Seasonalities in Liquidity and Price Discovery[*]

Mark Kamstra[†] **Lisa Kramer**[‡] **Andriy Shkilko**[§]

November 6, 2024

Abstract: Liquidity costs and the flow of information into prices exhibit persistent yet distinct seasonal patterns. Using a multi-year sample of high-frequency U.S. data, we show that liquidity costs peak in early winter and are lowest in the spring. Meanwhile, price discovery is most intensive in early fall and least intensive in the spring. These patterns are associated with seasonal changes in market participants' impatience and risk aversion, which are in turn related to daylight exposure. Using low-frequency liquidity data from a number of markets around the world, we show that the seasonal liquidity patterns are observed globally and are predictably offset by six months in the Southern Hemisphere.

Key words: time-varying liquidity, adverse selection, seasonal behavioral effects

JEL: G14; G15; G41

---

[*]We are grateful to Ferhat Akbas, Bidisha Chakrabarty, Hugh Kim, Dmitriy Muravyev, seminar participants at the University of Mannheim and the University of California San Diego, and conference participants at the European Finance Association and Northern Finance Association meetings for their helpful comments. We acknowledge support of the Social Sciences and Humanities Council of Canada, and Shkilko acknowledges support of the Canada Research Chairs program. Alice Wang and Zi Wang provided excellent research assistance.

[†]York University, Canada, e-mail: mkamstra@schulich.yorku.ca

[‡]University of Toronto, Canada, e-mail: Lisa.Kramer@utoronto.ca

[§]Wilfrid Laurier University, Canada, e-mail: ashkilko@wlu.ca

1. Introduction

Two essential functions of financial markets are providing liquidity and facilitating price discovery. Liquid markets enable investors to enter and exit positions with minimal impact on prices, and through the interactions of multiple investors prices tend to reflect the relevant information about the underlying asset values. While academic research has extensively examined these two functions, little is known about their variation within the year. In this study, we demonstrate that both liquidity and price discovery exhibit persistent and distinct seasonal patterns, and we provide evidence on the origin of these patterns.

Our investigation is motivated by literature on seasonal fluctuations in risk aversion and impatience, which are driven by Seasonal Affective Disorder (SAD) and its less clinically severe counterpart, winter blues. Kamstra, Kramer, and Levi (2003) and Kamstra, Kramer, Levi, and Wermers (2017) show that these fluctuations lead to distinct seasonal patterns in asset returns and investor fund flows. Considering the significant roles that risk aversion and impatience play in liquidity provision and in the way information enters prices, liquidity and price discovery may too exhibit seasonal patterns.

To investigate the existence of these patterns, we analyze a sample of equity trading data from the U.S. and around the globe. We begin by focusing on two components of trading costs – the *price impact* and the *realized spread* – and show that both vary seasonally, although their variation patterns differ. The price impact is a proxy for the adverse selection cost that market makers incur when providing liquidity to informed market participants, and can also be interpreted as the extent to which information is incorporated into prices through trades rather than through quotes.[1] Figure 1 shows that price impacts are the highest in early fall and lowest in the spring,

[1] For illustrative purposes, consider a scenario where at t_0 the bid and offer quotes are \$9.99 and \$10.00, respectively. An informed trader purchases at the offer, driving the quotes up to \$10.01 and \$10.02 by t_1. A market maker who sold to the trader at \$10.00 needs to close her position but can only do so at \$10.01 or higher, thus incurring a loss of 1 cent or more per share. This loss is typically referred to as the *adverse selection cost*, and price discovery resulting from this event is considered *trade-driven*. If the market maker had adjusted the quotes to \$10.01 and \$10.02 without a trade from the informed trader, the price discovery would have been considered *quote-driven*.

an intra-year difference that amounts to 6.3% in our most conservative tests. Stated differently, information enters into prices more via trades in September than it does in April.

[Figure 1]

In turn, the realized spread captures the compensation required by market makers for bearing inventory risk and, more generally, for committing capital to the market-making operation. The magnitude of the realized spread variation throughout the year is similar to that of the price impact, at 6.2%, but this variation follows a different pattern; the realized spread is highest in early winter and lowest in early summer. This and all other patterns we document are consistent year after year, persisting when we control for conventional liquidity cost determinants as well as stock and year fixed effects.

Together, the price impact and realized spread add up to the effective spread, which represents the per-share trading cost incurred by liquidity demanders. This cost is the lowest in the spring and highest in winter, with the variation amounting to 4.5% for an average stock. Compared to estimates from the prior literature, the 4.5% figure surpasses the spread changes attributed to technological innovations in trading, such as colocation and microwave signal transmission, by more than twofold (Brogaard, Hagströmer, Nordén, and Riordan (2015) and Shkilko and Sokolov (2020a)). Viewed in terms of economic significance, trading the average sample stock in December costs investors $69,140 more than trading the same stock in April. This incremental cost amounts to hundreds of millions of dollars marketwide.[2]

To explain the drivers behind these patterns and the differences between the patterns for price impacts and realized spreads, we turn to psychology literature, which has linked seasonal variation in daylight hours to changes in risk aversion and impatience. As much as ten percent of the world population suffers from seasonal depression (or SAD) during the fall and winter, with much of the remainder of the population experiencing a less clinically severe analogue,

[2] To obtain the dollar figures, we multiply the dollar effective spread differential between December and April by the total number of shares traded in an average stock in a typical month.

winter blues. Onset of seasonal depression is typically in the early fall, recovery is typically in the spring, and it is well accepted by medical professionals that the primary cause of the seasonal variation is a reduction in daylight, as opposed to other environmental variables such as rainfall or cloud cover (Young, Meaden, Fogg, Cherin, and Eastman (1997)). Meanwhile, studies in psychology and neuroscience have long associated seasonality in depression with seasonality in risk aversion and impatience, finding that depressed individuals tend to exhibit higher risk aversion and greater impatience.[3] Furthermore, even individuals who have not been diagnosed with depression exhibit seasonal variation in mood and risk preferences on average, consistent with the widespread influence of sub-clinical winter blues (Kramer and Weber (2012)). This implies all market participants, not just those with a clinical depression diagnosis, may be prone to exhibit seasonally varying risk aversion.

The literature recognizes two patterns in seasonal depression that are important in our context. *Onset/Recovery* refers to the proportion of the population newly affected or newly recovering from depression, while *Incidence* refers to the currently affected share of the population. Figure 2 shows that Onset/Recovery peaks in early fall, when most sufferers experience their symptoms for the first time, and dips in early spring, when most people begin to recover. In contrast, Incidence is highest in winter, when everyone susceptible has become affected and the number of people experiencing symptoms is greatest, and lowest in the summer. Put differently, Onset/Recovery is a flow variable that captures the rate of inflow/outflow of sufferers, and Incidence is a stock variable that captures the current number of sufferers. Note that the Onset/Recovery pattern closely resembles the pattern identified for price impacts in Figure 1, while the Incidence pattern aligns with that of realized spreads in the same figure.

[Figure 2]

In the financial setting, Kamstra, Kramer, and Levi (2003), Kamstra, Kramer, Levi, and

[3] See, for instance, Pietromonaco and Rook (1987), Harlow and Brown (1990), Carton, Morand, Bungenera, and Jouvent (1995), and Amlung, Marsden, Holshausen, Morris, Patel, Vedelago, Naish, Reed, and McCabe (2019).

Wang (2014), and Kamstra, Kramer, Levi, and Wermers (2017) report statistically significant and economically large seasonal patterns in security returns and investment flows, consistent with depression-driven changes in market participants', including market professionals', risk aversion and impatience.[4] They show that as many people experience depression onset in early fall, their risk aversion increases, prompting them to reallocate investment holdings from riskier to safer assets, then they switch back when their symptoms subside in the spring. The reallocation pattern generates substantial fund flows and aligns with the Onset/Recovery pattern. Not only does the onset of seasonal depression increase risk aversion, but it also heightens impatience to establish new investment positions. Together, these factors likely result in traders moving in and out of positions while demanding liquidity in the fall and supplying liquidity in the spring.

Theory dating back to Stoll (1978), Ho and Stoll (1981, 1983), and Glosten and Milgrom (1985) suggests that the magnitude of liquidity costs (spreads) is influenced by both the impatience of informed traders and the risk aversion of market makers. Market structure literature generally recognizes that when a trader's impatience increases, she shifts her order mix from non-marketable limit orders (which provide liquidity and do not require immediate execution) to marketable orders (which demand liquidity and immediacy). Using stock-level data from the MIDAS database maintained by the U.S. Securities and Exchange Commission, we show that the share of marketable orders indeed increases in the fall and decreases in the spring. It is highest when market participants succumb to seasonal depression and begin impatiently adjusting portfolio holdings in the fall and lowest when market participants regain patience in the spring.

To further substantiate the potential link between informed trader impatience and seasonal depression, we use the Informed Trading Intensity (ITI) metric developed by Bogousslavsky, Fos, and Muravyev (2024). The ITI is a product of a machine learning technique trained to recognize informed trading in conventional datasets. It measures the rate at which information is assimilated

[4] Kamstra, Kramer, Levi, and Wang (2014) show in a representative agent equilibrium asset pricing model that seasonal variation in both impatience and risk aversion is necessary to match observed seasonality in equity and Treasury returns.

into prices and has two variations: the intensity of informed trading by those who seek to trade promptly and by those who do not typically place much emphasis on immediacy.

By design, the former metric captures informed trading on short-lived information (ITSL), while the latter captures informed trading on long-lived information (ITLL).[5] To clarify the difference between the two metrics, consider two investment institutions: a fundamentals-driven fund and an event-driven fund. The former generates alpha through fundamental analysis, while the latter profits from short-lived information about upcoming news announcements. The event-driven fund is typically time-constrained and tends to trade via marketable orders, whereas the fundamentals-driven fund is more flexible with timing and tends to trade via non-marketable orders.

Event-driven strategies are typically the domain of hedge funds, so we expect ITSL to primarily reflect hedge fund activity.[6] Hedge funds are known for restrictive redemption policies (Aragon, 2007) and are therefore considerably less affected by seasonal fund flows documented by Kamstra, Kramer, Levi, and Wermers (2017).[7] In the meantime, we expect ITLL to mainly reflect the activity of a mix of remaining institutions that pursue less time-sensitive investment strategies and experience seasonal asset reallocations by investors.

Examining seasonal variations in ITLL supports our expectations. The metric follows an inverse pattern relative to price impacts; it drops in early fall when investor reallocations and impatience cause fund managers to generate abnormal marketable order volumes. Conversely, when investors reallocate again in the spring, managers are more patient, leading to increased

[5] Bogousslavsky, Fos, and Muravyev (2024) refer to these metrics as ITI impatient and ITI patient, respectively. In their context, such labels are appropriate because they consider only one source of impatience – the amount of time before information becomes public. In our context, there is an additional source of impatience: seasonally varying depression. To avoid ambiguity, we refer to the two metrics as ITSL and ITLL.

[6] Almazan, Brown, Carlson, and Chapman (2004) report that many mutual funds are constrained in their use of short sales, deivatives, and margin trading – important components of event-driven investing. See also "Event-Driven Investing Strategies and Examples" by W. Kenton, April 21, 2022 (https://bit.ly/3LEC1DH) and "Understanding Event-Driven Investing" by BarclayHedge (https://bit.ly/3y0ImGE).

[7] See also “Investor Bulletin: Hedge Funds” by Securities and Exchange Commission, Office of Investor Education and Advocacy (https://bit.ly/3A9pfdR).

trading volume that is relatively heavy on non-marketable orders. These variations predictably result in higher price impacts in the fall and lower price impacts in the spring.

Meanwhile, ITSL is highest in early winter when the incidence of seasonal depression is high and lowest in the summer when the incidence is low, though its fluctuations are more subdued compared to ITLL. Overall, ITLL accounts for most seasonal variation in price impacts, with ITSL playing a minor role. ITSL's low magnitude is unsurprising; investors trading on short-lived information are already impatient due to the brief trading window available for executing positions. While seasonal depression may add to this impatience, the data suggest the incremental increase is not large.

In addition to the ITI metrics, we examine the Price Jump Ratio (PJR), which measures the efficiency of information incorporation into prices. Weller (2018) suggests that if informed trading forces prices to reflect corporate earnings before their announcements, price reactions to the announcements should be less pronounced, resulting in smaller PJRs. Brogaard, Hendershott, and Riordan (2019), Hagströmer and Menkveld (2023), and Kwan, Philip, and Shkilko (2024) demonstrate that marketable orders transmit information more effectively than non-marketable orders. In our context, increased use of marketable orders by impatient informed traders during the fall and winter should improve information incorporation. The data confirm this notion, with PJRs becoming smaller in fall and winter months.

Our discussion has so far focused on the price impacts and their seasonal variation. Next, we turn to realized spreads with an aim to understand fluctuations in the compensation market makers require for assuming inventory risk and committing capital. We propose that seasonal depression-induced risk aversion and impatience may impact realized spreads in three ways. First, in the course of providing liquidity, market makers accumulate inventories, which are subject to price fluctuation risk. Increased risk aversion requires greater compensation for this risk. Second, greater impatience leads to using more liquidity demanding orders to manage inventories, increasing costs. Finally, greater risk aversion demands greater compensation for the capital at risk.

The data support this reasoning, with realized spreads peaking when the incidence of seasonal depression is highest.

Seasonal fluctuations in price impacts and realized spreads lead to changes in both displayed and realized liquidity costs, measured by quoted spreads, depths, and effective spreads. While our main results derive from high-frequency data, they persist with low-frequency metrics such as end-of-day quoted spread and the effective spread estimator of Corwin and Schultz (2012). Furthermore, these results hold in the cross-section, with more pronounced effects in smaller stocks, where trades are less frequent, making timely executions via non-marketable orders less likely and inventory management more difficult. Consequently, increased seasonal risk aversion and impatience lead to stronger reactions in price impacts and realized spreads. Quoted and effective spreads exhibit a similar cross-sectional pattern.

Our conclusion that seasonality in depression is a key determinant of seasonality in spreads is supported by two additional pieces of evidence. First, we find that the seasonality of spreads in the international cross-section varies with latitude.[8] The most northern markets exhibit the largest seasonal variation in spreads, markets in the north subtropics exhibit relatively smaller variation, and markets located in the tropics exhibit virtually no seasonal variation. Second, just as the seasons are shifted by six months in the Southern Hemisphere, so is the seasonal pattern in spreads. That is, in the Northern Hemisphere spreads are widest in December, while in the Southern Hemisphere they are widest in May.

Our study contributes to several strands of market structure literature that examine the impacts of automation on liquidity costs and price discovery. O'Hara (2015) describes the contemporary trading landscape as highly automated and ultra-fast, where human abilities to react and process information are substantially surpassed by those of machines. Hendershott, Jones, and

[8] Because seasonal variation in light exposure is a key determinant of seasonality in depression, risk aversion, and impatience, several studies of financial market seasonality exploit variation in hours of daylight across different geographic latitudes in their empirical tests. These studies tend to find stronger seasonal variation in economic quantities the higher the latitude of the market.

Menkveld (2011) and Brogaard, Hagströmer, Nordén, and Riordan (2015) show that automation often leads to significant reductions in both price impacts and realized spreads, as algorithms are more efficient in avoiding adverse selection and managing inventory. Particularly notable in our context, Chakrabarty and Moulton (2012), Chakrabarty, Moulton, and Wang (2022) show that another important advantage of trading automation is the significant reduction of human attention constraints, ultimately leading to more efficient market making. It nonetheless remains unclear if automation has entirely eliminated the impacts of human behavior on the trading process.

