UNINFORMED TRADER RISK AND MARKET INEFFICIENCY

Stephen Bertone

A Thesis

In

The John Molson School of Business

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Science in Administration (Finance) at

Concordia University

Montreal, Quebec, Canada

April 2011

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Stephen Bertone

Entitled: Uninformed Trader Risk and Market Inefficiency

and submitted in partial fulfillment of the requirements for the degree of

Master of Science Administration (Finance)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. Harjeet Bhabra ____________ Chair

Dr. David Newton ____________ Examiner

Dr. Nilanjan Basu ____________ Examiner

Dr. Rahul Ravi ____________ Supervisor

Approved by Dr. Sandra Betton ____________ Chair of Department

April 14, 2011 Dr. Sanjay Sharma ____________ Dean of Faculty

ABSTRACT

Uninformed trader risk and market inefficiency

Stephen Bertone

This study examines the relationship between uninformed liquidity and intraday market efficiency. We use the SPDR exchange traded fund and its underlying index, the S&P 500, as the instruments for this investigation, and provide evidence showing that uninformed liquidity can impede the price discovery process, thereby making the market relatively inefficient. We find that uninformed trader risk is significant at very short time intervals and it seems to dissipate relatively quickly. The results suggest that this risk is largely related to two systematic factors. First, short term market volatility: larger the volatility, higher the risk. Second, systematic adverse selection problem in the market: higher adverse selection is related to lower liquidity and lower noise trader activity.

Acknowledgements

I would like to express my sincere gratitude towards my supervisor Dr. Rahul Ravi for his guidance and dedication. A heartfelt thank you to my loving parents and brother for all their support, and to my future wife for her endless encouragement.

Table of Contents

List of Figures

List of Tables

Introduction

This study examines the intraday relationship between liquidity (buying and selling of stocks) and efficiency of asset prices. Extant literature identifies liquidity as a largely desirable asset characteristic.[1] There is strong theoretical and empirical support relating higher liquidity to more efficient price formation. However, these assertions largely see liquidity as the transmitter of information. In other words they rely on the role of informed traders. The role of uninformed traders has been largely ignored in these studies. Uninformed trades have the potential to deviate prices away from the fundamentals and to that extent impede price discovery in the market (De Long, Shleifer, Summers and Waldmann (1990), Campbell, Grossman, and Wang, (1993)). While these deviations should cancel out over time, we expect their magnitude to be significant on very short time horizons (intraday).

The findings of this study have important implications for investments because, in the current market where high frequency trades have become the norm and trading time is being measured in milliseconds, it might be possible for at least some investors to exploit this short term market inefficiency.

Any study attempting to explore the role of either the uninformed or the informed traders is likely to be faced with two complications. First, it is observationally difficult to distinguish one group of traders from the other. And second, their activities could be interrelated and therefore endogenous. For example, information flow increases informed traders' activities in the market, which might attract uninformed traders' attention leading to an increase in uninformed trading. We use the deviation between S&P 500 and SPDR as our instruments of exploration. Fundamentally the two assets are identical and therefore any information flow should affect them

[1] One can find excellent review of this literature in Amihud, Mendelson, and Pedersen (2005).

both similarly. Therefore, we argue that non-fundamentals related trading in the market should drive the deviation between the two assets. This allows us to avoid both the above-mentioned issues.

SPDR was designed to track the performance of the S&P 500 index by holding its constituent stocks in the same proportion as the index. By construction each unit of SPDR represents 10% of the index unit. The trust has an open fund structure, which allows for the creation and redemption of shares.[2] The units of SPDRs can be traded in the same manner as regular stocks in that they can be bought on margin sold short, and are option eligible. Unlike mutual funds however, there are no fractional units. Thus the minimum trade size is one unit of the SPDR. Also, the management fees for SPDRs are lower than those charged by the index mutual funds.[3] ETFs in general and SPDRs in particular are popular and fairly liquid assets. According to BlackRock, owner and manager of iShares, the US ETF Industry has grown from 500 million $ of assets invested in 1993 to over 900 billion $ invested in January 2011. As of January 2011 SPDR had over 93 billion $ under management making it the largest ETF in the world.[4]

