The Best Time of Day to Trade – An Analysis of Intraday Spread

In this paper, I conduct an analysis of intraday spreads and prove the conventional wisdom that the best time of day to trade is near market open and close, is only partially true: spread decreases throughout the day, and the best time to trade is right before market close. In doing so, traders can save around $1 for every $10,000 traded if transacting near market close rather than market open. I also show the widely held belief that spreads are large during times of high volatility to be true.

The premise holds that market open and close are times of high liquidity, and would therefore have smaller spreads, implying they are the ideal times to trade if the goal is to lose as little as possible to the spread 'premium'.

This study was performed on second-timestamped quotes for a popular ETF, IVE (the iShares S&P 500 Value ETF), from September 2009 to November 2022. I performed grouping on the data at various intervals and calculated correlations, significance, and created a rudimentary predictive model for spread based on trading volumes.

Jake Li
lijake@umich.edu
2484803953
11/30/2022

Contents

INTRODUCTION I
Conventional Wisdom 1
Relevance 1
Conclusion 1

METHOD AND DATA I
Data Procurement 1
Methodology 1
Cleaning Data 1
Calculations 2

RESULTS 2
Objective and Hypothesis 2
Insights 3
Spread and Volume 3
Spread and Volatility 4
Statistical Significance 5
Application 5

CONCLUSION 5

Tables

Table 1. Pearson Correlation of Spread and Volume 5

Figures

Figure 1. Spread and Volume Throughout Trading Hours 3
Figure 2. Spread and Volume Correspondence 4
Figures 3 and 4. Spread and Volatility Across Years and Months 4

Introduction

Conventional Wisdom

It's generally believed that the best time to buy an ETF is when it is most liquid, i.e. when trading volume is highest, or at the beginning and end of the trading day. The premise is that with high volume comes a narrowing of the ETF's spread: the difference in its bid and ask prices (what buyers are willing to pay and what sellers are willing to sell for, respectively). Ideally, one buys when the spread is small and sells when the spread is large – this way, one minimizes transaction losses to the spread.

From basic economic theory, higher quantities of a good mean there are more 'substitutes' for the good and therefore price elasticity of the good increases; the good becomes more 'commoditized' and surplus decreases. In trading, this suggests that, when transacting in the market during times of higher liquidity, buyers *should* pay less of a spread 'premium' and sellers *should* receive less of that 'premium' as well.

Empirically, this seems to be the case – placing a limit order (an order to buy or sell a share for a specific price or better) at market open when liquidity is high, for example, seems to execute much more consistently than in the middle of the day when liquidity is lower.

Relevance

The importance of this study is reflected in the implications of being on the 'wrong side of the spread' for large-volume institutional investors like market makers and high frequency traders (HFTs). These firms often employ delta-neutral strategies, and under those circumstances, must carefully analyze spread. Though spread is tiny compared to the underlying asset's price, firms who trade frequently, even with market-neutral positions, can expose themselves to unnecessary losses by trading at the wrong times and overpaying for an asset.

Conclusion

Surprisingly, the conventional wisdom is only half right. In fact, reality is almost the exact opposite – spread and volume correlate positively, meaning that, barring a few outliers, the smallest spreads are in the middle and end of the day when volume is lower.

Method and data

Data Procurement

Due to the high cost of granular historical intraday ticker data containing bid/ask prices, I am using a popular ETF (ticker IVE, the iShares S&P 500 Value ETF) as a case study. The results can loosely be generalized to other similar ETFs. I downloaded a dataset from Kibot for the IVE ETF with second timestamps, dates, close prices, bid and ask prices, and volumes. Kibot's data is NBBO (National Best Bid and Offer) and is taken from multiple exchanges and ECNs, making it relatively trustworthy. The data spans 4,810 days (of which 3,316 had transactions) from 09/28/2009 to 11/28/2022 and contains 10.4 million entries.

Methodology

My analysis involved the following steps: loading data, cleaning data, performing calculations, then performing grouping calculations and visualizations on several interval lengths.