Accordingly, our findings shed new light on the extent of automation reducing human effects in liquidity provision and demand. While the use of machines is certainly widespread and pervasive, the influence of humans remains important. Even the most tech-savvy trading firms rely on humans to set trading model parameters and calibrate algorithms.[9] In addition, humans periodically override system defaults. These interventions provide ample opportunity for human behavior to continue to exert a significant influence on liquidity generation and consumption, even in the age of machines.

A small related literature reports, but does not examine in detail, seasonality in liquidity in samples that precede the recent advent of automated trading. Like us, Chordia, Sarkar, and Subrahmanyam (2005) and Hameed, Kang, and Viswanathan (2010) find that spreads are narrowest in the summer. However, unlike us, these studies do not explore reasons for the seasonal variation, nor do they examine spread components such as price impacts and realized spreads, or metrics that capture the intensity of information incorporation into prices. Equally importantly, they study periods before modern automation, focusing on the 1990s and early 2000s when trading was still mainly conducted by humans.

[9]See, "The Intelligence Paradox: AI May Make Markets Less Rational," by A. Brynjolfsson and E. Brynjolfsson, The Wall Street Journal, January 31, 2024 (https://on.wsj.com/3Kr0mfS).

2. Data, sample, and metrics

Our data come from three sources. First, we use the Trade and Quote (TAQ) database to compute high-frequency intraday liquidity metrics for U.S. firms. These metrics include the quoted, effective, and realized spreads as well as price impacts. Second, we use CRSP and Datastream to compute the low-frequency alternatives to the TAQ quoted and effective spreads. These are discussed by Corwin and Schultz (2012) and Abdi and Ranaldo (2017).[10] The low-frequency metrics allow us to expand the analyses to non-U.S. markets, for which we do not have intraday data. In turn, these markets let us examine variation in SAD incidence patterns and severity as they vary across geographic latitudes. Finally, to compute market capitalization, returns, and volatility for the U.S. sample, we use data from the Center for Research in Security Prices (CRSP).

The sample period spans ten years, from 2010 through 2019. When selecting the sample of U.S. firms, we begin with 1,000 largest firms traded on the largest U.S. exchange, NYSE, as of January 2010 and drop those for which prices fall below \$5 or rise above \$500 at any time during the sample period. This procedure leaves us with the final sample of 939 firms.

2.1 Liquidity and impatience metrics

When analyzing the U.S. sample, we rely on conventional high-frequency metrics of displayed liquidity and trading costs. To examine displayed liquidity, we estimate the *quoted spread* as the difference between the lowest offer and the highest bid across all exchanges. Regulation requires that liquidity-seeking buy orders be sent to the exchange with the lowest offer quote (the National Best Offer) and sell orders to the exchange with the highest bid quote (the National Best Bid). Quoted spreads, often called the bid-ask spreads or National Best Bid and Offer (NBBO) spreads, capture liquidity costs based on posted prices and are among the most commonly studied

[10] Corwin and Schultz (2012) observe that an advantage of their spread estimator is its suitability for use across different markets with different market structures, which is useful in our context where we study spreads from countries around the world.

liquidity metrics. In addition to quoted spreads, researchers often estimate *quoted depths*, that is the number of shares available at the best bid and offer quotes. When market making costs increase, quoted spreads typically widen, and quoted depths decline, as market makers put smaller amounts of capital at risk.

While posted prices are commonly used as benchmarks, many traders time their liquidity consumption to periods when it is cheap to do so. Consequently, they often obtain average execution prices that are better than the average quoted spreads. In addition, execution prices may be better due to better-priced hidden liquidity or price improvement offered by liquidity providers. Still in some cases, liquidity demanders may receive prices worse than those posted, particularly if their demand exceeds the share quantities available at the best quotes. With these nuances in mind, we measure the actual trading costs incurred by liquidity demanders by computing the *effective spread*. This metric is typically computed as the difference between the traded price and the quote midpoint for trades initiated by buyers and as the difference between the midpoint and the traded price for trades initiated by sellers.

Quote midpoints used to compute the effective spreads are the averages of the best bid and ask prices. They are considered representations of the stock's intrinsic value at a given moment. For instance, if the quoted spread is \$9.99 on the bid and \$10.00 on the offer, the midpoint is \$9.995. A buyer who executes at \$10.00 pays \$0.005 per share more than the intrinsic value. The \$0.005 amount is the effective half-spread or the cost the buyer is willing to incur in exchange for immediacy.

The TAQ data do not directly distinguish between the buyer-initiated and seller-initiated trades. As is common, we infer trade direction using the Lee and Ready (1991) algorithm. This algorithm posits that trades with prices greater than the midpoint are likely buyer-initiated because impatient buyers are willing to pay for liquidity by accepting prices slightly above intrinsic values. Conversely, trades with prices below the midpoint are likely seller-initiated. For a small number of trades executed at midpoint prices, the algorithm copies the initiator from the previous trade. Despite its development in the early 1990s, the algorithm continues to be widely used today. Chakrabarty, Pascual, and Shkilko (2015) demonstrate its continued high efficacy in modern high-speed markets.

Early market structure research argues that informed traders tend to be impatient as they compete with others to incorporate short-lived information into prices (e.g., Glosten and Milgrom (1985), Kyle (1985)). Such traders use market and marketable orders that demand liquidity and immediacy. More recent studies show that the informed may also use limit orders, thus supplying liquidity (e.g., Kumar and Seppi (1994); Kaniel and Liu (2006); Goettler, Parlour, and Rajan (2009), Brolley and Malinova (2017), Roşu (2020), Bhattacharya and Saar (2021), Riccó, Rindi, and Seppi (2022)). When the informed seek liquidity, their trades push quotes in the direction of their information. Conversely, when they provide liquidity, quotes adjust in the direction of information without trades.

An impatient informed buyer typically places marketable buy orders until the price rises to a level at which further purchasing becomes unprofitable. The trades resulting from such orders are said to generate *price impact*. For buyer-initiated trades, price impact is computed as the difference between a future midpoint and the midpoint at the time of the trade. For seller-initiated trades, price impact is computed as the difference between the quote midpoint at the time of the trade and a future midpoint. In modern high-speed markets, quotes adjust to trades quickly, so we use a 60-second horizon for future midpoints. We, however, recognize that in a sample of over 900 securities, there are bound to be a few that have longer midpoint adjustment periods, so we use an additional 300-second horizon for robustness.

Price impacts hold significant importance in our analyses, as we anticipate them to increase when informed traders affected by seasonal depression display greater impatience. In addition to capturing impatience, price impacts represent an important market-making cost that factors into quoted and effective spreads. Known as adverse selection, it denotes the loss a market maker incurs while offering liquidity to an informed market participant. Consequently, more impatient informed trading when the incidence of seasonal depression is high will likely prompt market makers to widen spreads, compensating for the increased adverse selection cost.

The market structure literature is yet to explore behavioral changes in market participants' patience. In the meantime, the literature proposes several non-behavioral reasons for the informed to be impatient. These include (i) competition among traders whose information is homogeneous (Holden and Subrahmanyam (1992)), (ii) high information value that raises the opportunity costs of non-execution (Kaniel and Liu (2006)), and (iii) uncertainty of information revelation timing that increases the risk of information becoming public before an informed trader may act on it (Chau and Vayanos (2008)). We believe that these determinants of impatience are unlikely to change seasonally. It is difficult to imagine why, for instance, each year informed traders would obtain more valuable information in October and November as compared to April and May, or that such information would be more homogeneous and incite more competition.

Adverse selection is an important cost, but not the only one, incurred by market makers. Others include inventory and fixed costs. Inventory costs arise from non-zero inventory positions due to changes in asset prices. For example, when a market maker acquires a long position from a seller, even if the seller is uninformed, the position may lose value over time if the asset price falls. Market makers factor in the expected value of such losses and the liquidity costs of closing inventory positions into the price of liquidity. In turn, fixed costs primarily represent the expenses on sophisticated technology required for market making.

Capturing these two costs separately is not feasible using standard microstructure data. Consequently, the literature estimates them jointly as *realized spreads*, computed as the difference between effective spreads and price impacts. We note that it is unlikely that technology costs fluctuate seasonally, so seasonality in realized spreads is likely attributable to variation in inventory costs. In addition to reflecting inventory and fixed costs, realized spreads include market making profits (e.g., Hendershott, Jones, and Menkveld (2011) and Brogaard, Hagströmer, Nordén, and Riordan (2015)). As we mention earlier, seasonal depression tends to cause both increased impatience and risk aversion. Both these factors may affect realized spreads in our context. First, heightened impatience might prompt market makers to seek quicker ways to close out inventory positions, leading to greater position management costs and greater realized spreads.[11] Second, increased risk aversion may necessitate greater compensation for the expenses and risks associated with market making also resulting in larger realized spreads.

When computing the high-frequency metrics, we follow the procedure suggested by Holden and Jacobsen (2014). In addition, all high-frequency metrics are scaled by the corresponding quote midpoints to allow for comparability in the cross-section. Also, to make the effective and realized spreads, as well as price impacts, visually comparable to the quoted spreads, we multiply them by two. Finally, we drop the first and last five minutes of the trading day to reduce the effects of the opening and closing procedures.

Table 1 contains sample summary statistics, starting with stock characteristics in Panel A. The average stock has market capitalization of $14.709 billion and trades at $51.03 per share. The average daily volume of shares traded is nearly 2.5 million. There is a notable variability across sample stocks as should be expected from a sample of over 900 equities, with market capitalization ranging between $2.76 billion in the 25$^{\text{th}}$ percentile and over $13.7 billion in the 75$^{\text{th}}$ percentile. Prices and share volumes exhibit similar variations.

[Table 1]

When it comes to high-frequency liquidity metrics, in Panel B, we find that the average quoted spread is 7.46 bps, while the average effective spread is 5.79 bps. The effective spread captures trading costs incurred by traders who take liquidity. It is usually smaller than the quoted spread, because liquidity takers often come to the market when liquidity is cheaper and may also receive price improvement relative to the displayed quotes. In turn, the average 60-second price impact is 4.27 bps and increases to 4.63 bps when we extend the measurement horizon to

[11] An impatient market maker is more likely to use liquidity-demanding orders to manage positions. Such orders execute relatively quickly, but incur liquidity costs.

300 seconds. This result is expected, as information often drifts into prices for some time after the trade (Conrad and Wahal (2020)). Finally, the average realized spread is 1.51 bps at the 60-second horizon and 1.16 bps at the 300-second horizon. Similarly to stock characteristics, there is a non-trivial cross-sectional variation in liquidity metrics. In a later section, we explore this variation by examining the results in the cross-section.

While we have high-frequency data for the U.S., our liquidity analyses in a later section extend to other countries, requiring us to rely on low-frequency daily data. The literature has put forth several low-frequency liquidity proxies, including the end-of-day quoted spread denoted as *EOD*, which is computed as the difference between the closing bid and ask quotes scaled by the midpoint of these quotes. Additionally, two low-frequency estimators, as proposed by Corwin and Schultz (2012) and Abdi and Ranaldo (2017), labeled *CS* and *AR* respectively, have been shown to correlate with effective spreads. Notably, Abdi and Ranaldo (2017) demonstrate that the *EOD* quoted spread is the most accurate low-frequency liquidity proxy. Still, given that our high-frequency metrics differentiate between quoted and effective spreads, we include *CS* and *AR* alongside *EOD* for the sake of completeness.

Panel C of Table 1 contains summary statistics on three low-frequency liquidity metrics for the U.S. sample. In a later section, we demonstrate that these metrics successfully capture the SAD effects, similar to the high-frequency metrics, which enables us to extend the analysis to non-U.S. markets. We note that it is common for the low-frequency metrics to have magnitudes distinct from their high-frequency counterparts. The former metrics were designed to capture time-series and cross-sectional variations in liquidity rather than precisely represent the true spread values. Jahan-Parvar and Zikes (2023) find that *CS* and *AR* frequently yield estimates that are considerably larger than those from high-frequency data. Our data align with this finding. While the *EOD* quoted spread closely resembles the magnitude of its high-frequency counterpart (i.e., 5.76 bps compared to 7.46 bps), the *CS* and *AR* estimates are 96.27 and 64.62 bps, respectively. We reiterate that while the magnitude of these estimates is not the primary focus of our analyses, their ability to capture fluctuations in liquidity costs over time is crucial. As we demonstrate in a subsequent sections, the three metrics perform fairly well for this purpose.

2.2 Seasonal Affective Disorder metrics

We use two metrics related to SAD in this analysis: the proportion of SAD sufferers currently affected by SAD (which we refer to as *Incidence*) and the proportion of SAD sufferers newly affected or recovered from SAD (*Onset/Recovery*). We refer to the sum of these measures as SAD Composite.

2.2.1 SAD Incidence. To create the SAD Incidence variable, we adopt a metric based on the clinical timing of symptoms among people who experience seasonal depression. Young, Meaden, Fogg, Cherin, and Eastman (1997) and Lam (1998) studied hundreds of patients and recorded the date when each patient's SAD symptoms first arose in the late summer or fall and the date when their symptoms dissipated. We use the data sets made available by them to calculate the fraction of people susceptible to SAD who are actively exhibiting symptoms in a given month.

Following Kamstra, Kramer, and Levi (2015), we then use a spline function to smoothly interpolate the monthly variable to daily frequency, resulting in a daily measure of Incidence. The value of Incidence is zero in the summer, when virtually no one experiences depression symptoms. It increases most rapidly around the fall equinox in mid-September, when daylight hours are diminishing at their fastest rate, and the proportion of sufferers experiencing the onset of their symptoms is very high. Incidence peaks near 100% in winter, reflecting the fact that close to 100% of the people who are prone to suffer from seasonal depression have begun experiencing symptoms by the time winter begins. Finally the metric decreases most rapidly around the spring equinox in March, when hours of daylight are increasing at their fastest rate, and the proportion of recovering SAD-sufferers is very high, reaching a low of zero again the subsequent summer.

The SAD Incidence variable reflects the *stock* of people who are actively experiencing symptoms, and so we use it as a proxy for seasonal variation in these symptoms. Because this proxy measures the true incidence of SAD with error, using it directly could impart an errors-in-variables bias. To address this issue, we follow Kamstra, Kramer, Levi, and Wermers (2017) and use an instrumented version of the proxy which we produce as follows. After using a spline function to smoothly interpolate the monthly Incidence variable to daily frequency, we run a logistic regression of the daily Incidence measure on length of day. The fitted value from this regression yields the instrumented version of SAD Incidence.

2.2.2 SAD Onset/Recovery. To create the Onset/Recovery variable, we use the net flow of people becoming affected by or recovering from seasonal depression, i.e., the change in the proportion of people actively experiencing symptoms. We compute this metric by calculating the change in the Incidence variable. Onset/Recovery takes on its highest (positive) value around the fall equinox in September, when the number of sufferrers experiencing their symptoms for the first time that year is highest. Onset/Recovery takes on its lowest (negative) value around the spring equinox in March, when the rate of recovery is highest.

Kamstra, Kramer, Levi, and Wermers (2017) report that the onset and recovery from SAD symptoms lead to substantial seasonal variation in fund flows in September and March of each year. In the subsequent sections, we ask if these flows affect broader market liquidity.

2.2.3 SAD Composite. Summing SAD Incidence and Onset/Recovery yields the Composite variable intended to capture the combined effects of seasonally varying risk aversion and impatience. We primarily use the SAD Composite variable in figures, for illustration purposes.