SPDR units trade like common stock and therefore they are susceptible to sentiments and attention biases in the market (Barber and Odean, 2008). The underlying assets in the SPDRs are common stocks (more specifically 500 stocks which makeup the S&P 500 index). Each of these stocks would also be individually susceptible to sentiments and attention biases of their own. Assuming that these biases across all stocks are less than perfectly correlated, we should see a divergence between the value of the underlying basket and the value of the fund. The level of

---

[2] There are some frictions in the process. For example SPDR units can be created or redeemed only in multiples of 50,000 units and there is a fee charged for this transaction

[3] However, brokerage fees may be incurred by retail investors when trading SPDR

[4] "ETF Landscape: Industry Highlights" (January 2010) BlackRock retrieved from : http://www.blackrockinternational.com/content/groups/internationalsite/documents/literature/etfl_industryhilight_jan11.pdf

divergence should be related to the level of trading. To the extent that non-fundamental related price movements are expected to correct relatively quickly, these divergences should be greater at shorter time horizons. In other words, the comovement between the index and the ETF returns should increase as we increase the measurement time interval. An equivalent characterization of this argument is that the differing price processes of the index and the underlying ETF should create distinct intraday return volatility for the two assets. As the measurement time length increases, the standard deviations of the assets should converge.

The standard deviation of daily returns over entire sample period (1996-2003) for SPDR and the S&P 500 are found to be 0.012685 and 0.012189 respectively. These values are not statistically different from one another. When we look at intraday returns (second by second), SPDR standard deviation is found to be 0.00025 and S&P 500 standard deviation is 0.00072. The intraday volatility of the index is three times that of the corresponding ETF. The disappearance of this difference when the measurement time is increased from one second to one day, lends some preliminary support to our conjecture that non-informational trades are likely to drive prices away from fundamentals for short time intervals and that these deviations should correct relatively quickly.

To explore the SPDR ability to track the S&P 500, we use a simple market model regression with the SPDR return as the dependent variable and the index return as the independent variable. The ratio of the explained sum of squares to the total sum of squares (R^2) from this model is used as a measure of comovement between the two assets.[5] We find that at the daily level, SPDR tracks the index reasonably well with an average R^2 value of 0.922. We observe a sharp decline in the R^2 values when the two assets are compared at intraday levels. We find hourly R^2 values to

[5] This measure is inspired by Morck, Yeung and Yu (2000)

be 0.9, minute by minute R^2 values to be 0.132 and second by second R^2 values to be essentially zero.[6]

The remainder of this paper is organized as follows. Section 2 provides a brief review of the literature and develops the theoretical basis for this study. Section 3 describes the data analyzed as well as the methodology employed. Section 4 describes and discusses the results and section 5 concludes the paper with a brief discussion and the implications of the findings of this study.

1. Background

Fama (1970) suggests that all markets are not equal as it pertains to information efficiency. He notes three forms of efficiency, weak form, semi-strong form and strong form. In the weak form market prices only reflect past prices. In the semi strong form market prices reflect all publicly available information and in the strong form all information both public and non public is reflected in prices. Markets are largely considered to be semi-strong form efficient. Introductory finance textbooks present trading activities by informed agents (arbitrageurs) as the conduit of market efficiency. However, these efficiency arguments ignore the role of noise traders (uninformed, not-necessarily rational agents). De Long, Shleifer, Summers and Waldmann (1990) present a model of noise traders risk in the market, where they argue that the unpredictability of noise traders' beliefs creates a risk in the price of the asset that deters rational arbitrageurs from aggressively betting against them. As a result, prices can diverge significantly from fundamental values even in the absence of fundamental risk.

