review of Asset Pricing Studies*, 12, 1–52.

Frazzini, A. and L. H. Pedersen (2014) "Betting Against Beta," *Journal of Financial Economics*, 111, 1–25.

Gao, X. and J. R. Ritter (2010) "The Marketing of Seasoned Equity Offerings," *Journal of Financial Economics*, 97, 33–52.

Gârleanu, N. and L. H. Pedersen (2011) "Margin-Based Asset Pricing and Deviations from the Law of One Price," *Review of Financial Studies*, 24, 1980–2022.

Gkionis, K., A. Kostakis, G. Skiadopoulos, and P. S. Stilger (2021) "Positive Stock Information In Out-Of-The-Money Option Prices," *Journal of Banking & Finance*, 128, 106112.

Goncalves-Pinto, L., B. D. Grundy, A. Hameed, T. van der Heijden, and Y. Zhu (2020) "Why Do Option Prices Predict Stock Returns? The Role of Price Pressure in the Stock Market," *Management Science*, 66, 3903–3926.

Goyenko, R. Y., C. W. Holden, and C. A. Trzcinka (2009) "Do Liquidity Measures Measure Liquidity?" *Journal of Financial Economics*, 92, 153–181.

Green, J., J. R. M. Hand, and X. F. Zhang (2017) "The Characteristics That Provide Independent Information about Average U.S. Monthly Stock Returns," *Review of Financial Studies*, 30, 4389–4436.

Hasbrouck, J. (2009) "Trading Costs and Returns for U.S. Equities: Estimating Effective Costs from Daily Data," *Journal of Finance*, 64, 1445–1477.

He, H. and D. M. Modest (1995) "Market Frictions and Consumption-Based Asset Pricing," *Journal of Political Economy*, 103, 94–117.

Hitzemann, S., M. Hofmann, M. Uhrig-Homburg, and C. Wagner (2018) "Margin Requirements and Equity Option Returns," SSRN Working Paper.

Hou, K., S. Kim, and I. M. Werner (2016) "(Priced) Frictions," Fisher College of Business Working Paper Series 2016-19, Ohio State University.

Hou, K. and T. J. Moskowitz (2005) "Market Frictions, Price Delay, and the Cross-Section of Expected Returns," *Review of Financial Studies*, 18, 981–1020.

Hou, K., C. Xue, and L. Zhang (2015) "Digesting Anomalies: An Investment Approach," *Review of Financial Studies*, 28, 650–705.

——— (2018) "Replicating Anomalies," *Review of Financial Studies*, 33, 2019–2133.

Jarrow, R. (2016) "Testing for Asset Price Bubbles: Three New Approaches," *Quantitative Finance Letters*, 4, 4–9.

Jegadeesh, N. and S. Titman (2001) "Profitability of Momentum Strategies: An Evaluation of Alternative Explanations," *Journal of Finance*, 56, 699–720.

Jones, C. M. and O. A. Lamont (2002) "Short-Sale Constraints and Stock Returns," *Journal of Financial Economics*, 66, 207–239.

Jylhä, P. (2018) "Margin Requirements and the Security Market Line," *Journal of Finance*, 73, 1281–1321.

Martin, I. and C. Wagner (2019) "What Is the Expected Return on a Stock?" *Journal of Finance*, 74, 1887–1929.

Mehra, R. (2012) "Consumption-Based Asset Pricing Models," *Annual Review of Financial Economics*, 4, 385–409.

Muravyev, D., N. D. Pearson, and J. M. Pollet (2017) "Understanding Returns to Short Selling Using Option-Implied Stock Borrowing Fees," SSRN Working Paper.

——— (2022) "Is There a Risk Premium in the Stock Lending Market? Evidence from Equity Options," *Journal of Finance*, 77, 1787–1828.

Nagel, S. (2012) "Evaporating Liquidity," *Review of Financial Studies*, 25, 2005–2039.

Novy-Marx, R. and M. Velikov (2016) "A Taxonomy of Anomalies and Their Trading Costs," *Review of Financial Studies*, 29, 104–147.