Cleaning Data

To clean the data, I removed all entries with NA values or 0's for the prices. These would have skewed the distribution of data significantly to the left. I also removed all entries whose timestamp fell outside of the normal 9:30AM-4:00PM trading day, as there was significantly less trading volume outside of this range. Trading ETFs after hours is generally risky and would be unrepresentative of typical trading behavior. The final step of cleaning was to remove all rows with uncharacteristically high spreads to reduce noise in the data. I arbitrarily set an upper limit of the spread being 2% of the asset's price. In total, I removed 19,624 entries with 'bad' data, or 0.2% of the dataset.

Calculations

To perform calculations, I used the bid and ask fields to calculate absolute spread for each entry. Since the price of the ETF changes over the years, I normalized the spread at each timestamp by converting it to a percentage of the price at that time.

I then identified several (arbitrary) interval lengths: 1-minute, 5-minute, 20-minute, 30-minute, and 1-hour blocks. I did not go below 1 minute because the data would be too granular and uneven, as large orders would cause spikes. Conversely, I did not go above 1 hour because there would not be enough blocks of time to visualize any trends, as a normal trading day only lasts 6.5 hours.

For each of these intervals, I performed a group-by operation to classify spreads into 'bins' of the chosen interval length, then took the mean for each 'bin'. I repeated a similar process for volumes and normalized the bins' sums to a percentage of the total trading day volume for readability and unit consistency, as total volumes at times of the day across years was less useful. Plots were generated from the outputs measuring spread over the course of the trading day, volume over the course of the trading day, and correlation between volume and spread. I used Pearson's formula, the most common method to calculate linear correlation between 2 numeric sets of data, to calculate the correlation coefficient between spread and volume.

Under the assumption that asset values follow a lognormal distribution, continuously compounded returns follow a normal distribution, which reflects the real world. I calculated the continuously compounded return by grouping the intraday data into daily data, took the opening price of each trading day as the price, and computed the natural logarithm of ratios of prices from one day to the next. Continuously compounded returns were grouped (by month/year) and the daily standard deviation of log returns was taken for each group. Lastly, the daily volatility for each group was scaled to monthly/yearly volatilities.

Finally, I checked spread variance based on macro time periods, such as by year and by month. I performed a 2-level group-by operation on the macro period and a chosen time interval, like before. From this, along with the above volatility data, a grouped bar chart was generated that showed changes in intraday spreads (monthly/yearly) together with volatility fluctuations throughout the macro periods (months/years).

Results

Objective and Hypothesis

The main objective of this study was to analyze whether conventional wisdom of buying at the start and end of a trading period (the trading day) holds true.

My hypothesis was that conventional wisdom is correct. I expect the highest volumes at market open and close; this would imply tighter spreads, and therefore I expect the market-neutral (regardless of whether the price will rise or fall) best time to transact is at market open and close.

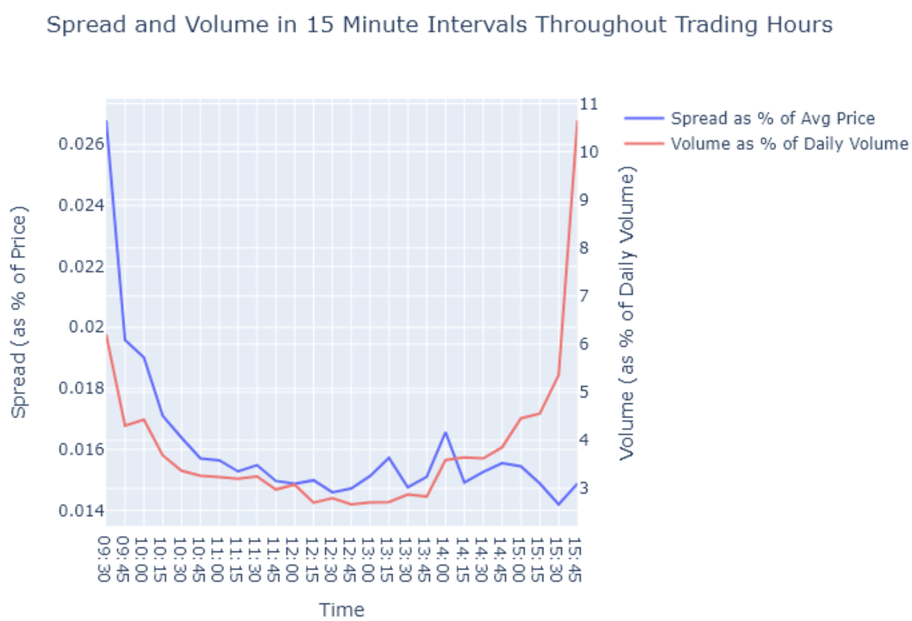
I also expect that higher volatilities correlate with higher spreads because volatility usually increases during periods of rapid market decline or advancement. At these times, high risk and uncertainty and risk create an environment where market makers take advantage by charging higher premiums, thereby widening the spread.