Going through the extent liquidity literature leaves the impression that liquidity in all forms is a desirable asset characteristic. Greater liquidity is always desirable over lesser liquidity.

[6] Second by second R^2 values are not reported in the tables. For sake of brevity, we limit our reported analysis to, starting with per minute and increasing to per day comparison of the two asset returns.

Companies prefer greater liquidity, as it leads to lower costs of capital (Amihud and Mendelson 1986; Brennan, Chordia and Subrahmanyam 1998; Liu 2006). Market makers prefer liquidity as it would potentially reduces their risk of market making, and investors prefer liquidity because higher liquidity would allow them to adjust or close their position faster and cheaper.[7] Branch and Freed (1977) as well as Copeland and Galai (1987) use trading volume to proxy for asset liquidity. They find that transaction cost is directly related to trading volume. As trading volume increases, measures of transaction cost such as brokerage fees, execution costs and bid ask spreads decrease. While these findings might be true on average, noise trading models such as De Long, Shleifer, Summers and Waldmann (1990) seem to present a note of caution particularly including periods of high uninformed trader activity.

The motivation for this study comes from the understanding that the price discovery and the riskiness (volatility) of a security should be a function of trading volume (Karpoff, 1987). Based on its source and its effect on asset prices, we can broadly classify trading volume into two categories. First, trading volume as a transmitter of information, whereby it aids price discovery, and second, trading volume as a transmitter of noise, in which case it would inject noise into the price process and thereby, could potentially drive prices away from the fundamentals. Trading volume generated by informed traders may be classified among those serving as the information transmitters, while the uninformed trading (noise trading or liquidity trading) volume could potentially be the transmitter of noise.

Trading volume in total has increased exponentially over the past 50 years. Volatility of asset returns has also increased with the increased trading (Wei and Zhang 2006 and Irvine and Pontiff

[7] A caveat is in order here, whereby volume captures just one aspect of liquidity (depth). Order-imbalance could make the asset less desirable even in the presence of high one sided demand/supply. Liquidity in its entirety has three dimensions: depth, immediacy and resilience (Kyle, 1985)

2009). While it is difficult to attribute this increase in trading volume to either informed or uninformed sources, some evidence exists which argue that the increase has come largely from traders who might be considered somewhere in between the informed and the uninformed (Chordia Huh and Subrahmanyam, 2007).[8] This study explores the impact of uninformed trading in the market by exploring the tracking errors between SPDR and its underlying index (S&P 500). Fundamentally the two assets are identical and therefore any fundamental information flow should affect the two assets identically. Therefore, non-fundamentals related trading in the market should drive any deviation between the values of the two assets. To the extent that the above-mentioned gray-zone trades have very little to nothing to do with the asset fundamentals, the design of our study would capture their effects among the non-fundamental driven trades.

2. Data and Methodology

2.1 Data

The study involves analyzing intraday trade and quote data for each of the S&P 500 constituent stocks. We start by selecting January 1st 1996 to December 31st 2003 as the sample period for this study. The choice of the time-period is driven by our desire to keep the data analysis manageable, without sacrificing any generalisability of the results. The sample period spans across a bull market (rise of the technology bubble) and a bear market (period after the collapse of the bubble on March 10th, 2000) and therefore in some sense covers a full business cycle. It allows the study to span across various important financial market changes, which came into effect in the late 90s and which, arguably had important implications for publicly traded stocks. The decimalization process in the US stock markets started in August 2000 and was

---

[8] Algorithmic traders have been classified among these 'in-between' or gray zone traders. Hendershott, Jones and Menkveld (2010) note that with the rise of algorithmic trading, computer trading now accounts for over 70% of all trading.

completed by April 9th 2001. Regulation fair disclosure came into existence in August 2000 and the Sarbanes-Oxley Act was enacted on July 30th 2002.