Ofek, E., M. Richardson, and R. F. Whitelaw (2004) "Limited Arbitrage and Short Sales Restrictions: Evidence from the Options Markets," *Journal of Financial Economics*, 74, 305–342.

Pasquariello, P. (2014) "Financial Market Dislocations," *Review of Financial Studies*, 27, 1868–1914.

Patton, A. J. and A. Timmermann (2010) "Monotonicity in Asset Returns: New Tests with Applications to the Term Structure, the CAPM, and Portfolio Sorts," *Journal of Financial Economics*, 98, 605–625.

Roll, R. (1984) "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market," *Journal of Finance*, 39, 1127–1139.

Santa-Clara, P. and A. Saretto (2009) "Option Strategies: Good Deals and Margin Calls," *Journal of Financial Markets*, 12, 391–417.

Stambaugh, R. F., J. Yu, and Y. Yuan (2015) "Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle," *Journal of Finance*, 70, 1903–1948.

Stambaugh, R. F. and Y. Yuan (2017) "Mispricing Factors," *Review of Financial Studies*, 30, 1270–1315.

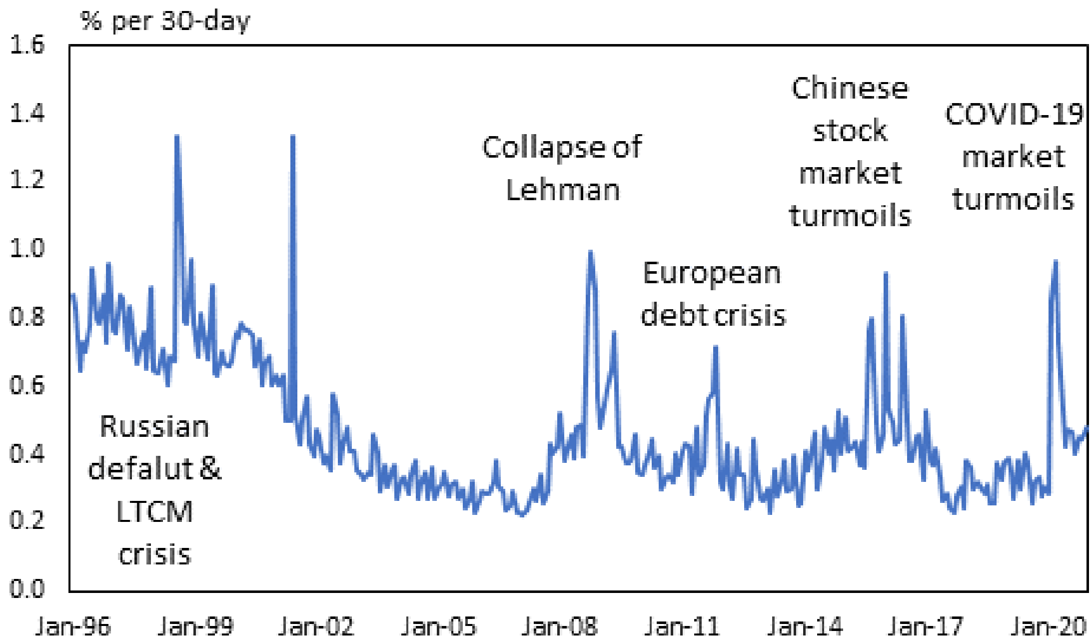

Figure 1. Time Series of the monthly IQR of SSD_t

Description: Figure 1 illustrates the time-series of the monthly interquartile range (IQR) of SSD_t. At the end of each month, we calculate the IQR of the individual stocks' SSD_t values as the difference between 75th and 25th percentile points of the distribution of SSD calculated from the cross-section of optionable common U.S. stocks. The unit of the y-axis is % per 30-day. The period spans January 1996-December 2020.

Interpretation: The dispersion of SSD widens during periods of distress as expected, given that transaction costs magnify over stock market crises.

Table 1. Spot-synthetic difference: Summary statistics

Description: Entries report the summary statistics of the spot-synthetic difference measure, SSD_t, at the end of each month. N is the total number of observations, s.d. is the standard deviation, skew is the skewness, kurt is the kurtosis, 5P and 95P are the 5th and 95th percentile points, respectively, med is the median, IQR is the interquartile range (75th minus 25th percentile values), and neg is the proportion of observations with a negative value. The estimation period spans January 1996-December 2020 (300 months). The unit of statistics (except skewness, kurtosis, and neg) is % per 30-day.