3. Empirical results

To understand the relationship between seasonal depression and trading costs, it is essential to clarify its effects on the two cost components: price impact and realized spread. Figure 1 shows that while both components vary seasonally, their patterns differ. Therefore, in the subsequent sections, we examine each component separately.

3.1 Price impacts

As we discuss earlier, the increased impatience associated with SAD may prompt informed traders to rely more heavily on liquidity-demanding (marketable) orders, potentially leading to greater adverse selection of liquidity provider quotes and manifesting in greater price impacts. In Figure 1, trade price impacts indeed exhibit a distinct seasonal pattern that appears aligned with SAD Onset/Recovery in Figure 2.

The two figures, however, do not account for cross-sectional and time-series fixed effects. Nor do they consider well-known adverse selection determinants such as volume and volatility. Prior research (e.g., Hendershott, Jones, and Menkveld (2011), O'Hara and Ye (2011)) reports associations between these two variables and price impacts and therefore using them as controls appears warranted. More specifically, greater volatility is often linked to greater adverse selection, while greater volume – holding volatility constant – is associated with uninformed trading and therefore lower adverse selection.[12]

To account for these effects, we conduct a formal regression analysis by estimating the following model for each stock i on each day t:

$$\begin{aligned} DepVar_{i,t} = {} & \alpha_i + \gamma_{year} + \beta_1 Onset/Recovery_t + \beta_2 Incidence_t \\ & + \beta_3 Volume_{i,t-1} + \beta_4 Volatility_{i,t-1} + \varepsilon_{i,t}, \end{aligned} \tag{1}$$

where $DepVar$ is the price impact in stock i on day t, $Onset/Recovery$ is the onset/recovery variable, $Incidence$ is the incidence variable, $Volume$ is the lagged natural logarithm of daily

[12] Controlling for volume and volatility also allows us to account for variation in investor disagreement. Kandel and Pearson (1995) and Banerjee and Kremer (2010) propose that periods of disagreement are often associated with high volume and volatility.

number of shares traded, and *Volatility* is lagged volatility computed as the standard deviation of intraday midquotes.[13] We estimate this model using ordinary least squares, controlling for stock and year fixed effects, and clustering the standard errors by firm and date.

The results appear in Table 2, with price impacts measured at the 60-second and 300-second horizons. Price impacts increase significantly in seasonal depression, and this effect comes mainly from the Onset/Recovery variable. Although Incidence is significant at the 5% level in the Full models in columns [2] and [4], its coefficient is considerably smaller than that of Onset/Recovery and the maximum increase (decrease) in price impact occurs in September (March), reflecting the predominance of SAD flows impacting market transactions.

[Table 2]

Overall, these results are consistent with the notion that when the value of the SAD Onset/Recovery variable is high, informed trader impatience increases, and they tilt their order submission mix to marketable orders. These orders, combined with increased reallocation flows, in turn increase adverse selection of liquidity provider quotes. The overall price impact arising from Onset/Recovery and Incidence peaks around September (when the value of the Onset/Recovery variable reaches its annual maximum of 0.38, and the value of Incidence also happens to be about 0.38), with price impacts increasing by about 0.2 basis points for each of models [1] through [4].[14] This translates into an overall seasonal variation in price impact of about 6% in all cases.[15]

[13]Our results are qualitatively unchanged if we use contemporaneous values of volume and volatility or an alternate measure of volatility based on the difference between the highest and lowest daily prices scaled by the average of the two prices.

[14]For this and subsequent sets of tabulated regression results, the basis point changes for a given month are calculated by taking (i) the value of the SAD Incidence variable in that month times the SAD Incidence coefficient estimate plus (ii) the value of the SAD Onset/Recovery variable in that month times the SAD Onset/Recovery coefficient estimate.

[15]For this and subsequent sets of tabulated regression results, the economic magnitudes are calculated as follows. For a given month, we compute (i) the Onset/Recovery coefficient estimate times the value of the Onset/Recovery variable and (ii) the Incidence coefficient estimate times the value of the Incidence variable. We then (iii) sum those two products for each month, (iv) compute the difference between the maximum and minimum sums over the year, and (v) divide that difference by the mean spread value from Table 1.

3.2 Prevalence of marketable orders

As we suggested earlier, increased impatience driven by the onset of seasonal depression likely results in greater use of marketable orders, especially when markets experience increases in investment flows in early fall. Conversely, as sufferers recover in early spring and investment flows rise again, lower levels of impatience should lead to a reduced use of marketable orders.

We are not aware of a comprehensive multi-year U.S. dataset that would allow researchers to observe orders by type, so we develop a proxy for marketable order prevalence. For much of our sample period (2012-2019), the MIDAS dataset maintained by the U.S. Securities and Exchange Commission provides monthly traded volume and order volume statistics for each stock. It is plausible that while most marketable orders result in trades, a smaller share of non-marketable orders do. Therefore, dividing traded volume by the difference between order volume and traded volume approximates the prevalence of marketable orders.

Panel D of Table 1 suggests that the prevalence is about 0.03. This figure aligns with similar findings in recent studies. For instance, using Canadian data, Brogaard, Hendershott, and Riordan (2019) report a prevalence of 0.05 in 2012-2013. Considering the more recent nature of our sample period and the fact that the U.S. market is significantly more fragmented than the Canadian market, leading to more frequent order revisions as traders adjust to signals from multiple venues, it is conceivable that the U.S. prevalence figure would be slightly lower than its Canadian counterpart.

Finally, as we show in Table 3, marketable order prevalence, MKTBL, is indeed positively related to the Onset/Recovery variable. Note that when estimated jointly with Onset/Recovery, the Incidence variable acquires a negative sign. This indicates that MKTBL declines a little more sharply from its September peak to the March trough compared to the pattern shown in Figure 1 for price impacts.

[Table 3]

3.3 Informed trading intensity

Bogousslavsky, Fos, and Muravyev (2024) propose a set of metrics that allows researchers to gauge the patience with which informed traders open their positions. The authors train a machine learning algorithm to recognize informed trading using an observed sample of activist investor trades and obtain a set of non-linear combinations of variables that determine the prevalence of informed trading. Subsequently, they use these variable combinations to compute informed trading intensity (ITI) for the universe of stock-days. Importantly, they obtain two additional metrics based on periods when informed investors trade more aggressively, pressed for time, and periods when they trade less aggressively when time is abundant.

The latter two metrics capture what the literature commonly refers to as informed trading on short-lived information (ITSL) and informed trading on long-lived information (ITLL). For instance, traders who are concerned that their information may quickly become known by others tend to trade relatively aggressively. Conversely, traders who rely on information that others are unlikely to discover until much later have the luxury of trading slowly and stealthily, often through limit orders, with the aim of avoiding detection by the rest of the market. The former trading style is typical for event-driven hedge funds, whose informational advantage is often very short-lived, while the latter style more accurately reflects trading by traditional investors such as mutual funds.

By definition, trading on short-lived information should be conducted rather aggressively. It is possible that this aggressiveness further intensifies in SAD Incidence, as the numbers of impatient traders using event-driven strategies grows. However, such an increase is contingent on the aggressiveness not already being at a high level. Examining this possibility is an empirical exercise that we will delve into shortly.

Meanwhile, trading on long-lived information, as represented by mutual funds and similar institutions, is likely to be the most susceptible to SAD Onset/Recovery. Shares in most mutual funds can be redeemed at will throughout the year, while redemptions from hedge funds are considerably more restricted. Kamstra, Kramer, Levi, and Wermers (2017) show that changes in risk aversion prompt many mutual fund investors to redeem shares in riskier funds in September in exchange for cash or shares in less risky funds. Naturally, the redemptions, combined with depression-driven impatience, lead fund managers to generate a lot of volume that is more aggressive than usual. Conversely, when fund investors restore their positions in the spring, managers tend to be more patient, and the resulting increase in trading volume is predominantly driven by limit orders. So, when Onset/Recovery is at its highest, ITLL should be at its lowest and vice versa.

Table 3 examines how ITLL and ITSL correlate with the SAD variables.[16] The data confirm our expectations; ITSL loads primarily on Incidence, and ITLL loads negatively on Onset/Recovery. We observe that among all the SAD-ITI combinations, the pairing of Onset/Recovery with ITLL produces the most significant economic effect, with SAD Incidence and ITSL following as a distant second. These results are consistent with the earlier findings, indicating that Onset/Recovery has a much more substantial impact on price impacts than Incidence. In other words, the seasonal variations in price impact are primarily driven by fluctuations in patience among those trading on long-lived information in September and again in March.

3.4 Information incorporation into prices

As market participants research firm fundamentals, value-relevant information flows into prices through their trading activity. The more impatient such market participants are, the more direct price pressure they create, and the more likely prices will reflect their information. In the case of earnings, the more information is incorporated into prices prior to an announcement, the smaller should be the market reaction to the announcement itself.

To measure this effect, Weller (2018) introduces the price jump ratio, PJR, that divides the

[16] We thank Dmitriy Muravyev for sharing the ITI data with us.

earnings announcement return by the total return plausibly attributable to the announcement. The latter includes three weeks of pre-announcement price changes. A low PJR is consistent with high levels of price discovery, as it implies that a substantial portion of earnings information is incorporated into prices in the weeks prior to the announcement. In our setting, if informed trader impatience indeed increases in SAD, PJR should decline during the fall and winter months.

To compute PJR, we follow Weller (2018) and let T be the earnings announcement date. We then define the *announcement window* as $[T-1, T+2]$, *event window* as $[T-21, T+2]$, and *pre-event window* as $[T-255, T-90]$. For each day t and each stock i, we compute the close-to-close return, r_{it}, and the return on each the market index, r_{mt}. We then obtain the abnormal return, abr_{it}, as the difference between the stock i return on day t and the expected return according to the market model estimated in the pre-event window, that is,

$$abr_{it} = r_{it} - \hat{\alpha}_i - \hat{\beta}_i r_{mt}. \tag{2}$$

Next, we define cumulative abnormal return as the sum of abnormal returns from t_1 to t_2,

$$CAR_i^{t_1,t_2} = \sum_{t=t_1}^{t_2} abr_{it}, \tag{3}$$

and compute PJR as the ratio of the announcement-window CAR and the event-window CAR,

$$PJR_i = 100 * \frac{CAR_i^{T-1,T+2}}{CAR_i^{T-21,T+2}}. \tag{4}$$

One notable implementation issue when computing PJR is that the denominator of the metric may occasionally be close to zero. To account for this issue, Weller (2018) drops the announcements for which the absolute event-window CAR is smaller than $\sqrt{24}\sigma_i$, where σ_i is the standard deviation of r_i over the preceding month. We do the same.

To reiterate, if informed investors shift their order submissions from non-marketable to marketable limit orders due to the increased impatience induced by seasonal depression, we anticipate a more efficient incorporation of information into prices. The results presented in the last two columns of Table 3 align with our expectations. As informed trading becomes less patient during in the fall and winter months, information is incorporated into prices more efficiently, resulting in smaller PJRs.

3.5 Realized spreads

As we mention previously, increased risk aversion and impatience associated with seasonal depression may lead liquidity providers to require additional compensation for assuming inventory risk and for committing capital to the market making operation. We examine this possibility in Table 4. For the Base models in columns [1] and [3], the 60-second and 300-second realized spreads increase significantly with the Incidence variable but not with the Onset/Recovery variable. Results are similar for the Full models, in columns [2] and [4], which control for volatility and volume, although for the 300-second horizon Onset/Recovery becomes statistically significant at the 5% level, and negative, which somewhat dampens the impact from Incidence in the fall and amplifies it in winter. In December, when the value of the Incidence variable is close to one and the value of Onset/Recovery is close to zero, the realized spreads increase by 10 to 12 basis points for each of models [1] through [4]. The overall seasonal variation in realized spreads is about 8 to 10% for the Base models and about 6 to 8% for the Full models. For each model, untabulated tests strongly reject the null (at the 0.1% level) that Onset/Recovery and Incidence coefficient estimates are equal.

[Table 4]

3.6 Trading costs

Having established the relationships between SAD and the trading cost components, we proceed to examine trading costs themselves. The results in Table 5 are consistent with our expectations, in that quoted and effective spreads increase with Onset/Recovery and Incidence. In the month of December, when the value of SAD Incidence is nearly one and the value of SAD Onset/Recovery is close to zero, we calculate that the quoted spreads increase by about 0.4 bps and the effective spreads increase by 0.2 bps in Base regressions (columns [1] and [3]). The estimates are slightly smaller in Full regressions that control for the effects of volume and volatility (columns [2] and [4]). When it comes to quoted depth (columns [5] and [6]), it decreases consistently with our expectations.

We note that the economic magnitudes of the changes in liquidity metrics, reported in the bottom row of the table, are consistent with those observed in Base results, even after controlling for the effects of volatility and volume. Specifically, based on the full models, quoted spreads vary by 5.5%, effective spreads vary by 4.5%, and quoted depths vary by 7.1%. For quoted spreads, effective spreads, and quoted depth, untabulated tests fail to reject the null that the Onset/Recovery and Incidence coefficient estimates are equal at the 23.2%, 4.2%, and 30% levels respectively. While this implies that we could restrict the coefficients on these two variables to be identical and use SAD Composite as our main metric, this null is very strongly rejected for some of the dependent variables we consider below.

[Table 5]

3.7 Low-frequency liquidity metrics

In a later section, we expand our analysis to international markets because the magnitude of the seasonal depression and its seasonality should exhibit considerable variation across geographic latitudes. Locations closer to the equator receive less variable amounts of light exposure during the year and therefore people living in such locations experience seasonal depression to a lesser extent. Also, the timing of the SAD cycle in countries located in the Southern Hemisphere is six months removed from that in the Northern Hemisphere. These variations allow us to verify if the effects documented in the United States extend to other jurisdictions and to confirm that they are less likely to be driven by confounding factors.

For the international markets we lack high-frequency data, so we must instead resort to the low-frequency proxies. These include the end-of-day (EOD) quoted spreads, the Corwin-Schultz (CS) effective spread estimator, and the Abdi-Ranaldo (AR) effective spread estimator. Abdi and Ranaldo (2017) show that when the quote data are available, the EOD spreads are the most reflective of liquidity conditions. Even though these low-frequency estimators have been shown to work in previous research, we would like to test whether they pick up the same seasonal patterns as those picked up by the high-frequency metrics. To do so, in this section we repeat the earlier analyses using the low-frequency metrics for US data.

Figure 3 shows that seasonal correlations between the low-frequency metrics and the Composite variable resemble those identified earlier for the high-frequency spread metrics. That is, both the low-frequency metrics and the SAD Incidence variable dip in late spring and peak in late fall. In turn, the equation 1 results appear in Table 6. The Onset/Recovery and Incidence coefficient estimates are strongly statistically significant for the EOD and CS spreads, like they were for the high-frequency spread estimates, while insignificant for the AR spread. Due to dropping negative spread estimates as prescribed by the AR method, we have only a third as many observations for AR compared to the EOD case, a shortfall which may explain the lack of power to identify the SAD effect here.[17]

The magnitude of the seasonal changes are again large. For example, in September, when the value of the Onset/Recovery variable is at its annual high, the end-of-day spreads increase

[17] When we compute AR for international markets in a later section, the share of negative AR estimates is smaller, and the metric performs better in regression models.

by about 0.1 basis points in the Base model and in the Full model. In December, when SAD Incidence is nearly 1 and SAD Onset/Recovery is close to zero, the end-of-day spreads increase by about 0.2 basis points for the Base model and 0.1 basis points for the Full model. On a proportional basis, the seasonal variation in EOD spreads due to the SAD Onset/Recovery and SAD Incidence variables is roughly 4% for the Base model and 2% for the Full model. The seasonal variation in CS spreads is about 7 to 8%. Altogether, it appears that both the high-frequency and at least two out of three low-frequency proxies are sufficiently sensitive to identify the seasonal relations between liquidity costs and seasonal depression.