The study requires comparing the intraday levels and returns of the S&P 500 index with the SPDR unit. Therefore, our sample consists of all the constituents of the S&P 500 index (in the sample period) and the SPDR. We compile the list of daily constituents of the S&P500 index for all dates in the sample period using the index additions and deletions information obtained from the Standard and Poor's. Intraday trades and quotes for all the sample stocks are obtained from the NYSE TAQ database. Several filters are employed to ensure the validity of the trade and the quote data.[9] The TAQ database does not eliminate auto quotes (passive quotes by secondary market dealers), which can cause quoted spreads to be artificially inflated. Since reliable filtering out of auto quotes in the TAQ data is not possible, only BBO eligible (best bid or offer) primary market quotes are used.[10] Quotes established before market open or close are also discarded.

Daily shares outstanding for the index constituents are obtained from CRSP. We use several measures of market-wide volatility, systematic liquidity and market sentiments to explore the deviations between the SPDR return and the return on the underlying index (S&P 500). Sadka (2006) permanent variable factor is used as a measure of systematic liquidity. This data is obtained from Wharton Research Data Services (WRDS) 'Fama French & Liquidity Factors' dataset. We use investor sentiment index as calculated in Baker and Wurgler (2006). The data is

[9] We drop all trades with correction indication other than 0 and 1, retain only those trades for which the condition is B,J,K or S. We also drop all trades with a non positive trade size or price. Finally we omit all trades recorded before market opening time or after market closing time. Negative bid ask spreads and transaction prices are also eliminated. We eliminate all quotes where quoted spread is greater than 20% or quote midpoint, where quote midpoint is greater than 10$ or quoted spread is greater than 2$ when quote midpoint is less than 10$. We also eliminate all quotes where ask or bid moves by more than 50%. Trades with non standard settlement conditions (A,C,D,N,O,R and Z) are excluded.

[10] All quotes with conditions 5,7,8,9,11,13,14,16,17,19,20,27,28,29 are excluded.

obtained from Jeffrey Wurgler's website at Stern, NYU. The volatility index (VIX) is used as measure of short-term volatility in the market. Other measures of sentiment used in the study include the CBOE put call ratios and the Advance Decline data is obtained from Bloomberg.

2.2 Methodology

This study attempts to explore the impact of non-informational trades in the price process of financial assets in the market. The tracking error between SPDR and the S&P 500 index is used as the instrument for this exploration. The first step in this analysis involves constructing a second by second time-series of the S&P 500 index for all trading days in the sample period. We obtain details of 1,979,964,070 trades across all the index constituents in the sample period, from TAQ. We lose 61,494 trades on imposition of the validity filters. Thus the second by second index time series is constructed using the remaining 1,979,902,576 trades. We also reconstruct a second by second time series for the SPDR ETF. This series is constructed using 13,415,847 SPDR trades, which occurred between January 1st 1996 and December 31st, 2003. We also constructed a second by second time series of the SPDR using the quotes mid-point instead of the transaction prices. 97,291,760 SPDR quotes obtained from TAQ are used to construct this series.

The S&P 500 index is a market cap weighted index where the constituents' market caps are summed and then divided by a divisor to get an index level.[11] The divisor is used in order to scale the index and keep the index comparable over time by maintaining a link to the base period

[11] The index changed from market cap weighted to free-float weighted in March 2005.

value of the index. Note that the divisor as well as shares outstanding used is constant throughout the day. [12] Equation (1) provides the formula for the index calculation.

$$Index = \frac{\sum \left(P_j \times S_j\right)}{Divisor} \tag{1}$$

Where P_j is the price of stock j and S_j is the shares outstanding for stock j. Since the divisor remains constant through the day, one-second intraday index return r_t can be calculated using equation (2).

$$r_t = \ln\left(\frac{\sum \left(P_{t,j} \times S_j\right)}{\sum \left(P_{t-1,j} \times S_j\right)} \right) \tag{2}$$

Where $P_{t,j}$ is the price of stock j at time t. If there is no trade at time t we use the most recent trade price before time t. [13] We assume that the share outstanding for a particular index constituent remains constant through a given day. Daily shares outstanding are obtained from CRSP daily dataset.