Interpretation: SSD can become sizable and its distribution has a longer left tail due to short-sale costs, as the theoretical range of TC-alpha, equation (5), suggests. Stocks subject to greater short-sale costs have more negative SSD.

N	mean	s.d.	skew	kurt	min	5P	med	95P	max	IQR	neg
432,837	-0.07	0.93	-2.15	73.4	-30.6	-1.24	-0.02	0.92	25.1	0.47	53.0

Table 2. Transition probability matrix of SSD bins

Description: This table reports the transition probability matrix of SSD bins. Bin 1 (3) contains stocks whose SSD value is very negative (positive), whereas Bin 2 contains stocks whose SSD value is close to zero. The (i, j) element of each panel represents the transition probability from bin i in month $t-1$ to bin j in month t. The unit is percentage.

Interpretation: SSD is not stable for certain stocks and its value can fluctuate significantly month by month. The variation in SSD is significant for stocks which have a big absolute value of SSD.

	1	2	3
1	39.4	29.6	31.0
2	26.5	43.7	29.7
3	30.9	32.9	36.2

Table 3. Characteristics of value-weighted decile portfolios sorted by SSD_t

Description: Entries report the average value of the proxies of transaction costs and related firms' and stocks' characteristics of decile portfolios formed by SSD_t. On the last trading day of each month t, stocks are sorted in ascending order based on SSD_t, and then value-weighted decile portfolios are formed. Then, we calculate the value-weighted value of each characteristic. See Appendix C for the detailed description of each variable. The unit of SSD_t to VoV(%Spread) is percent. Amihud and three Extended Amihud (EA) measures are multiplied by 10^9 so that the unit is basis point price impact caused by $10,000 order flow. The unit of S_t is dollar and that of $R_{t-1,t}$ is percent. B/M stands for book-to-market and N stands for the average number of stocks in each decile portfolio. The data period spans January 1996-December 2020 (300 months).

Interpretation: There is a U-shape relation between the cross-sectional SSD and each one of the employed proxies of transaction costs; a large value of SSD (either very positive and negative) is associated with greater transaction costs.

	Value-weighted decile portfolios sorted by SSD_t									
	1 (L)	2	3	4	5	6	7	8	9	10 (H)
SSD_t	-1.29	-0.50	-0.29	-0.17	-0.08	0.00	0.08	0.18	0.36	0.96
$range(SSD_t)$	2.82	1.49	1.06	0.88	0.80	0.77	0.82	0.95	1.33	2.56
$std(SSD_t)$	0.76	0.40	0.28	0.23	0.21	0.20	0.21	0.25	0.35	0.69
Bid-ask spread	0.47	0.37	0.32	0.29	0.27	0.25	0.25	0.28	0.32	0.41
Roll	1.30	1.10	1.02	0.94	0.94	0.94	0.96	0.98	1.06	1.23
High-Low	0.87	0.70	0.64	0.60	0.59	0.58	0.60	0.62	0.66	0.80
FHT	0.15	0.10	0.08	0.06	0.06	0.05	0.06	0.07	0.09	0.14
VoV(%Spread)	0.27	0.17	0.14	0.12	0.11	0.10	0.11	0.12	0.15	0.24
Amihud	3.16	0.96	0.57	0.33	0.23	0.20	0.23	0.34	0.73	2.44
EA FHT	0.49	0.13	0.06	0.04	0.02	0.02	0.02	0.03	0.09	0.38
EA High-Low	1.22	0.36	0.19	0.12	0.09	0.08	0.09	0.13	0.28	0.96
EA VoV	0.70	0.18	0.09	0.05	0.03	0.03	0.03	0.05	0.14	0.55
Size	15.45	16.44	16.96	17.30	17.56	17.61	17.58	17.30	16.72	15.81
B/M	0.53	0.46	0.44	0.42	0.41	0.40	0.40	0.42	0.45	0.50
S_t	42.38	64.00	81.90	85.99	97.08	95.20	95.86	75.50	62.09	46.64
$R_{t-1,t}$	2.62	2.33	2.12	1.95	1.75	1.72	1.39	1.12	0.59	-0.01
N	142.5	142.7	142.5	142.6	142.3	142.9	142.6	142.7	142.6	142.6