[Figure 3 and Table 6]

3.8 Cross-sectional analysis

To explore cross-sectional differences in the U.S. data, we split our sample into three groups on the basis of firm size, re-sorted daily based on the previous day's market capitalization. Tercile 1 contains the largest firms, and tercile 3 contains the smallest firms. Summary statistics appear in Table 7. The group of largest firms has a mean market capitalization above $28 billion, and the group of the smallest firms has a mean market capitalization below $2.5 billion. The high- and low-frequency liquidity cost metrics are consistently smallest for tercile 1 and increase as firm size decreases. This result is anticipated, as the costs related to providing liquidity are higher in smaller stocks due to increased information asymmetries, longer inventory holding periods, and higher fixed costs per share resulting from lower trading volumes (e.g., Dyhrberg, Shkilko, and Werner (2023)).

[Table 7]

We examine the relationship between seasonal depression and the various liquidity metrics for each of the terciles in Tables 8 and 9. With an increase in impatience and risk aversion, measured by Onset/Recovery and Incidence, we anticipate that price impacts and realized spreads will increase more in the stocks where they hold greater significance, and where achieving timely executions via non-marketable orders is more difficult. For instance, in small stocks where information asymmetries are relatively high and informed traders' incorporation of information into prices is more pronounced, a rise in the impatience of informed traders should lead to a more significant increase in price impacts compared to their larger counterparts, where information asymmetries are lower. Likewise, in the case of smaller stocks where trading volumes are relatively low, establishing informed positions and managing inventory via non-marketable orders is more challenging, and therefore an increase in impatience and risk aversion should lead to a more substantial increase in price impacts and the expected compensation for inventory costs and committing capital.

In the results presented earlier, we find that of the two SAD variables, price impact loads primarily on Onset/Recovery and realized spreads load primarily on Incidence. Hence for the tercile analysis, we present models that include only the most relevant SAD measure. Models incorporating both SAD measures yield similar results. Price impact and realized spreads appear in Table 8. As expected, price impact increases more with Onset/Recovery for small firms (tercile 3) than for large firms (tercile 1). The pattern is monotonic through the terciles and statistically significant in all terciles, with coefficients of 0.180 pbs, 0.326 bps, and 0.484 bps for the large through small terciles with price impacts measured at 60-second horizons and similar estimates for the 300-second horizons. In the month of September when the value of Onset/Recovery is highest, price impacts for the set of smallest firms increase by about 0.2 bps for both models [1] and [2]. The annual seasonal variation in price impact for small firms is 6.5% relative to the average value of price impacts for tercile 3 for price impacts measured at both the 60- and 300-second horizons. It is somewhat smaller for terciles 1 and 2, at about 5% and 6%, respectively.

Turning to the realized spreads measured at 60-second horizons, they increase by 0.022 bps, 0.117 bps, and 0.195 bps for the large through small terciles, respectively, with similar figures for the 300-second horizons. Statistical significance of the Incidence coefficient is observed at the 5% level or better for all but the largest-firm tercile. The overall annual seasonal variation in realized spreads for the smallest firms is 7.4% relative to the mean value of realized spreads for the 60-second horizon and 9.6% for the 300-second horizon. It is 4.2% to 4.5% for tercile 2.

[Table 8]

Table 9 contains regression results for the high-frequency quoted and effective spreads, quoted depth, the low-frequency end-of-day quoted spreads, Corwin-Schultz effective spreads, and Abdi-Ranaldo effective spreads. Overall, consistent with our expectations, the tercile results suggest that the impatience and risk aversion associated with SAD have greater economic impact on the spreads of smaller firms compared to larger firms.

[Table 9]

3.9 International liquidity metrics

To provide further evidence identifying the effect of SAD on spreads, we turn our attention to the analysis of data from markets located in countries other than the United States. SAD varies in intensity and prevalence based on latitude, and therefore by considering spreads data from markets around the world at different latitudes, we can test the identification of spread seasonality arising due to seasonal light exposure.

We consider a collection of large, broad-based markets that provide representation across different latitude groupings that span the globe. The group furthest to the north is the northern temperate zone, located at latitudes above 40 degrees north. Exchanges in Norway, Germany, the United Kingdom, France, Canada, and Italy are located in this zone. The northern sub-tropics region spans 23.5 degrees north to 40 degrees north, and markets in China, Israel, Japan, and Hong Kong are located in this region. The tropical zone is between 23.5 degrees north and 23.5 degrees south, and includes Brazil, Thailand, the Philippines, and Indonesia. Finally, the southern sub-tropics and temperate zone countries, at latitudes 23.5 degrees south and higher, are New Zealand, Argentina, Australia, Chile, and South Africa.

For each country in our sample, we collect stock-level data from Datastream for all available firms, yielding millions of firm-day observations for each latitude grouping: 9 million for the most northern group, over 15 million for the northern subtropics, and about 3 million for the tropics region and the southern sub-tropics/temperate zone. Summary statistics appear in Table 10; more granular summary statistics, on a country-by-country basis, are tabulated in an online appendix (Table A1).

[Table 10]

Starting with the stock characteristics in Table 10, we see the average firm market capitalization, converted to U.S. dollars, is over $1.7 billion for the northern temperate zone and northern subtropic groupings, and is a little below $1 billion for the tropics and southern sub-tropics/temperate zone regions. The average share price is highest for the most northern latitude group at $13.90 and drops monotonically through the groups to a low of $2.47 for the most southern latitude group.

Turning to formal analysis of the spreads, we estimate equation 1 for each of the four regions. Results appear in Table 11. Panels A, B, C, and D correspond to the northern temperate region, the northern sub-tropics, the tropics, and the southern sub-tropics/temperate zones respectively. That is, results appear from furthest north to furthest south. For the southern region, we shift the Onset/Recovery and Incidence variables by six months to adjust for the fact that the timing of daylight exposure in the Southern Hemisphere is offset by six months relative to the Northern Hemisphere. In the interest of brevity, we present results for the Full models only; results based on the Base models are qualitatively similar.

[Table 11]

In Panel A, which covers the northern temperate region, we see the low-frequency spreads measures vary significantly with one or both of the SAD measures. In September, when SAD/Onset recovery is at its annual peak, the end-of-day spreads increase by about 2 bps, CS spreads increase by about 1 bp, and AR spreads increase by about 3 bps. In Panels B and C, the northern sub-tropics and tropics, we mostly find no discernible SAD effect. This is expected in light of the fact that medical research finds the effects of SAD are most noticeable at latitudes above 40 degrees. The southern regions in Panel D, a blend of sub-tropical and temperate countries, exhibit significantly increased spreads with seasonal depression in all cases. We note that the negative coefficient on Onset/Recovery does not alter the overall impact of seasonal depression on spreads. Rather, this coefficient leads to a delayed impact; however, the peak spreads are still during the southern winter. Overall, the international results are consistent with those observed based on U.S. data.

4. Robustness

To supplement our main analysis, we performed a variety of robustness checks. For the 2010 to 2019 sample period, we considered a broader cross-section including all NYSE and Nasdaq firms. We also considered various NYSE data starting in 1926 and Nasdaq data starting in 1972, examining various sub-periods, including pre-2000, the 1990s specifically, and the 2000s. Importantly, the sign and significance of the Onset/Recovery and Incidence variables is similar to that found on our primary sample of NYSE firms over 2010 to 2019. The stability of findings across these various sample periods is at odds with the perhaps intuitive idea that markets ought to become less prone to human influence with the introduction of algorithmic trading. We find no evidence that the widespread use of algorithms reduces the presence of seasonal effects in liquidity costs. The fact that humans write the code underlying algorithmic trading and/or often manually override the recommendations of programs may help explain this finding.

We also explored using contemporaneous volume and volatility in places of lagged values in our regression models, and found similar results. We included local weather variables, such as precipitation and temperature, as additional control variables and found similar results. Including a turn-of-the year indicator variable as a control in the regression models also yielded similar results.

Furthermore, we considered alternate liquidity cost metrics, including the Amihud (2002) lliquidity measure and the ratio of marketable orders to non-marketable orders, and found significant SAD-related seasonality in both cases.

The SAD-related seasonal effect remains evident when we add various ad hoc variables to our empirical models, such as a January indicator, a turn-of-the-year indicator, or summer vacation indicator variables. The SAD-related effect is also robust to inclusion of various weather variables, such as precipitation, daily low temperature, and daily high temperature, remaining economically large and statistically significant, losing only about a third of its magnitude for all the dependent variables we consider. In our models with liquidity cost dependent variables, precipitation, when statistically significant, has a positive coefficient estimate, consistent with the work of Shkilko and Sokolov (2020b), and daily low temperature typically has a negative statistically significant coefficient estimate. In our model with depth as the dependent variable, precipitation is negative and statistically significant, consistent with Shkilko and Sokolov (2020b).[18]

Potential concerns about non-stationarity of our dependent variables of interest are addressed two-fold. First, the dependent variable in many of our models is a percentage difference in bid and ask prices, comparable to a rate of return. To the extent any of our dependent variables nevertheless exhibit autocorrelation or heteroskedasticity, our use of standard errors clustered by firm and date mitigates the issue.

[18]We are grateful to Lai and Dzombak (2019) for the weather data. Unfortunately their data set does not include cloud cover or sunshine; we view precipitation as a reasonable proxy for both.

5. Conclusion

This study documents distinct seasonal patterns in liquidity costs and price discovery. These patterns are driven by seasonal fluctuations in risk aversion and impatience, influenced by Seasonal Affective Disorder (SAD) and its clinically milder counterpart, winter blues. We show that adverse selection, one of the largest costs of market making, is highest in the early fall and lowest in the spring. Meanwhile, realized spreads, representing market maker required compensation for inventory risk and committed capital, peak in early winter and dip in early summer. These seasonal variations in trading costs persist when controlling for conventional liquidity cost determinants and stock and year fixed effects.

Our evidence is consistent with the notion that seasonal depression leads to increased impatience and risk aversion, prompting greater use of marketable orders in fall and winter. This behavior, combined with increased fund flows, aligns with greater price impacts during these periods. In contrast, spring sees a return to patience and a shift toward non-marketable orders, resulting in lower price impacts. As a result, information is incorporated into prices more efficiency in the fall and winter months than in the spring and summer months.

The robustness of the results is corroborated by their predictability across different latitudes. Northern markets show larger seasonal variations in spreads compared to tropical markets, where sunlight exposure is relatively stable throughout the year, and the spread pattern shifts by six months in the Southern Hemisphere.

References

Abdi, F., and A. Ranaldo, 2017, "A simple estimation of bid-ask spreads from daily close, high, and low prices," *Review of Financial Studies*, 30(12), 4437–4480. 9, 14, 25, 39, 45, 51

Almazan, A., K. C. Brown, M. Carlson, and D. A. Chapman, 2004, "Why constrain your mutual fund manager?," *Journal of Financial Economics*, 73(2), 289–321. 5

Amihud, Y., 2002, "Illiquidity and stock returns: Cross-section and time series effects," *Journal of Financial Markets*, 5(1), 31–56. 31

Amlung, M., E. Marsden, K. Holshausen, V. Morris, H. Patel, L. Vedelago, K. R. Naish, D. D. Reed, and R. E. McCabe, 2019, "Delay discounting as a transdiagnostic process in psychiatric disorders: A meta-analysis," *JAMA Psychiatry*, 76(11), 1176–1186. 3

Aragon, G. O., 2007, "Share restrictions and asset pricing: Evidence from the hedge fund industry," *Journal of financial economics*, 83(1), 33–58. 5

Banerjee, S., and I. Kremer, 2010, "Disagreement and learning: Dynamic patterns of trade," *The Journal of Finance*, 65(4), 1269–1302. 17

Bhattacharya, A., and G. Saar, 2021, "Limit Order Markets under Asymmetric Information," *Working Paper*, University of Chicago and Cornell University. 11

Bogousslavsky, V., V. Fos, and D. Muravyev, 2024, "Informed trading intensity," *Journal of Finance*. 4, 5, 20, 41

Brogaard, J., B. Hagströmer, L. Nordén, and R. Riordan, 2015, "Trading fast and slow: Colocation and liquidity," *The Review of Financial Studies*, 28(12), 3407–3443. 2, 8, 12

Brogaard, J., T. Hendershott, and R. Riordan, 2019, "Price discovery without trading: Evidence from limit orders," *Journal of Finance*, 74, 1621–1658. 6, 19

Brolley, M., and K. Malinova, 2017, "Informed trading in a low-latency limit order market," *Working Paper*, Wilfrid Laurier University. 11

Carton, S., P. Morand, C. Bungenera, and R. Jouvent, 1995, "Sensation-seeking and emotional disturbances in depression: relationships and evolution," *Journal of Affective Disorders*, 34(3), 219–225. 3

Chakrabarty, B., and P. C. Moulton, 2012, "Earnings announcements and attention constraints: The role of market design," *Journal of Accounting and Economics*, 53(3), 612–634. 8

Chakrabarty, B., P. C. Moulton, and X. F. Wang, 2022, "Attention: How high-frequency trading improves price efficiency following earnings announcements," *Journal of Financial Markets*, 57, 100690. 8

Chakrabarty, B., R. Pascual, and A. Shkilko, 2015, "Evaluating trade classification algorithms: Bulk volume classification versus the tick rule and the Lee-Ready algorithm," *Journal of Financial Markets*, 25, 52–79. 11

Chau, M., and D. Vayanos, 2008, "Strong-form efficiency with monopolistic insiders," *Review of Financial Studies*, 21(5), 2275–2306. 12

Chordia, T., A. Sarkar, and A. Subrahmanyam, 2005, "An empirical analysis of stock and bond market liquidity," *The Review of Financial Studies*, 18(1), 85–129. 8

Conrad, J., and S. Wahal, 2020, "The term structure of liquidity provision," *Journal of Financial Economics*, 136(1), 239–259, Publisher: Elsevier. 14

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(2), 719–760. 7, 9, 14, 39, 45, 51

Dyhrberg, A. H., A. Shkilko, and I. M. Werner, 2023, "The retail execution quality landscape," *Working Paper*, Wilfrid Laurier University. 26

Glosten, L. R., and P. R. Milgrom, 1985, "Bid, ask and transaction prices in a specialist market with heterogeneously informed traders," *Journal of Financial Economics*, 14(1), 71 – 100. 4, 11

Goettler, R. L., C. A. Parlour, and U. Rajan, 2009, "Informed traders and limit order markets," *Journal of Financial Economics*, 93(1), 67–87. 11

Hagströmer, B., and A. J. Menkveld, 2023, "Trades, quotes, and information shares," *Available at SSRN 4356262*. 6

Hameed, A., W. Kang, and S. Viswanathan, 2010, "Stock market declines and liquidity," *The Journal of Finance*, 65(1), 257–293. 8

Harlow, W., and K. C. Brown, 1990, "Understanding and assessing financial risk tolerance: A biological perspective," *Financial Analysts Journal*, 46(6), 50–62. 3

Hendershott, T., C. M. Jones, and A. J. Menkveld, 2011, "Does Algorithmic Trading Improve Liquidity?," *Journal of Finance*, 66(1), 1–33. 7, 12, 17

Ho, T. S. Y., and H. R. Stoll, 1981, "Optimal dealer pricing under transactions and return uncertainty," *Journal of Financial Economics*, 9, 47–73. 4