2.2.1 *Comovement of S&P 500 and SPDR*

Our measure of comovement is inspired by the R^2 measure proposed by Morck, Yeung and Yu (2000). We estimate this measure as the ratio of the explained sum of squares to the total sum of squares (R^2) in the market model described in equation (3).

$$r_{SPDR,t} = \alpha + \beta \times r_{index,t} + \varepsilon_t \tag{3}$$

[12] "Index Mathematics: a very short course" by David M. Blitzer, Managing Director and Chairman of the Index Committee. Standard & Poor's retrieved from : http://www2.standardandpoors.com/spf/pdf/index/Index%20Mathematics%2012-05.pdf

[13] If multiple trades occurred at same second we used mean trade price (results do not change when using medians.)

Where $r_{SPDR,t}$ is the return on SPDR at time t, and $r_{index,t}$ is the return on the S&P 500 Index at time t. Given that the SPDRs are by design constructed to track the index, ideally R^2 should always be very close to one. We expect R^2 to deviate from one due to differences in the level of trading between the two assets.

We calculate the time series of one-second holding period returns for both the index and the SPDR. Since we are using continuously compounded returns (equation 2), one-minute return is calculated by adding the sixty one-second returns within that minute. Similarly we create five-minute, ten-minute, fifteen-minute, thirty-minute, sixty-minute and daily holding period return series of the SPDR and the S&P 500 index.[14] Equation (3) is used to estimate the comovement (R^2) between the index and the SPDR for the different periods (one-second, one-minute, etc.).

The first set of R^2 values are estimated using a pooled regression across the complete sample period. We repeat the analysis by year, across each month and across each day of the week. Since, the SPDR and the underlying index are fundamentally identical, all less than perfect comovement should be related to noise trading (non fundamental/ informational trades). These should cancel out over longer time periods. Therefore, we expect to see higher level of comovement between longer holding-period returns. In other words, we expect the R^2 to increase as the holding period increases from one-second to one-minute, five-minutes etc.

[14] Market opens at 9:30 and closes at 16:00 hours. In creating the sixty-minute return series, we define the first interval as 9:30 to 11:00. This is an interval of 90 minutes. This should not affect the study because we are not carrying out any inter-interval comparison.

2.2.2 Tracking errors

The comovement between the SPDR units and the underlying index is found to be less than perfect. In order to explore this less than perfect comovement, we use three related measures of tracking errors. The first measure is the sum of the absolute error term from the equation (3) residual.

$$TE_i = \sum_t \left| \varepsilon_{t,i} \right| \tag{4}$$

Where TE_i is the tracking error for month i. We estimate equation (3) for each of the pairs of the return series (SPDR and S&P 500) over various time intervals: one minute five minute etc. for each month in the sample period. $\varepsilon_{t,i}$ are the residuals series generated for month i. The tracking error for month i is calculated as the sum of the absolute values of all the residuals for month i. Measure 2 is the standard deviation of the residuals $\varepsilon_{t,i}$. The third measure is $\left(- R_i^2 \right)$, where R_i^2 is the coefficient of determination estimated from equation 3 for month i. We find similar results for all three measures. For the sake of brevity, we present only the results pertaining to the first measure.