Table 4. Summary statistics of low-frequency liquidity proxies

Description: Entries report the summary statistics of the two dispersion statistics of SSD_t and low-frequency liquidity proxies. For each stock-month, we calculate the range (the difference between maximum and minumum) and the standard deviation of daily SSD_t data within the month. For the definition of the low-frequency liquidity proxies, See Appendix C. The sample period spans January 1996-December 2020. We restrict our sample to stock-month observations that have a valid SSD_t data. The unit of the two dispersion statistics of SSD_t and the proxies of effective spreads (FHT to BAS) is percent. The price impact proxies (Amihud and four Extended Amihud proxies) are multiplied by 10^{10}.

Interpretation: Both the dispersion statistics of SSD and low-frequency liquidity proxies have a large variation and a longer right tail.

	N	mean	s.d.	skew	5P	median	95P
$range(SSD_t)$	369,199	1.66	1.88	4.47	0.33	1.06	4.92
$std(SSD_t)$	369,199	0.45	0.50	4.34	0.09	0.29	1.33
FHT	419,825	0.14	0.36	5.98	0.00	0.00	0.76
VoV(%Spread)	419,825	0.32	0.26	2.39	0.08	0.24	0.81
Roll	419,821	1.29	1.76	2.79	0.00	0.78	4.57
High-Low	416,700	0.86	0.55	1.96	0.28	0.72	1.93
BAS	418,045	0.37	0.73	4.25	0.02	0.10	1.74
Amihud	419,844	49.87	233.87	59.97	0.40	8.23	200.63
EA (FHT)	419,825	7.34	60.24	36.05	0.00	0.00	23.19
EA (Roll)	419,821	27.66	148.82	28.99	0.00	1.00	112.00
EA (VoV)	419,825	13.28	82.23	38.28	0.03	1.10	48.33
EA (High-Low)	416,700	20.10	84.21	20.86	0.16	3.19	79.46

Table 5. Average cross-sectional correlations between the dispersion statistics of SSD and other proxies of ρ parameter

Description: Entries report the average cross-sectional Pearson correlation coefficient between the dispersion statistics of SSD_t, either the range or the standard deviation of daily SSD_t within each month, and each low-frequency liquidity proxy. The correlation coefficients are calculated as the time-series average of cross-sectional correlations in each month. The sample period spans January 1996-December 2020.

Interpretation: The average over time cross-sectional correlation between the monthly dispersion statistics of daily SSD and transaction costs is positive, for each proxy, and it can reach a high value (an example is the case of the VoV(%Spread) of Fong et al. (2017a)).

	$std(SSD_t)$	$range(SSD_t)$
FHT	0.24	0.23
VoV(%Spread)	0.62	0.59
Roll	0.17	0.17
High-Low	0.45	0.44
BAS	0.43	0.41
Amihud	0.44	0.41
EA (FHT)	0.28	0.26
EA (Roll)	0.32	0.30
EA (VoV)	0.35	0.33
EA (High-Low)	0.42	0.39

Table 6. Performance of value-weighted SSD_t-sorted portfolios: Bivariate dependent sorts controlling for relative bid-ask spread or RSI

Description: Entries in Panel A report the result of the bivariate dependent sort in value-weighted portfolios, where we first sort stocks based on the relative bid-ask spread (BAS), and then within each group of the BAS level, we further sort stocks into quintile portfolios by SSD_t. Rows correspond to the level of the first sorting variable, BAS, and the first to the fifth columns correspond to the level of the second sorting variable, SSD_t. Sixth to the last columns report the average returns, Carhart (1997) four-factor alpha, and the average SSD_t, respectively, of the spread portfolio (the highest SSD_t portfolio minus the lowest SSD_t portfolio). Entries in Panel B report the result, where we use RSI (relative short interest) as the first sorting variable, instead of BAS. The estimation period spans January 1996-December 2020 (300 months). t-statistics are adjusted for heteroscedasticity and autocorrelation (HAC-adjusted t-stat). The unit of the average returns and alphas (average SSD) is % per month (per 30-days).