——— , 1983, "The dynamics of dealer markets under competition," *Journal of Finance*, 38(4), 1053–1074. 4

Holden, C. W., and S. Jacobsen, 2014, "Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions," *Journal of Finance*, 69(4), 1747–1785. 13

Holden, C. W., and A. Subrahmanyam, 1992, "Long-lived private information and imperfect competition," *Journal of Finance*, 47(1), 247–270. 12

Jahan-Parvar, M. R., and F. Zikes, 2023, “When do low-frequency measures really measure transaction costs?,” *Review of Financial Studies*. 14

Kamstra, M. J., L. A. Kramer, and M. D. Levi, 2003, “Winter blues: A SAD stock market cycle,” *American Economic Review*, 93(1), 324–343. 1, 3

Kamstra, M. J., L. A. Kramer, and M. D. Levi, 2015, “Seasonal variation in Treasury returns,” *Critical Finance Review*, 4(1), 45–115. 15

Kamstra, M. J., L. A. Kramer, M. D. Levi, and T. Wang, 2014, “Seasonally varying preferences: Theoretical foundations for an empirical regularity,” *The Review of Asset Pricing Studies*, 4(1), 39–77. 3, 4

Kamstra, M. J., L. A. Kramer, M. D. Levi, and R. Wermers, 2017, “Seasonal asset allocation: Evidence from mutual fund flows,” *Journal of Financial and Quantitative Analysis*, 52(1), 71–109. 1, 4, 5, 16, 21

Kandel, E., and N. D. Pearson, 1995, “Differential interpretation of public signals and trade in speculative markets,” *Journal of Political Economy*, 103(4), 831–872. 17

Kaniel, R., and H. Liu, 2006, “So what orders do informed traders use?,” *Journal of Business*, 79(4), 1867–1914. 11, 12

Kramer, L. A., and J. M. Weber, 2012, “This is your portfolio on winter: Seasonal affective disorder and risk aversion in financial decision making,” *Social Psychological and Personality Science*, 3(2), 193–199. 3

Kumar, P., and D. J. Seppi, 1994, “Information and index arbitrage,” *The Journal of Business*, 67(4), 481–509. 11

Kwan, A., R. Philip, and A. Shkilko, 2024, “The conduits of price discovery: A machine learning approach,” *Working Paper*, University of New South Wales. 6

Kyle, A. S., 1985, "Continuous Auctions and Insider Trading," *Econometrica*, 53(6), 1315–1335. 11

Lai, Y., and D. A. Dzombak, 2019, "Use of historical data to assess regional climate change," *Journal of Climate*, 32(14), 4299–4320. 31

Lam, R. W., 1998, "Seasonal affective disorder: diagnosis and management," *Primary Care Psychiatry*, 4, 63–74. 15

Lee, C., and M. Ready, 1991, "Inferring trade direction from intraday data," *Journal of Finance*, 46(2), 733–746. 10

O'Hara, M., 2015, "High frequency market microstructure," *Journal of Financial Economics*, 116(2), 257–270. 7

O'Hara, M., and M. Ye, 2011, "Is market fragmentation harming market quality?," *Journal of Financial Economics*, 100(3), 459–474. 17

Pietromonaco, P. R., and K. S. Rook, 1987, "Decision style in depression: The contribution of perceived risks versus benefits.," *Journal of Personality and Social Psychology*, 52(2), 399. 3

Riccó, R., B. Rindi, and D. J. Seppi, 2022, "Information, Liquidity, and Dynamic Limit Order Markets," *Working Paper*, Bocconi University. 11

Roşu, I., 2020, "Liquidity and information in limit order markets," *Journal of Financial and Quantitative Analysis*, pp. 1792–1839. 11

Shkilko, A., and K. Sokolov, 2020a, "Every Cloud Has a Silver Lining: Fast Trading, Microwave Connectivity and Trading Costs," *Journal of Finance*, 75(6), 2899–2927. 2

——— , 2020b, "Every cloud has a silver lining: Fast trading, microwave connectivity and trading costs," working paper, Available at SSRN: https://ssrn.com/abstract=2848562. 31

Stoll, H. R., 1978, "The supply of dealer services in securities markets," *Journal of Finance*, 33(4), 1133–1151. 4

Weller, B. M., 2018, "Does algorithmic trading reduce information acquisition?," *Review of Financial Studies*, 31(6), 2184–2226. 6, 21, 22, 41

Young, M. A., P. M. Meaden, L. F. Fogg, E. A. Cherin, and C. I. Eastman, 1997, "Which environmental variables are related to the onset of seasonal affective disorder?," *Journal of Abnormal Psychology*, 106(4), 554. 3, 15

Table 1
Summary Statistics

The table reports summary statistics for the sample period staring in January 2010 through December 2019. The data are from CRSP and TAQ databases. The top portion of the table, Panel A, contains summary statistics for stock characteristics such as market capitalization, share price, daily trading volume, and volatility. Panel B reports on high-frequency liquidity metrics obtained from TAQ, including quoted, effective, and realized spreads as well as price impacts. We compute price impacts and realized spreads for two horizons, 60 and 300 seconds after the trade. Panel C reports on three low-frequency liquidity metrics, including the end-of-day (EOD) quoted spread computed using CRSP quotes as well as two effective spread estimators proposed by Corwin and Schultz (2012) and Abdi and Ranaldo (2017), respectively, CS and AR. Panel D contains summary statistics on other metrics, including quoted depth (in share hundreds), prevalence of marketable orders, variables related to informed trading intensity, and the price jump ratio. When aggregating, we first compute the averages of all variables for each stock and then compute sample characteristics across stocks. The full sample contains over 2.1 million stock-day observations. For some variables, fewer than the full sample number of observations are available, most notably CS and AR which discard negative estimates and price jump ratio which is based on periods immediately surrounding earnings announcements.

	Mean	St. dev.	Median	25th	75th
Panel A: Stock characteristics:					
Market capitalization, $ millions	14,709	29,121	5,509	2,760	13,785
Price, $	51.03	41.56	42.27	27.56	61.41
Volume, thousands of shares	2,594	5,160	1,270	615	2,832
Natural log volume	12.30	2.11	12.59	11.10	13.74
Volatility	11.58	37.40	8.13	5.36	11.79
Panel B: High-frequency liquidity metrics, bps:					
Quoted spread	7.46	16.71	5.31	3.61	7.91
Effective spread	5.79	11.51	4.30	2.91	6.09
Price impact, 60s	4.27	2.57	3.77	2.66	5.04
Price impact, 300s	4.63	3.36	3.96	2.68	5.41
Realized spread, 60s	1.51	9.89	0.40	0.14	1.03
Realized spread, 300s	1.16	9.23	0.30	0.14	0.67
Panel C: Low-frequency liquidity metrics, bps:					
EOD quoted spread	5.76	16.26	3.41	2.44	5.54
CS effective spread	96.27	34.53	87.75	72.16	110.86
AR effective spread	64.62	20.73	60.38	50.11	73.77
Panel D: Other metrics:					
Quoted depth	31.72	102.95	9.39	5.55	162.98
Marketable orders	0.03	0.01	0.03	0.02	0.03
Informed trading intensity	0.294	0.022	0.294	0.284	0.304
Informed trading intensity, long-lived	0.213	0.027	0.210	0.201	0.220
Informed trading intensity, short-lived	0.432	0.018	0.432	0.425	0.440
Price jump ratio	45.58	117.03	44.92	26.36	63.59

Table 2
Trade Price Impacts

The table examines the relation between the SAD metrics and price impacts. The sample period spans January 2010 through December 2019. We compute price impacts for two horizons, 60 and 300 seconds after the trade. The reported coefficients are obtained from the regression of the following form:

$$DepVar_{i,t} = \alpha_i + \gamma_{year} + \beta_1 Onset/Recovery_t + \beta_2 Incidence_t + \beta_3 Volume_{i,t-1} + \beta_4 Volatility_{i,t-1} + \varepsilon_{i,t}.$$

DepVar is the price impact in stock i on day t, *Onset/Recovery* is the SAD onset/recovery variable, *Incidence* is the SAD incidence variable, *Volume* is the lagged natural logarithm of daily number of shares traded, and *Volatility* is the lagged quote-based intraday volatility (expressed as a standard deviation). In specifications [1] and [3], we report the results from the Base models that do not include the control variables. In specifications [2] and [4], we report the results from the Full models with control variables. The models are estimated using ordinary least squares, controlling for stock and year fixed effects, and the standard errors are clustered by firm and date. The magnitudes reported in the bottom row of each panel are calculated as described in footnote 15. *** and ** indicate statistical significance at the 1% and 5% levels. The sample contains over 2.1 million stock-day observations.

	60 seconds		300 seconds	
	Base	Full	Base	Full
	[1]	[2]	[3]	[4]
Onset/Recovery	0.347***	0.353***	0.390***	0.379***
	(0.055)	(0.051)	(0.056)	(0.053)
Incidence	0.102***	0.066**	0.102***	0.065**
	(0.030)	(0.028)	(0.030)	(0.030)
Volatility		0.028***		0.036***
		(0.009)		(0.011)
Volume		0.371***		0.174***
		(0.054)		(0.064)
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Adj. R^2	0.48	0.49	0.44	0.45
Magnitude as a % of Mean	6.3%	6.3%	6.5%	6.3%

Table 3
Informed Trading Intensity and Information Incorporation into Prices

The table examines the relationship between the SAD metrics, the prevalence of marketable orders (MKTBL), and the two proxies for informed trading intensity (ITI) proposed by Bogousslavsky, Fos, and Muravyev (2024), and the price jump ratio (PJR) proposed by Weller (2018). MKTBL is computed as the ratio between traded volume and the difference between order volume and traded volume from the SEC MIDAS database. The ITI proxies are obtained using a machine learning technique trained on a sample of informed institutional transactions and extrapolated to the entire stock-day universe. *ITLL* and *ITSL* distinguish between informed trading on long-lived information and short-lived information. In turn, PJR is computed as the return immediately surrounding an earnings announcement divided by the return that includes three weeks preceding the announcement,

$$PJR_i = \frac{CAR_i^{T-1,T+2}}{CAR_i^{T-21,T+2}},$$

where $CAR_i^{T-1,T+2}$ is the cumulative market-adjusted return for the announcement i from day $T-1$ to day $T+2$, with T being the announcement date, and $CAR_i^{T-21,T+2}$ is the same metric computed from day $T-21$ to day $T+2$. The sample period spans January 2010 through December 2019. The reported coefficients are obtained from the regression of the following form:

$$DepVar_{i,t} = \alpha_i + \gamma_{year} + \beta_1 Onset/Recovery_t + \beta_2 Incidence_t + \beta_3 Volume_{i,t-1} + \beta_4 Volatility_{i,t-1} + \varepsilon_{i,t}.$$

DepVar is one of the three above-mentioned ITI metrics or the PJR metric in stock i on day t, *Onset/Recovery* is the SAD onset/recovery variable, *Incidence* is the SAD incidence variable, *Volume* is the lagged natural logarithm of daily number of shares traded, and *Volatility* is the lagged quote-based intraday volatility (expressed as a standard deviation). For each dependent variable, we report the results from the Base regression model, which does not include the control variables, and the Full model, which includes the control variables. The models are estimated using ordinary least squares, controlling for stock and year fixed effects, and the standard errors are clustered by firm and date. The magnitudes reported in the bottom row of the table are calculated as described in footnote 15. *** and ** indicate statistical significance at the 1% and 5% level.The ITI sample contains over 1.4 million stock-day observations, whereas the PJR sample is stock-earnings announcement based and therefore contains fewer, 95.5 thousand, observations.

	MKTBL		ITLL		ITSL		PJR	
	Base	Full	Base	Full	Base	Full	Base	Full
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Onset/Recov.	0.081*	0.197***	−0.021***	−0.018***	−0.009**	−0.006	−2.812	3.180
	(0.043)	(0.047)	(0.003)	(0.003)	(0.004)	(0.004)	(3.435)	(3.353)
Incidence	0.021	−0.072**	−0.003	−0.004**	0.012***	0.009***	−1.569	−3.457***
	(0.025)	(0.028)	(0.002)	(0.002)	(0.003)	(0.002)	(1.370)	(1.305)
Volatility		0.013***		0.000***		0.000**		0.071
		(0.002)		(0.000)		(0.000)		(0.046)
Volume		1.057***		0.048***		0.069***		17.274***
		(0.020)		(0.005)		(0.007)		(1.728)
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.35	0.52	0.02	0.05	0.03	0.12	0.05	0.07
Magnitude as a % of Mean	0.8%	0.2%	0.5%	0.2%	<0.1%	<0.1%	5.3%	8.0%

Table 4
Realized Spreads

The table examines the relation between the SAD metrics and realized spreads. The sample period spans January 2010 through December 2019. We compute realized spreads for two horizons, 60 and 300 seconds after the trade. The reported coefficients are obtained from the regression of the following form:

$$DepVar_{i,t} = \alpha_i + \gamma_{year} + \beta_1 Onset/Recovery_t + \beta_2 Incidence_t + \beta_3 Volume_{i,t-1} + \beta_4 Volatility_{i,t-1} + \varepsilon_{i,t}.$$

DepVar is the realized spread in stock i on day t, *Onset/Recovery* is the SAD onset/recovery variable, *Incidence* is the SAD incidence variable, *Volume* is the lagged natural logarithm of daily number of shares traded, and *Volatility* is the lagged quote-based intraday volatility (expressed as a standard deviation). In specifications [1] and [3], we report the results from the Base models that do not include the control variables. In specifications [2] and [4], we report the results from the Full models with control variables. The models are estimated using ordinary least squares, controlling for stock and year fixed effects, and the standard errors are clustered by firm and date. The magnitudes reported in the bottom row of each panel are calculated as described in footnote 15. *** and ** indicate statistical significance at the 1% and 5% levels. The sample contains over 2.1 million stock-day observations.

	60 seconds		300 seconds	
	Base	Full	Base	Full
	[1]	[2]	[3]	[4]
SAD Onset/Recovery	0.038	−0.042	−0.003	−0.066**
	(0.038)	(0.032)	(0.039)	(0.033)
SAD Incidence	0.127***	0.099***	0.128***	0.100***
	(0.031)	(0.032)	(0.030)	(0.031)
Volatility		0.056***		0.048***
		(0.010)		(0.008)
Volume		−0.758***		−0.563***
		(0.156)		(0.141)
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Adj. R^2	0.58	0.61	0.52	0.53
Magnitude as a % of Mean	8.0%	6.2%	10.4%	8.3%

Table 5
Displayed Liquidity and Trading Costs

The table examines the relationship between the SAD metrics, quoted spreads, quoted depths, and effective spreads. The sample period spans January 2010 through December 2019. The reported coefficients are obtained from the regression of the following form:

$$DepVar_{i,t} = \alpha_i + \gamma_{year} + \beta_1 Onset/Recovery_t + \beta_2 Incidence_t + \beta_3 Volume_{i,t-1} + \beta_4 Volatility_{i,t-1} + \varepsilon_{i,t}.$$

DepVar is the effective or quoted spread, or quoted depth in stock i on day t, *Onset/Recovery* is the SAD onset/recovery variable, *Incidence* is the SAD incidence variable, *Volume* is the lagged natural logarithm of daily number of shares traded, and *Volatility* is the lagged quote-based intraday volatility (expressed as a standard deviation). In specifications [1], [3], and [5], we report the results from the Base models that do not include the control variables. In specifications [2], [4], and [6], we report the results from the Full models with control variables. The models are estimated using ordinary least squares, controlling for stock and year fixed effects, and the standard errors are clustered by firm and date. The magnitudes reported in the bottom row of the table are calculated as described in footnote 15. *** and ** indicate statistical significance at the 1% and 5% level. The sample contains over 2.1 million stock-day observations.