The SPDR and the underlying index are fundamentally identical, therefore, tracking errors should not be related to the asset fundamentals. As argued earlier, we expect the tracking errors to be driven by noise and systematic factors that could potentially translate into trading biases. The variables of interest identified here include Sadka permanent variable factor, a measure of systematic liquidity; Baker and Wurgler Investor sentiment index; short term market volatility, as measured by the volatility index VIX, and two other investor sentiment measures: CBOE put call ratios and the Advance Decline ratio. The last measure is the ratio of MAX(number of stocks advancing, number of stocks declining), and the total number of stocks trading in the market. This variable should measure the level of herding in the market.[15] The results presented are computed using monthly tracking errors (estimation process explained above). The choice of month as the measurement unit is determined by the nature of the Sadka liquidity data and the Baker and Wurgler's Investor sentiment index. Both these series are monthly. We propose a simple OLS model (equation 5) to explore the properties of the tracking error.

$$TE_i = \quad + \quad \times \quad + \quad \times \quad + \quad \times \quad + \quad \times \quad + \quad \times \quad + \quad \ldots \quad \ldots \quad +$$

Where TE_i is the tracking error in month i, $\ln VIX$ is the natural logarithm of VIX, PVF is the Sadka (2006) Permanent Variable Factor, PC is the CBOE putcall ratio, AD is the advance decline ratio, SI is the Baker & Wurgler (2006) sentiment index and $D_{96\ldots}$ represent a set of seven year dummies.

2.2.3 A naïve trading strategy

This section attempts to develop a simple trading strategy to see if the less than perfect comovement between the two assets (SPDR and the index) can be profitably exploited. By construction each SPDR unit is supposed to represent 10% of the index level. If the relationship between the SPDR and the index deviates from this 10%, it would suggest a deviation from fundamental value and therefore a potential arbitrage opportunity. In order to implement this strategy, we estimate the relationship ratio (equation 6) at every one-second interval in a given trading day.

---

[15] The word herding is being loosely used here to denote the trading behavior of the majority market participants. The variable will be high if the fraction of stocks moving in one direction (increasing or declining) goes up. Therefore, herding as used here may be understood as investor comovement.

$$IS_t = \frac{Index_t}{SPDR_t} \tag{6}$$

Where IS_t is the index to SPDR ratio at time t. A value of greater than 10 for this ratio would suggest that either the index is too high or the SPDR is too low. A correction would require either the index level to move down or the SPDR level to increase. The situation can be exploited by taking a long position in the SPDR and simultaneously a short position in the S&P 500. The opposite positions can be taken if the index SPDR ratio is less than 10. At the end of the pre-determined time interval we would close the existing positions and simultaneously open a new position, depending on the calculated ratio at that time. We attempt to implement this strategy every minute, every 5 minutes, every 10 minutes, every 15 minutes, every 30 minutes, every 60 minutes and every day. Since we expect greater level of mispricing at shorter time interval, we expect the profitability of the strategy to increase as the trading time interval decreases. Therefore, among the stated time intervals, trading every minute should be most profitable and strategy returns should monotonically decrease as the trading interval increases to 5-minute, 10-minute, etc.

3. Results

Table I presents descriptive statistics pertaining to the data set used in our study. It described the average number of trades per day of the week, per month and per year as well as it notes the median values and standard deviations. The table also includes the average share price of SPDR and the average market cap of S&P 500 firms per year. The average price per share of S&P 500 component stock is 43.00$ although the prices range from 0.36$ to 659.00$. The average shares outstanding is about 410 million shares with a range of 1.9 million to 11 billion shares outstanding. The average number of trades are 275 million per year with a range of 75 million to 464 million trades per year. The SPDR share prices average 106.91$ with a range of 59.97$ to 153.56$. The average shares outstanding are 131 million with a range of 9.2 million to 465 million. There are on average 1.7 million trades per year with a range of 41,695 trades to 5.9 million trades per year.

Table II presents Pearson and Spearman's correlations between the various variables use in this study in exploring the cause of intraday deviation between the SPDR and the S&P 500. In general the correlations are not too high. We find a negative correlation between VIX and PVF which is consistent with the existing literature whereby volatility and liquidity are positively correlated and adverse selection risk (PVF) and liquidity are negatively correlated.