Insights

Spread and Volume

Figure 1 displays the spread (as a percent of the asset price) and volume (as a percent of the total daily volume) throughout trading hours, in 'bins' of 15 minutes.

Figure 1. Spread and Volume Throughout Trading Hours

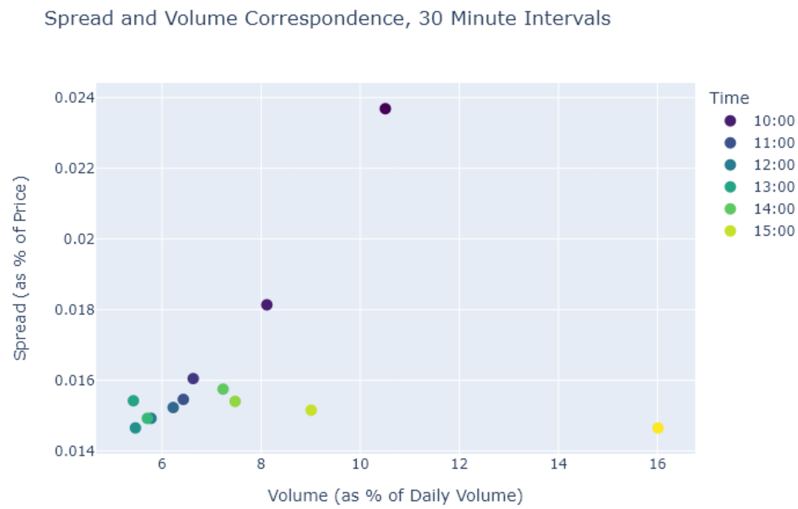


As expected, volume roughly follows a 'U' shaped curve: the highest trading volume is indeed in the periods right after 9:30AM (market open) and right before 4PM (market close). The half hours after market open and before market close together account for over 26% of the daily volume.

Spread is highest at market open, and quickly shrinks until around 11AM, when it stabilizes. From then on, it trends downwards slowly for the remaining of the trading day. This might be due to the fact that while ETFs can be traded after hours (like stocks), lower liquidity means orders may not be filled until the next market open, creating a situation with a wide array of prices investors are willing to buy and sell at.

Interestingly, spread appears to be positively correlated with volume, especially from market open until around 2PM. Figure 2 further illustrates the spread-volume correspondence using 'bins' of 30 minutes.