	Quoted spread		Effective spread		Quoted depth	
	Base	Full	Base	Full	Base	Full
	[1]	[2]	[3]	[4]	[5]	[6]
Onset/Recovery	0.548***	0.434***	0.385***	0.311***	−1.679**	−1.088
	(0.081)	(0.071)	(0.060)	(0.056)	(0.835)	(0.850)
Incidence	0.412***	0.322***	0.230***	0.165***	−2.025***	−2.370***
	(0.062)	(0.067)	(0.043)	(0.046)	(0.707)	(0.706)
Volatility		0.122***		0.085***		0.011
		(0.026)		(0.018)		(0.016)
Volume		−0.647***		−0.397**		10.011***
		(0.245)		(0.181)		(1.665)
Firm FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Adj. R^2	0.64	0.67	0.62	0.65	0.67	0.67
Magnitude as a % of Mean	7.0%	5.5%	5.7%	4.5%	6.5%	7.1%

Table 6
Low-Frequency Liquidity Metrics

The table examines the relation between the SAD metrics and each of three low-frequency liquidity proxies: the end of day spread (EOD), which proxies for displayed liquidity, the Corwin-Schultz (CS) metric – a proxy for trading costs, and the Abdi-Ranaldo (AR) metric – also a proxy for trading costs. The sample period spans January 2010 through December 2019. The reported coefficients are obtained from the regression of the following form:

$$DepVar_{i,t} = \alpha_i + \gamma_{year} + \beta_1 Onset/Recovery_t + \beta_2 Incidence_t + \beta_3 Volume_{i,t-1} + \beta_4 Volatility_{i,t-1} + \varepsilon_{i,t}.$$

DepVar is one of the three above-mentioned low-frequency metrics in stock i on day t, $Onset/Recovery$ is the SAD onset/recovery variable, *Incidence* is the SAD incidence variable, *Volume* is the lagged natural logarithm of daily number of shares traded, and *Volatility* is the lagged quote-based intraday volatility (expressed as a standard deviation). For each dependent variable, we report the results from the Base regression model that does not include the control variables and from the Full model that includes the control variables. The models are estimated using ordinary least squares, controlling for stock and year fixed effects, and the standard errors are clustered by firm and date. The magnitudes reported in the bottom row of the table are calculated as described in footnote 15. *** and ** indicate statistical significance at the 1% and 5% level. The EOD sample contains over 2.1 million stock-day observations, and the CS and AR samples contain about 1.3 million and 0.7 million observations respectively because both the CS and AR methods discard negative estimates.

	EOD		CS		AR	
	Base	Full	Base	Full	Base	Full
	[1]	[2]	[3]	[4]	[5]	[6]
Onset/Recovery	0.200***	0.126**	8.096***	9.844***	3.040	3.573
	(0.060)	(0.054)	(2.108)	(1.876)	(3.138)	(3.109)
Incidence	0.188***	0.113**	4.229***	3.234***	3.508*	2.741
	(0.054)	(0.054)	(1.396)	(1.219)	(2.043)	(2.010)
Volatility		0.092***		0.392**		0.555***
		(0.013)		(0.169)		(0.077)
Volume		−0.273		26.418***		13.768***
		(0.188)		(1.857)		(1.619)
Firm FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Adj. R^2	0.54	0.55	0.19	0.23	0.05	0.06
Magnitude as a % of Mean	3.7%	2.3%	7.1%	8.1%	5.7%	5.3%

Table 7
Cross-Sectional Summary Statistics

The table reports summary statistics for the data sorted daily into size terciles over the sample period January 2010 through December 2019. The data are from CRSP and TAQ databases. Panel A corresponds to the largest firms (tercile 1), Panel B corresponds to smaller firms (tercile 2), and Panel C corresponds to the smallest firms (tercile 3). Summary statistics appear for the following stock characteristics: market capitalization, share price, daily trading volume, and volatility. Summary statistics also appear for the following low-frequency liquidity metrics: the end-of-day (EOD) quoted spread, the Corwin and Schultz (2012) (CS) effective spread, and the Abdi and Ranaldo (2017) (AR) effective spread. When aggregating, we first compute the averages of all variables for each stock and then compute sample characteristics across stocks. Each tercile contains over over 700,000 stock-day observations. For some variables, fewer than the full sample number of observations are available, most notably CS and AR which discard negative estimates.

Panel A: Tercile 1

	Mean	St. dev.	Median	25th	75th
Stock characteristics:					
Market capitalization, $ millions	28,077	36,746	16,938	11,588	27,517
Price, $	70.36	49.92	58.54	40.02	83.04
Volume, thousands of shares	4,241	7,759	2,359	1,344	4,326
Volatility	5.93	4.67	5.32	3.95	6.72
High-frequency liquidity metrics, bps:					
Quoted spread	4.10	3.31	3.34	2.73	4.58
Effective spread	3.31	2.66	2.73	2.22	3.58
Price impact, 60s	2.94	1.68	2.53	2.06	3.34
Price impact, 300s	2.96	1.81	2.50	2.03	3.36
Realized spread, 60s	0.37	1.64	0.16	0.05	0.41
Realized spread, 300s	0.35	1.45	0.20	0.08	0.40
Low-frequency liquidity metrics, bps:					
EOD quoted spread	3.21	4.83	2.21	1.77	3.23
CS effective spread	84.59	25.32	80.00	67.09	97.52
AR effective spread	58.22	31.01	54.25	45.43	64.66

Panel B: Tercile 2

	Mean	St. dev.	Median	25th	75th
Stock characteristics:					
Market capitalization, $ millions	6,054	3,211	5,472	4,137	7,246
Price, $	50.85	42.46	42.41	27.35	61.04
Volume, thousands of shares	2,553	4,312	1,165	641	2,390
Volatility	8.97	4.92	8.22	6.30	10.19
High-frequency liquidity metrics, bps:					
Quoted spread	8.80	67.72	5.05	4.01	6.88
Effective spread	7.52	67.35	4.12	3.26	5.58
Price impact, 60s	4.13	1.91	3.73	3.00	4.72
Price impact, 300s	4.84	12.21	3.92	3.04	5.06
Realized spread, 60s	2.65	47.91	0.36	0.11	0.87
Realized spread, 300s	2.35	46.63	0.25	0.04	0.61
Low-frequency liquidity metrics, bps:					
EOD quoted spread	5.99	33.34	3.26	2.56	4.99
CS effective spread	99.42	40.68	90.38	72.63	116.02
AR effective spread	68.09	36.29	62.66	49.38	78.18

(Table 7 continues on the next page)

(Table 7 continued)

Panel C: Tercile 3					
	Mean	St. dev.	Median	25th	75th
Stock characteristics:					
Market capitalization, $ millions	2,449	1,541	2,256	1,736	2,825
Price, $	33.29	35.69	27.01	15.84	41.02
Volume, thousands of shares	2,026	4,271	778	388	1,734
Volatility	16.85	48.93	12.01	9.10	15.35
High-frequency liquidity metrics, bps:					
Quoted spread	10.80	21.17	7.78	5.80	11.38
Effective spread	8.35	14.49	6.03	4.72	9.04
Price impact, 60s	5.85	2.95	5.02	4.03	6.96
Price impact, 300s	6.48	3.87	5.52	4.26	7.59
Realized spread, 60s	2.49	12.68	0.87	0.36	1.99
Realized spread, 300s	1.86	11.85	0.55	0.16	1.36
Low-frequency liquidity metrics, bps:					
EOD quoted spread	8.56	20.88	5.28	3.83	9.14
CS effective spread	115.88	46.88	104.69	83.28	140.13
AR effective spread	77.85	32.92	71.29	57.58	94.60

Table 8
Cross-Sectional Results: Trading Cost Components

The table examines the relationship between SAD Incidence and various spreads and trading cost metrics for each of three size terciles over the sample period January 2010 through December 2019. Panel A corresponds to the largest firms (tercile 1), Panel B corresponds to smaller firms (tercile 2), and Panel C corresponds to the smallest firms (tercile 3). The reported coefficients are obtained from the regression of the following form in the case of price impacts:

$$DepVar_{i,t} = \alpha_i + \gamma_{year} + \beta_1 Onset/Recovery_t + \beta_3 Volume_{i,t-1} + \beta_4 Volatility_{i,t-1} + \varepsilon_{i,t}.$$

and the following form in the case of realized spreads:

$$DepVar_{i,t} = \alpha_i + \gamma_{year} + \beta_2 Incidence_t + \beta_3 Volume_{i,t-1} + \beta_4 Volatility_{i,t-1} + \varepsilon_{i,t}.$$

$DepVar$ is the effective or quoted spread in stock i on day t; $Onset/Recovery$ is the SAD onset/recovery variable, $Incidence$ is the SAD incidence variable; $Volume$ is the lagged natural logarithm of daily number of shares traded; and $Volatility$ is the lagged quote-based intraday volatility (expressed as a standard deviation). The models are estimated using ordinary least squares, controlling for stock and year fixed effects, and the standard errors are clustered by firm and date. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively. Each tercile contains over 730,000 stock-day observations. For some variables, fewer than the full sample number of observations are available.

Panel A: Tercile 1				
	Price impact, 60s	Price impact, 300s	Realized spread, 60s	Realized spread, 300s
	[1]	[2]	[3]	[4]
Onset/Recovery	0.180***	0.195***		
	(0.033)	(0.033)		
Incidence			0.022	0.031
			(0.021)	(0.020)
Volatility	0.030**	0.037**	0.046	0.038
	(0.012)	(0.017)	(0.035)	(0.028)
Volume	0.308***	0.238***	−0.284*	−0.216
	(0.036)	(0.047)	(0.165)	(0.151)
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Adj. R^2	0.556	0.453	0.417	0.274
Magnitude as a % of Mean	4.8%	5.2%	5.6%	8.4%

(Table 8 continues on the next page)

(Table 8 continued)

Panel B: Tercile 2				
	Price impact, 60s	Price impact, 300s	Realized spread, 60s	Realized spread, 300s
	[1]	[2]	[3]	[4]
Onset/Recovery	0.326***	0.341***		
	(0.055)	(0.057)		
Incidence			0.117***	0.113***
			(0.020)	(0.020)
Volatility	0.029***	0.031***	0.013***	0.012***
	(0.006)	(0.006)	(0.004)	(0.004)
Volume	0.377***	0.238***	−0.382***	−0.244***
	(0.030)	(0.036)	(0.031)	(0.027)
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Adj. R^2	0.526	0.486	0.614	0.317
Magnitude as a % of Mean	6.2%	5.5%	4.2%	4.5%

Panel C: Tercile 3				
	Price impact, 60s	Price impact, 300s	Realized spread, 60s	Realized spread, 300s
	[1]	[2]	[3]	[4]
Onset/Recovery	0.484***	0.527***		
	(0.080)	(0.088)		
Incidence			0.195**	0.189**
			(0.080)	(0.076)
Volatility	0.025***	0.032***	0.055***	0.048***
	(0.009)	(0.011)	(0.009)	(0.008)
Volume	0.388***	0.067	−1.430***	−1.112***
	(0.111)	(0.144)	(0.295)	(0.278)
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Adj. R^2	0.381	0.344	0.620	0.562
Magnitude as a % of Mean	6.5%	6.5%	7.4%	9.6%

Table 9
Cross-Sectional Results: Displayed Liquidity and Trading Costs

The table examines the relationship between SAD Incidence and various quoted and effective spreads for each of three size terciles over the sample period January 2010 through December 2019. Panel A corresponds to the largest firms (tercile 1), Panel B corresponds to smaller firms (tercile 2), and Panel C corresponds to the smallest firms (tercile 3). The reported coefficients are obtained from the regression of the following form:

$$DepVar_{i,t} = \alpha_i + \gamma_{year} + \beta_1 Onset/Recovery_t + \beta_2 Incidence_t + \beta_3 Volume_{i,t-1} + \beta_4 Volatility_{i,t-1} + \varepsilon_{i,t}.$$

DepVar is the effective or quoted spread in stock i on day t, $Onset/Recovery$ is the SAD onset/recovery variable, *Incidence* is the SAD incidence variable, *Volume* is the lagged natural logarithm of daily number of shares traded, and *Volatility* is the lagged quote-based intraday volatility (expressed as a standard deviation). The models are estimated using ordinary least squares, controlling for stock and year fixed effects, and the standard errors are clustered by firm and date. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively. Each tercile contains over 700,000 stock-day observations. For some variables, fewer than the full sample number of observations are available, most notably CS and AR which discard negative estimates.

Panel A: Tercile 1						
	Quoted [1]	Effective [2]	Depth [3]	EOD [4]	CS [5]	AR [6]
Onset/Recovery	0.102**	0.086**	−3.550***	0.057	8.222***	3.209
	(0.046)	(0.039)	(0.902)	(0.053)	(1.736)	(2.926)
Incidence	0.152***	0.062**	−1.836***	0.041	2.481**	1.677
	(0.037)	(0.028)	(0.691)	(0.031)	(1.144)	(1.874)
Volatility	0.100*	0.077*	0.151	0.061	0.712**	0.459
	(0.058)	(0.046)	(0.136)	(0.042)	(0.306)	(0.314)
Volume	−0.067	0.020	14.366***	0.100	28.484***	15.297***
	(0.220)	(0.188)	(3.933)	(0.183)	(1.166)	(1.352)
Firm FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Adj. R^2	0.58	0.59	0.71	0.54	0.20	0.05

(Table 9 continues on the next page)

(Table 9 continued)

Panel B: Tercile 2

	Quoted [1]	Effective [2]	Depth [3]	EOD [4]	CS [5]	AR [6]
Onset/Recovery	0.356***	0.252***	−2.380	−0.012	9.040***	2.158
	(0.070)	(0.055)	(2.019)	(0.052)	(1.892)	(3.068)
Incidence	0.364***	0.196***	−0.817	0.120***	2.800**	3.096
	(0.044)	(0.032)	(0.598)	(0.024)	(1.225)	(1.991)
Volatility	0.061***	0.042***	−0.130*	0.028***	0.806***	0.433***
	(0.014)	(0.009)	(0.071)	(0.007)	(0.137)	(0.083)
Volume	−0.126**	−0.010	7.690***	−0.011	27.599***	14.147***
	(0.062)	(0.047)	(2.220)	(0.096)	(1.036)	(1.139)
Firm FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Adj. R^2	0.73	0.74	0.66	0.64	0.22	0.05

Panel C: Tercile 3

	Quoted [1]	Effective [2]	Depth [3]	EOD [4]	CS [5]	AR [6]
Onset/Recovery	0.786***	0.522***	0.901	0.230*	11.905***	5.114
	(0.152)	(0.111)	(0.880)	(0.129)	(2.271)	(3.746)
Incidence	0.528***	0.281***	−2.931***	0.240*	3.968***	3.447
	(0.153)	(0.101)	(0.899)	(0.144)	(1.423)	(2.391)
Volatility	0.116***	0.080***	−0.018	0.091***	0.282*	0.550***
	(0.026)	(0.018)	(0.014)	(0.012)	(0.150)	(0.064)
Volume	−1.624***	−1.063***	6.954***	−0.831**	28.412***	15.964***
	(0.524)	(0.367)	(1.135)	(0.391)	(1.411)	(1.272)
Firm FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Adj. R^2	0.66	0.64	0.81	0.54	0.26	0.06

Table 10
International Summary Statistics

The table reports summary statistics for each of the latitude groupings over the sample period January 2010 through December 2019. The data are from Datastream. Summary statistics for each latitude grouping appear for the following stock characteristics: market capitalization, share price, daily trading volume, and volatility is computed as the the difference between the high and low prices of the day scaled by their average and multiplied by 100. Summary statistics for each latitude grouping also appear for the following low-frequency liquidity metrics: the end-of-day (EOD) quoted spread, the Corwin and Schultz (2012) (CS) effective spread, and the Abdi and Ranaldo (2017) (AR) effective spread. When aggregating, we first compute the averages of all variables for each stock and then compute sample characteristics across stocks. Panel A corresponds to the Northern Temperate Zone countries: Norway, Germany, the United Kingdom, France, Canada, and Italy. Panel B corresponds to the Northern Sub-Tropics countries: China, Japan, and Hong Kong. Panel C corresponds to the Tropics countries: Brazil, Thailand, the Philippines, and Indonesia. Panel D corresponds to the Southern Sub-Tropics and Temperate Zone countries: New Zealand, Argentina, Australia, Chile, and South Africa. The number of stock-day observations in each full sample is as follows: 9,059,434 for Panel A, 15,699,170 for Panel B, 3,066,550 for Panel C, and 3,148,766 for Panel D. For some variables, fewer than the full sample number of observations are available, most notably CS and AR which discard negative estimates.