Interpretation: The power of SSD to predict future stock returns is greater for stocks which are subject to greater transaction costs and short-sale constraints. In addition, Panel B suggests that severer short-sale constraints imply larger stock underperformance.

	α_{FFC} of SSD_t-sorted portfolios					Spread portfolio (5-1)		
	1 (L)	2	3	4	5 (H)	Ave ret	α_{FFC}	Ave SSD
Panel A: Relative bid-ask spread-sorted dependent bivariate sort								
BAS 1	-0.50	-0.27	0.01	-0.10	0.13	0.54	0.63	0.71
(narrow)	(-2.37)	(-1.65)	(0.10)	(-0.58)	(0.80)	(2.24)	(2.75)	(12.48)
2	-0.58	-0.21	0.04	0.15	0.24	0.83	0.81	0.98
	(-3.22)	(-1.49)	(0.25)	(1.08)	(1.11)	(3.33)	(3.26)	(9.34)
3	-0.60	-0.26	0.10	0.22	0.51	1.00	1.11	1.25
	(-3.50)	(-1.64)	(0.74)	(1.38)	(2.45)	(4.13)	(4.35)	(15.06)
4	-0.69	-0.12	-0.05	0.17	0.37	1.02	1.06	1.68
	(-3.14)	(-0.57)	(-0.31)	(0.97)	(1.99)	(4.21)	(4.01)	(15.50)
BAS 5	-1.70	-0.64	0.06	0.07	0.38	2.06	2.08	2.98
(wide)	(-6.05)	(-2.69)	(0.27)	(0.37)	(2.06)	(7.00)	(7.11)	(16.46)
Panel B: Relative short interest-sorted dependent bivariate sort								
RSI 1	-0.07	0.06	0.03	0.46	0.60	0.68	0.67	0.84
(small)	(-0.44)	(0.52)	(0.27)	(3.17)	(3.94)	(3.33)	(2.99)	(13.52)
2	-0.18	-0.03	-0.04	0.05	0.41	0.63	0.60	1.00
	(-1.13)	(-0.28)	(-0.26)	(0.36)	(2.11)	(2.88)	(2.40)	(19.61)
3	-0.26	-0.22	0.02	0.16	0.03	0.23	0.29	1.20
	(-1.42)	(-1.34)	(0.15)	(1.28)	(0.17)	(1.02)	(1.32)	(22.54)
4	-0.86	-0.13	0.11	0.03	0.02	0.82	0.88	1.39
	(-4.06)	(-0.68)	(0.61)	(0.17)	(0.12)	(3.31)	(3.48)	(22.61)
RSI 5	-1.38	-0.39	-0.17	-0.07	-0.10	1.30	1.29	1.88
(large)	(-5.54)	(-1.68)	(-1.04)	(-0.31)	(-0.42)	(4.65)	(4.38)	(16.77)

Table 7. Cross-sectional predictability of SSD_t

Description: Entries report the average SSD_t, average post-ranking return of underlying stocks, and results for the risk-adjusted underlying stock returns (α) of the SSD_t-sorted value-weighted decile portfolios and the spread portfolio, with respect to the Carhart (1997) four-factor model (FFC), the Stambaugh and Yuan (2017) mispricing factor model (SY), and the Hou et al. (2015) q-factor model (HXZq). On the last trading day of each month t, stocks are sorted in ascending order based on SSD_t, and then value-weighted decile portfolios are formed. Then, we calculate the underlying return of these portfolios and the spread portfolio in the succeeding month-($t+1$). The estimation period spans January 1996-December 2020. t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in parentheses. The unit of the average returns and alphas (average SSD_t) is percent per month (30-day). N is the average number of stocks in each decile portfolio.

Interpretation: SSD predicts stock returns in line with the theoretical setting of TC-alpha; the long (short) leg earns a positive (negative) alpha, and the long-short spread portfolio earns a positive and sizable alpha.