Figure 2. Spread and Volume Correspondence

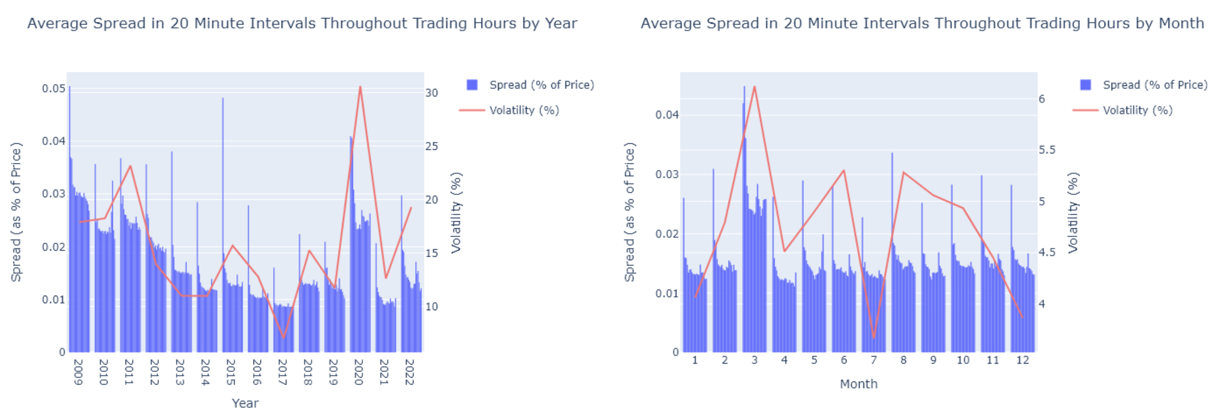


Notably, there is a clear trend among darker colored dots – all the outliers appear to be in the light green to yellow color range, or at the tail end of the trading day. At this time, as shown in Figure 1, volume increases dramatically whereas the spread stays relatively constant. With high liquidity and low spread, near market close would be a good time to execute trades.

Spread and Volatility

Volatility results along with daily spreads across macro time periods (years/months) are illustrated below.

Figures 3 and 4. Spread and Volatility Across Years and Months



There appears to be a slight correlation between general spread levels and volatility: for example, Figure 3 shows that generally, when the yearly volatility is high, spreads throughout the day tend to also be high on average. The correlation between monthly volatility and spread is less clear, as months' spreads do not vary tremendously (for the most part, they average ~0.014% of the asset price throughout the year) and yearly effects dominate (for example, the abnormally high March volatility is due in large part to the onset of COVID in 2020 and the significant disruptions it caused).

Statistical Significance

From Figures 1 and 2, it is clear that there is some degree of correlation between spread and volume, especially for the period before 2PM (14:00). For 30-minute intervals, the Pearson correlation coefficient, r, between spread and volume is 0.242 when considering all trading hours. When excluding times past 2PM, r is much higher at 0.977. From this, we can deduce that the variables are highly correlated for all hours save for near market close. Results are summarized in Table 1 below. It is important to note that while the p-values for pre-2PM data are statistically significant, the caveat is the sample sizes are quite small.

Table 1. Pearson Correlation of Spread and Volume

	Pearson correlation coefficient (r)		P-value	
Interval size (min)	All trading hours	Trading hours before 2PM	All trading hours	Trading hours before 2PM
10	0.286	0.974	0.077	<0.001
15	0.268	0.975	0.186	<0.001
30	0.242	0.977	0.425	<0.001

I also made a rudimentary model for a given interval of time that takes volumes and predicts the spread at any given time in that interval (testing and training sets were randomly sampled at a ratio of 30:70). The coefficient of determination, R^2, is highest for intervals of 10 minutes, at 0.93 (but values varied widely depending on the interval size). The model coefficient was 0.004 with an intercept of 0.007, meaning for a 1% increase in share of daily volume for a particular 10-minute interval, spread is expected to increase 0.004% of the asset price. Larger intervals were unable to be tested because there were not enough data points (bins) to split into testing and training data.

Application

Many players in the market must consider spreads when trading, including HFTs and algorithmic traders, market makers, and mutual funds (to name a few).

HFTs move billions of shares every day, and on every share, there is a little bit lost to the spread. With algorithms that can make thousands of trades every second in an attempt to scalp pennies, the small spread becomes a much more prevalent concern. A similar concept applies to market makers, who profit from the spread by providing liquidity. Market makers remain delta-neutral, and because they don't take directional positions on assets, they must pay close attention to fluctuations in bid/ask prices.

Mutual funds that charge fees with the goal of beating an index by a small percentage often rebalance their entire portfolio, sometimes multiple times a year. In such cases, fractions of percentage points lost to the spread can add up quickly and make it that much more difficult to meet the performance metric required.

Conclusion

The conventional wisdom of buying at market open and close, as well as avoiding trading when volatility is high, is only partially correct from a spread perspective.

Spread is highest at market open and decreases dramatically until around 11AM, after which it levels off until market close. The spread can vary by over 0.01% of the share price over the course of a day – $1 per $10,000 traded. This may not seem like a lot, but to institutional investors who trade millions or billions of dollars, the losses add up quickly.

Thus, to take advantage of small spreads, it is best to trade near market close, but not market open. Lastly, spreads do tend to be larger when volatility is high, as expected.