Panel A: Northern Temperate Zone (Above 40° N)					
	Mean	St. dev.	Median	25th	75th
Stock characteristics					
Market capitalization, \$ millions	1,768	8,107	107	28	493
Price, \$	13.90	31.05	4.06	0.99	13.10
Volume, thousands of shares	683.00	4734.10	46.83	6.83	258.61
Volatility	0.052	0.061	0.035	0.024	0.057
Low-frequency liquidity metrics, bps					
EOD quoted spread	532.66	948.71	259.12	100.83	560.92
CS effective spread	282.84	504.19	150.01	101.26	262.33
AR effective spread	394.36	596.38	227.88	150.28	389.48

Panel B: Northern Subtropics (23.5° N to 40° N)					
Stock characteristics					
Market capitalization, \$ millions	1,631	6,887	0.495	0.140	1.099
Price, \$	8.33	22.78	2.76	0.96	7.76
Volume, thousands of shares	7,028.53	15,371.57	2,648.88	208.287	8,068.13
Volatility	0.039	0.019	0.038	0.027	0.045
Low-frequency liquidity metrics, bps					
EOD quoted spread	88.74	137.80	33.44	12.88	107.81
CS effective spread	153.87	91.24	142.15	109.43	173.41
AR effective spread	210.33	117.46	185.35	153.34	231.39

(Table 10 continues on the next page)

(Table 10 continued)

Panel C: Tropics (Between 23.5° N and 23.5° S)					
	Mean	St. dev.	Median	25th	75th
Stock characteristics					
Market capitalization, $ millions	961	3,640	120	36	538
Price, $	4.61	23.24	0.20	0.05	1.40
Volume, thousands of shares	9,704.36	40,369.49	1,500.58	133.82	6,110.67
Volatility	0.041	0.028	0.034	0.025	0.048
Low-frequency liquidity metrics, bps					
EOD quoted spread	247.21	365.60	118.82	78.21	258.40
CS effective spread	214.34	184.60	171.65	122.35	257.59
AR effective spread	303.65	227.73	236.96	171.20	367.81

Panel D: Southern Sub-Tropics and Temperate Zone (23.5° S and Higher)					
Stock characteristics					
Market capitalization, $ millions	752	3,929	57	17	277
Price, $	2.47	9.27	0.32	0.10	1.39
Volume, thousands of shares	1,367.54	7,511.59	286.30	97.03	884.49
Volatility	0.068	0.071	0.053	0.293	0.079
Low-frequency liquidity metrics, bps					
EOD quoted spread	807.96	1,297.42	498.61	219.90	951.29
CS effective spread	546.54	826.21	300.83	142.48	575.38
AR effective spread	645.12	838.17	431.16	199.00	736.36

Table 11
International Low-Frequency Liquidity Metrics

The table examines, from an international perspective, the relation between the SAD metrics and each of three low-frequency liquidity proxies: the end of day spread (EOD), which proxies for displayed liquidity, the Corwin-Schultz (CS) metric – a proxy for trading costs, and the Abdi-Ranaldo (AR) metric – also a proxy for trading costs. Results appear for each latitude grouping: the northern temperate zone in Panel A (Norway, Germany, the United Kingdom, France, Canada, and Italy), the northern sub-tropics in Panel B (China, Japan, Israel, and Hong Kong), the tropics in Panel C (Brazil, Thailand, the Philippines, and Indonesia), and the southern sub-tropics and temperate zone in Panel D (New Zealand, Argentina, Australia, Chile, and South Africa). The sample period spans January 2010 through December 2019. The reported coefficients are obtained from the regression of the following form:

$$DepVar_{i,t} = \alpha_i + \gamma_{year} + \beta_1 Onset/Recovery_t + \beta_2 Incidence_t + \beta_3 Volume_{i,t-1} + \beta_4 Volatility_{i,t-1} + \varepsilon_{i,t}.$$

where *DepVar* is one of the three above-mentioned low-frequency metrics in stock i on day t, *Onset/Recovery* is the SAD onset/recovery variable, *Incidence* is the SAD incidence variable, *Volume* is the lagged natural logarithm of daily number of shares traded, and *Volatility* is computed as the the difference between the high and low prices of the day scaled by their average and multiplied by 100. The models are estimated using ordinary least squares, controlling for stock and year fixed effects, and the standard errors are clustered by firm and date. *** and ** indicate statistical significance at the 1% and 5% levels.

Panel A: Northern Temperate Zone (Above 40° N)			
	EOD [1]	CS [2]	AR [3]
Onset/Recovery	6.137***	0.891	5.712***
	(1.564)	(0.910)	(1.750)
Incidence	0.156	2.160***	2.400**
	(0.904)	(0.493)	(1.014)
Volume	−0.575***	−0.181***	−0.292***
	(0.014)	(0.010)	(0.011)
Volatility	0.356***	0.350***	0.361***
	(0.010)	(0.009)	(0.010)
Firm FE	Y	Y	Y
Year FE	Y	Y	Y
N (millions)	8.8	6.2	5.4
Adj. R^2	0.64	0.60	0.57

Panel B: North Sub-Tropics (23.5° N to 40° N)			
Onset/recovery	2.747***	−0.943	−1.308
	(0.749)	(1.117)	(2.917)
Incidence	0.591	−0.444	−0.866
	(0.410)	(0.631)	(1.591)
Volume	−0.276***	−0.005	−0.067***
	(0.017)	(0.020)	(0.022)
Volatility	0.090***	0.226***	0.221***
	(0.012)	(0.016)	(0.018)
Firm FE	Y	Y	Y
Year FE	Y	Y	Y
N (millions)	15.9	10.2	8.8
Adj. R^2	0.49	0.43	0.33

(Table 11 continues on the next page)

(Table 11 continued)

Panel C: Tropics (Between 23.5° N and 23.5° S)			
	EOD [1]	CS [2]	AR [3]
Onset/Recovery	2.397	1.474	−1.903
	(1.688)	(1.236)	(2.024)
Incidence	−1.187	−1.792**	1.520
	(1.065)	(0.695)	(1.126)
Volume	−0.402***	−0.129***	−0.232***
	(0.030)	(0.019)	(0.020)
Volatility	0.254***	0.340***	0.330***
	(0.041)	(0.028)	(0.031)
Firm FE	Y	Y	Y
Year FE	Y	Y	Y
N (millions)	3.0	1.2	1.1
Adj. R^2	0.51	0.54	0.45

Panel D: Southern Sub-Tropics & Temperate Zone (23.5° S and Higher)			
Onset/Recovery	−26.069***	−9.955***	−11.402***
	(2.789)	(1.758)	(2.400)
Incidence	15.277***	5.231***	6.616***
	(1.759)	(0.954)	(1.305)
Volume	−0.705***	−0.373***	−0.377***
	(0.019)	(0.015)	(0.015)
Volatility	0.494***	0.587***	0.518***
	(0.014)	(0.013)	(0.012)
Firm FE	Y	Y	Y
Year FE	Y	Y	Y
N (millions)	3.0	2.2	2.0
Adj. R^2	0.58	0.74	0.67

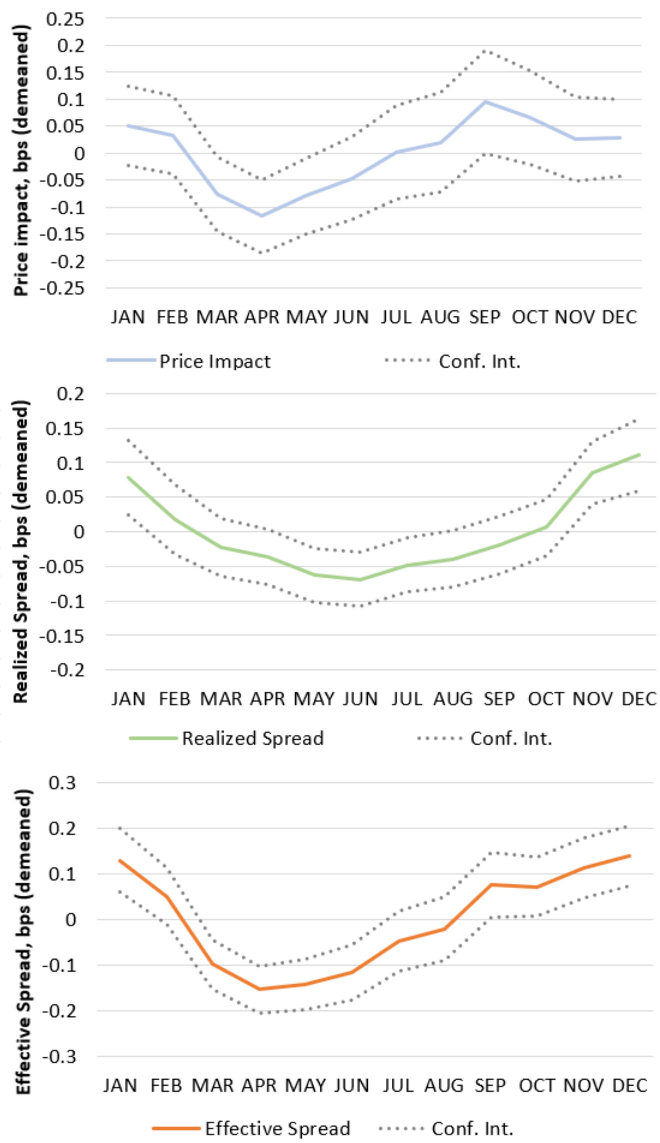

Figure 1

Seasonal patterns in price impacts, realized and effective spreads

The figure plots monthly estimates of price impacts (blue line), realized spreads (green line), and effective spreads (orange line) for the sample period January 2010 through December 2019. The effective spread is computed from intraday TAQ data as twice the signed difference between the trade price and the quote midpoint. The price impact is the signed difference between the midpoint at the time of the trade and the midpoint 60 seconds later. The realized spread is the difference between the effective spread and the price impact. All three metrics are three-month centered moving averages. All series have been demeaned for ease of comparison across plots. Dotted lines represent a plus-or-minus 10% confidence interval around the spread. The confidence intervals are based on clustered standard errors, clustered by date and firm.

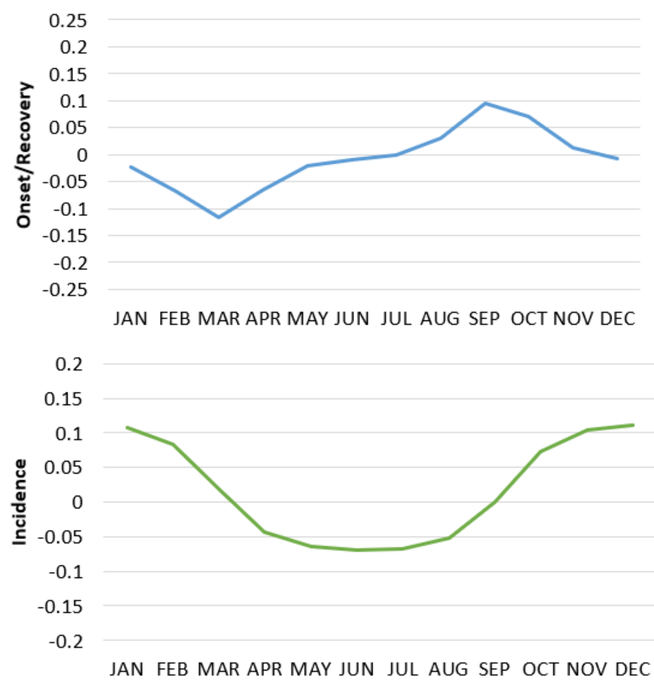

Figure 2
Seasonal depression

The figure plots monthly estimates of the proportion of the population newly affected or newly recovering from depression (onset/recovery, blue line) and the currently affected share of the population (incidence, green line) for the sample period January 2010 through December 2019.

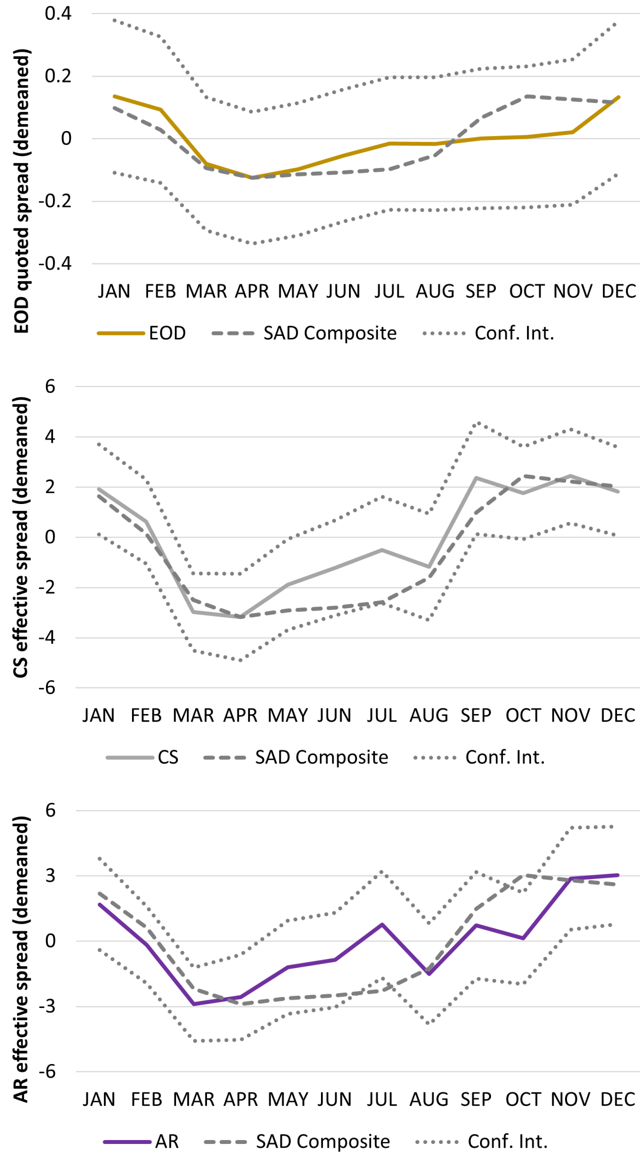

Figure 3
SAD and low-frequency liquidity metrics

The figure plots monthly estimates of end-of-day (EOD; solid yellow line in the top chart) quoted spreads, Corwin-Schultz (CS; solid grey line in the middle chart) effective spread estimate, Abdi-Ranaldo (AR; solid purple line in the bottom chart) effective spread estimate, and the SAD Composite measure (long-dashed line) for the sample period staring in January 2010 through December 2019. The spread measures are three-month centered moving averages. All series have been demeaned for ease of comparison across plots. Dotted lines represent a plus-or-minus 10% confidence interval around the spread. The confidence intervals are based on clustered standard errors, clustered by date and firm.