	Ave. SSD_t	Ave. Ret.	α_{FFC}	α_{SY}	α_{HXZq}	N
Value-weighted decile portfolios sorted by SSD_t						
1 (L)	-1.29	-0.04	-1.05	-0.78	-0.92	142.5
		(-0.08)	(-6.99)	(-4.67)	(-5.32)	
2	-0.50	0.41	-0.53	-0.51	-0.56	142.7
		(1.21)	(-3.58)	(-2.91)	(-3.63)	
3	-0.29	0.73	-0.17	-0.21	-0.17	142.5
		(2.39)	(-1.61)	(-1.77)	(-1.39)	
4	-0.17	0.65	-0.20	-0.20	-0.17	142.6
		(2.37)	(-2.20)	(-1.98)	(-1.66)	
5	-0.08	0.92	0.01	-0.08	-0.01	142.3
		(3.33)	(0.11)	(-0.80)	(-0.14)	
6	0.00	0.95	0.05	0.04	0.07	142.9
		(3.39)	(0.55)	(0.38)	(0.65)	
7	0.08	1.07	0.19	0.23	0.21	142.6
		(3.92)	(2.17)	(2.39)	(2.19)	
8	0.18	1.04	0.15	0.22	0.09	142.7
		(3.39)	(1.35)	(1.72)	(0.83)	
9	0.36	1.08	0.23	0.47	0.35	142.7
		(3.39)	(1.67)	(2.38)	(1.95)	
10 (H)	0.96	1.46	0.58	0.89	0.73	142.6
		(3.79)	(3.20)	(3.64)	(3.02)	
Long-short spread portflio (Port 10 - Port 1)						
Spread	2.25	1.49	1.63	1.67	1.65	—
		(6.37)	(6.67)	(4.96)	(5.67)	

Table 8. Pooled panel regressions: SSD_t and stock portfolio returns

Description: Entries reports the results of regressions $R_{t,T} - R^0_{t,T} = \alpha + \beta' f_{t,T} + \gamma SSD_t + \varepsilon_T$, where we employ the Carhart (1997) four-factor model (FFC), the Stambaugh and Yuan (2017) mispricing factor model (SY), and the Hou et al. (2015) q-factor model (HXZq), separately, as a set of risk-factors $f_{t,T}$. On the last trading day of each month t, stocks are sorted in ascending order based on SSD_t, and then value-weighted decile portfolios are formed. Then, we calculate the underlying return of each one of these portfolios and the respective portfolio's average SSD_t. Then, we run a pooled regression using these ten value-weighted portfolios' returns and SSD_t. The estimation period spans January 1996-December 2020. t-statistics are adjusted for heteroscedasticity and autocorrelation and reported in parentheses. The estimates of α (intercept) are multiplied by 100.

Interpretation: SSD is consistent with the cross-section of observed stock returns regardless of the chosen asset pricing model. The asset pricing equation under transaction costs setting, (4), cannot be rejected when SSD_t is used as a proxy of TC-alpha. This result renders further support to the fact that SSD_t accurately estimates TC-alpha.

	FFC	SY	HXZq
α	-0.02	0.08	0.03
	(-0.42)	(1.77)	(0.61)
γ	0.75	0.83	0.81
	(6.76)	(6.89)	(6.87)
β_{MKT}	1.04	1.03	1.05
	(100.63)	(84.16)	(80.91)
β_{SMB}	0.05	0.03	
	(3.29)	(1.67)	
β_{HML}	-0.00		
	(-0.24)		
β_{MOM}	-0.06		
	(-4.57)		
β_{MGMT}		-0.05	
		(-3.06)	
β_{PERF}		-0.06	
		(-4.75)	
β_{ME}			0.01
			(0.53)
β_{IA}			-0.00
			(-0.06)
β_{ROE}			-0.06
			(-2.36)
adj. R^2	0.85	0.84	0.84
	H_0: $\alpha = 0, \gamma = 1$		
p-value	0.025	0.155	0.117

School of Economics and Finance

This working paper has been produced by
the School of Economics and Finance at
Queen Mary University of London

School of Economics and Finance
Queen Mary University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 7356
Fax: +44 (0)20 8983 3580
Web: www.econ.qmul.ac.uk/research/workingpapers/