Table A1
Country-by Country Summary Statistics
(Calculated on Means of Variables Firm-by-Firm)

	Mean	Std	Min	Max	Skew	Kurt
Argentina:						
Market capitalization, $ millions	675.62	1419.3	3.28	11023	5.216	34.47
Local Currency Price	74.655	317.93	0.95	2904.6	8.613	77.18
Price, $	3.120	6.60	0.17	48.51	5.096	29.77
Return %	0.289	0.30	-0.25	1.94	3.311	14.46
Volume (millions of shares)	0.231	0.66	0.00	5.00	5.588	35.97
Natural log volume	3.264	1.77	-0.72	8.10	0.389	0.27
Volatility	0.039	0.01	0.02	0.08	1.160	3.95
EOD quoted spread	227.62	183.43	40.12	1347.0	3.266	16.46
CS effective spread	172.18	42.61	88.03	338.24	1.544	3.34
AR effective spread	240.42	48.05	136.76	447.33	1.501	4.18
Australia:						
Market capitalization, $ millions	626.02	4139.8	0.08	96829	16.003	306.17
Local Currency Price	2.784	12.35	0.00	187.40	7.807	67.31
Price, $	2.241	9.75	0.00	173.65	8.220	81.79
Return %	0.331	2.12	-20.05	66.19	16.271	438.18
Volume (millions of shares)	1.131	3.17	0.00	84.90	12.820	264.34
Natural log volume	5.082	1.66	-1.28	10.51	-0.191	0.65
Volatility	0.073	0.07	0.00	0.67	4.589	28.95
EOD quoted spread	880.94	1343.0	10.56	20000	7.920	92.39
CS effective spread	610.93	862.47	8.27	6666.7	4.167	21.56
AR effective spread	719.29	871.71	11.97	6931.5	4.139	21.88
Brazil:						
Market capitalization, $ millions	2051.5	6546.7	0.00	74516	7.097	60.27
Local Currency Price	62.691	166.93	0.11	1509.3	5.535	34.79
Price, $	19.380	47.09	0.03	408.72	5.139	29.34
Return %	0.305	1.18	-5.42	17.63	7.624	102.42
Volume (millions of shares)	1.940	19.46	0.00	424.56	20.646	446.14
Natural log volume	3.147	2.96	-2.10	11.43	0.384	-0.92
Volatility	0.040	0.04	0.00	0.32	4.135	24.54
EOD quoted spread	384.97	632.66	14.63	4996.9	3.060	11.38
CS effective spread	184.56	241.21	2.28	3155.5	7.418	79.24
AR effective spread	270.79	275.60	2.50	3228.6	5.466	47.17
Canada:						
Market capitalization, $ millions	1341.5	5367.5	0.00	92961	9.372	116.06
Local Currency Price	12.329	26.20	0.02	482.57	10.172	151.58
Price, $	10.887	23.56	0.02	431.06	11.113	176.75
Return %	0.108	0.87	-22.50	14.25	-6.989	338.16
Volume (millions of shares)	0.278	0.57	0.00	7.14	4.377	27.13
Natural log volume	3.563	1.92	-2.07	8.62	0.095	-0.72
Volatility	0.042	0.04	0.00	0.66	5.352	59.50
EOD quoted spread	346.79	497.29	10.69	7014.5	5.235	41.60
CS effective spread	236.15	347.92	10.34	6534.4	7.314	90.35
AR effective spread	323.24	389.70	16.42	6791.1	6.176	67.67

(Table A1 continues on the next page)

(Table A1 continued)

	Mean	Std	Min	Max	Skew	Kurt
Chile:						
Market capitalization, $ millions	1673.5	3041.7	0.00	20219	3.519	14.49
Local Currency Price	2237.8	5093.7	0.87	38084	4.650	25.72
Price, $	3.991	9.10	0.00	66.94	4.640	25.40
Return %	0.410	1.46	-4.52	7.41	2.564	11.88
Volume (millions of shares)	6.039	25.82	0.00	296.26	9.234	99.69
Natural log volume	5.151	2.44	0.36	12.13	0.269	-0.30
Volatility	0.026	0.02	0.00	0.16	3.834	17.25
EOD quoted spread	385.27	357.35	46.46	1934.2	1.949	3.99
CS effective spread	123.23	74.14	25.74	426.57	1.808	3.54
AR effective spread	163.94	97.32	31.50	833.82	3.138	15.51

China:						
Market capitalization, $ millions	1888.7	6691.2	0.00	221832	20.024	533.67
Local Currency Price	19.816	18.12	0.38	353.32	5.612	67.55
Price, $	2.997	2.64	0.13	53.35	5.598	68.78
Return %	0.022	0.35	-5.43	10.03	18.861	566.54
Volume (millions of shares)	13.520	19.14	0.01	382.77	6.727	76.70
Natural log volume	8.663	0.98	3.14	12.28	-0.370	1.98
Volatility	0.040	0.01	0.01	0.14	1.653	17.43
EOD quoted spread	14.799	19.69	1.25	309.56	7.672	71.63
CS effective spread	149.23	30.57	45.23	665.22	3.153	35.08
AR effective spread	193.19	65.27	13.65	1825.3	13.735	314.27

France:						
Market capitalization, $ millions	2487.5	10200	1.34	131380	7.776	73.80
Local Currency Price	27.962	41.20	0.01	298.39	3.022	11.63
Price, $	34.341	50.89	0.02	395.26	3.111	12.56
Return %	0.126	1.38	-27.69	18.32	-5.586	234.03
Volume (millions of shares)	0.218	1.19	0.00	27.14	15.308	312.40
Natural log volume	1.684	2.39	-2.30	9.78	0.789	0.10
Volatility	0.033	0.02	0.01	0.28	4.534	36.47
EOD quoted spread	265.57	507.66	3.83	9018.1	8.401	113.84
CS effective spread	168.19	232.95	1.21	4569.7	10.492	166.48
AR effective spread	231.87	236.27	10.01	4147.6	8.164	107.53

Germany:						
Market capitalization, $ millions	2197.7	9195.9	0.00	100000	7.059	56.82
Local Currency Price	18.388	32.78	0.01	418.27	5.228	41.70
Price, $	22.431	39.64	0.01	470.57	5.103	39.16
Return %	0.416	7.81	-50.00	300.00	31.718	1211.7
Volume (millions of shares)	0.103	0.66	0.00	13.19	13.746	224.11
Natural log volume	1.271	1.92	-2.23	9.39	1.121	1.47
Volatility	0.074	0.10	0.01	0.73	3.185	11.33
EOD quoted spread	798.25	1560.3	4.06	20000	4.735	32.48
CS effective spread	450.14	883.35	30.04	7564.3	3.985	18.27
AR effective spread	639.29	1024.8	53.98	9559.1	3.562	15.25

(Table A1 continues on the next page)

(Table A1 continued)

	Mean	Std	Min	Max	Skew	Kurt
Hong Kong:						
Market capitalization, $ millions	1369.2	8242.4	2.60	213362	18.964	431.62
Local Currency Price	4.201	12.28	0.02	251.38	9.233	124.36
Price, $	0.539	1.58	0.00	32.28	9.236	124.40
Return %	0.032	0.37	-5.74	5.41	-1.293	99.88
Volume (millions of shares)	5.326	14.12	0.00	474.26	19.415	571.10
Natural log volume	6.755	1.46	1.19	11.90	-0.052	0.27
Volatility	0.052	0.02	0.00	0.20	1.086	3.39
EOD quoted spread	234.52	164.18	8.30	1909.1	1.654	7.78
CS effective spread	212.58	96.37	31.24	1278.3	2.591	17.13
AR effective spread	293.74	133.16	50.46	1413.8	1.637	7.05

Indonesia:						
Market capitalization, $ millions	681.85	2496.4	0.83	27856	8.110	72.37
Local Currency Price	1947.5	5435.0	54.75	82217	8.368	93.60
Price, $	0.168	0.49	0.00	7.53	8.474	95.05
Return %	0.119	0.65	-5.89	7.65	-0.300	54.79
Volume (millions of shares)	15.337	41.24	0.00	464.65	5.960	43.57
Natural log volume	6.234	2.62	-0.19	12.48	-0.015	-0.79
Volatility	0.053	0.03	0.01	0.22	1.825	5.28
EOD quoted spread	260.12	260.51	25.90	2112.5	2.863	11.24
CS effective spread	230.44	119.48	52.20	909.19	1.754	4.54
AR effective spread	321.35	181.91	69.18	1481.9	1.974	5.99

Italy:						
Market capitalization, $ millions	1377.6	5240.7	1.40	73040	8.505	91.03
Local Currency Price	6.464	11.36	0.01	153.35	6.890	68.63
Price, $	7.732	13.50	0.02	176.73	6.680	63.71
Return %	0.023	0.20	-1.09	2.17	2.402	33.25
Volume (millions of shares)	1.888	12.00	0.00	189.05	11.815	157.66
Natural log volume	3.759	2.35	-1.33	11.75	0.714	0.25
Volatility	0.036	0.02	0.00	0.16	2.432	12.15
EOD quoted spread	223.24	192.91	8.26	1348.1	1.963	5.14
CS effective spread	171.88	139.94	20.10	1443.3	4.752	28.98
AR effective spread	225.01	143.76	26.98	1369.0	3.827	20.17

(Table A1 continues on the next page)

(Table A1 continued)

	Mean	Std	Min	Max	Skew	Kurt
Japan:						
Market capitalization, $ millions	1505.0	6067.9	2.44	181230	13.153	279.09
Local Currency Price	1969.7	3227.4	2.27	49995	6.864	64.59
Price, $	19.710	35.42	0.02	469.60	7.198	67.74
Return %	0.069	0.37	-8.44	7.53	-5.837	301.99
Volume (millions of shares)	0.603	4.17	0.00	213.22	41.078	2026.0
Natural log volume	3.959	2.02	-2.30	12.06	0.203	-0.13
Volatility	0.029	0.02	0.00	0.50	11.021	212.29
EOD quoted spread	77.962	121.48	9.96	3696.9	15.648	353.83
CS effective spread	120.30	113.33	14.87	3651.1	16.259	380.83
AR effective spread	174.92	126.80	18.11	3764.8	13.770	291.15
New Zealand:						
Market capitalization, $ millions	538.53	988.96	0.83	6647.7	3.287	12.80
Local Currency Price	2.474	2.70	0.02	17.29	2.509	9.20
Price, $	1.805	1.98	0.01	12.47	2.555	9.45
Return %	0.254	0.97	-1.61	10.13	7.429	69.94
Volume (millions of shares)	0.365	0.74	0.00	7.56	6.329	56.22
Natural log volume	4.198	1.57	0.30	8.63	0.162	-0.43
Volatility	0.040	0.07	0.01	0.55	5.193	32.17
EOD quoted spread	611.01	1757.6	53.17	18095	7.730	68.27
CS effective spread	399.54	890.88	29.02	6666.7	5.060	28.69
AR effective spread	418.84	780.49	56.55	5822.0	4.949	29.10
Norway:						
Market capitalization, $ millions	938.81	4443.4	0.98	70556	12.373	179.88
Local Currency Price	53.331	150.24	0.30	2615.8	14.474	244.05
Price, $	7.285	19.45	0.05	330.92	13.567	221.11
Return %	0.153	0.72	-6.43	8.02	2.406	64.61
Volume (millions of shares)	0.332	0.92	0.00	8.85	5.469	36.31
Natural log volume	3.270	1.96	-1.30	8.73	0.272	-0.20
Volatility	0.047	0.03	0.01	0.21	2.220	6.72
EOD quoted spread	299.83	359.49	8.99	2902.8	3.301	14.80
CS effective spread	226.78	176.03	33.63	1290.6	2.865	10.58
AR effective spread	325.44	246.70	39.70	1922.2	3.063	12.93
Philippines:						
Market capitalization, $ millions	871.62	1944.3	0.00	14821	3.858	17.39
Local Currency Price	47.260	191.33	0.00	2197.7	7.982	73.75
Price, $	1.017	4.15	0.00	48.50	8.125	76.58
Return %	0.094	0.61	-5.36	5.93	0.111	56.07
Volume (millions of shares)	12.683	83.08	0.00	1311.5	13.667	207.71
Natural log volume	5.531	2.45	-0.85	12.58	-0.168	0.00
Volatility	0.046	0.03	0.01	0.26	2.605	12.95
EOD quoted spread	313.78	339.30	28.18	2887.5	2.713	12.20
CS effective spread	227.08	194.32	43.73	2488.9	6.030	62.85
AR effective spread	318.21	229.03	46.34	2515.8	3.927	29.66

(Table A1 continues on the next page)

(Table A1 continued)

	Mean	Std	Min	Max	Skew	Kurt
South Africa:						
Market capitalization, $ millions	1221.4	3912.0	0.00	55447	8.230	93.14
Local Currency Price	35.775	98.49	0.01	1604.2	10.203	148.75
Price, $	3.302	8.47	0.00	129.09	8.729	114.06
Return %	0.377	1.42	-4.63	14.38	5.376	39.79
Volume (millions of shares)	1.563	10.84	0.00	209.10	16.680	309.84
Natural log volume	4.616	1.87	-0.46	12.06	0.325	0.42
Volatility	0.067	0.08	0.01	0.69	4.027	20.46
EOD quoted spread	728.06	1094.1	16.70	8242.0	2.907	10.64
CS effective spread	460.54	737.94	33.99	6074.1	4.162	21.10
AR effective spread	562.54	794.83	39.78	6292.2	3.776	17.52

Thailand:						
Market capitalization, $ millions	575.26	1968.8	2.64	31700	8.398	95.87
Local Currency Price	24.132	78.34	0.20	1626.4	12.627	223.14
Price, $	0.750	2.44	0.01	50.52	12.605	221.83
Return %	0.037	0.25	-2.56	2.72	-0.214	45.96
Volume (millions of shares)	8.682	21.00	0.00	228.53	6.134	48.01
Natural log volume	6.178	2.46	-0.73	11.95	-0.517	-0.19
Volatility	0.031	0.02	0.01	0.30	6.321	70.32
EOD quoted spread	141.22	175.07	35.21	2750.0	7.259	78.03

United Kingdom:						
Market capitalization, $ millions	1674.7	8854.4	0.23	175070	10.885	146.20
Local Currency Price	250.13	565.59	0.03	9676.7	6.708	72.05
Price, $	3.683	8.22	0.00	141.92	6.719	73.02
Return %	0.066	0.86	-21.11	13.98	-2.995	228.81
Volume (millions of shares)	1.409	6.20	0.00	165.90	14.572	298.12
Natural log volume	4.213	2.11	-1.56	11.92	0.250	-0.25
Volatility	0.055	0.04	0.00	0.37	2.472	9.36
EOD quoted spread	666.05	737.14	2.59	5833.3	2.302	7.85
CS effective spread	259.94	239.59	13.04	3491.7	3.935	28.97
AR effective spread	359.87	327.84	21.76	3566.7	3.184	16.45