3.1 Comovement between SPDR and S&P 500 index

Figure 1 presents the level of comovement between the SPDR and the S&P 500 index, as a function of holding period. Equation (3) is estimated for holding periods ranging from one minute to one day. The R^2 from these estimations gives a measure of comovement between the SPDR and the index returns. We find that as the holding period increases, the level of comovement increases monotonically. This result is in concurrence with our hypothesis that any deviation between the two asset values is likely to be caused by non-fundamental related trading in the market. By design, all fundamental related changes (permanent changes) in the two asset values should comove perfectly. All other changes should be transitory. As the length of time increases, the transitory effects should cancel out and the assets must converge towards their fundamental values leading to an increase in comovement.

Table III complements figure 1 by presenting the R^2 values by year (Panels A and B), by month (panels C and D) and by day of the week (panels E and F) for SPDR trades and SPDR quote midpoints. Panel G presents the level of comovement across the complete sample period. Overall, the findings are consistent with figure 1, across each year, month and day of the week subgroup. On average at one minute we find R^2 values for trades to be 0.184, at 5 minute intervals it increases to 0.569 and at 1 day the R^2 is 0.917. Using SPDR quote midpoints instead of SPDR transaction prices, the 1 minute average R^2 is 0.136 which increases to 0.545 at 5 minutes and 0.931 at one day. Slicing the sample period by year, we find considerable variation between years. However, within any given year, the pattern observed in figure 1 holds. At shorter holding periods, the level of comovement increases from 1996 through 2002 and then seems to decline in 2003. The results are consistent using quote mid-points. Panels C through F repeat the above exercise, controlling for any potential month of the year, or day of the week effects. All results are found to be consistent with the general findings of figure 1.

Ackert and Tian (2000) analyzed the relationship between SPDR and the S&P 500 between 1993-1996. Using daily prices they found that the SPDR ETF did not trade at economically significant discount to the S&P 500.[16] They conclude that arbitrage forces are strong enough to eliminate the impact of noise traders. They note that the redemption feature of SPDR potentially plays a role in keeping prices efficient. They attempt to explain the (economically insignificant) discount of SPDR vis-à-vis the S&P 500 and found investor sentiment to be an insignificant predictor. We find that the comovement between the SPDR and

[16] Low tracking error between ETF and underlying index, at daily holding period horizon is also supported by several other studies such as Tse & Martinez (2007) and Rompotis (2010). These studies find that unlike closed end funds, ETFs track their NAV rather closely and any observed discount is economically insignificant. They attribute this difference between the closed end funds and the ETFs to the redemption feature of ETFs.

the S&P 500 daily returns is significantly higher than the corresponding one-minute returns (all intra-day returns).

3.2 Explaining the low comovement

Using a simple polynomial regression (equation 5), this subsection attempts to explain the low comovement between the SPDR and the underlying index. Fundamentally these two assets are identical and therefore reduced comovement is unlikely to be related any asset fundamentals. We search for the explanation among various systematic factors. To the extent that the low comovement is likely to be caused by non-fundamental related trading in the market, any explanation for this should lie among factors which are likely to affect these trades. Noise traders are likely to be swayed by sentiments in the market.[17] We use three different measures of sentiment in the market, Baker & Wurgler (2006) sentiment index (SI), the ratio of the volume of put options traded to the volume of call options traded at the CBOE (PC), and the advance decline ratio (AD). Non-fundamental related traders are also likely to be sensitive to the general frothiness of the market. We use logarithm of the volatility index (lnVIX) to proxy for this effect. These traders are likely to be affected by market-wide news releases. Sadka (2006) permanent variable factor (PVF) is the priced component of the systematic adverse selection risk. We use this variable to capture any systematic information effect. The analysis in Table III, panels A and B show significant volatility in comovement levels across years. This could be related to market conditions (bull vs. bear markets) or it could be related to the market learning about the SPDRs. At very short holding periods, we do observe a somewhat increasing comovement across the

[17] Among studies that support this line of reasoning, Lee, Jiang, and Indro (2002) look at the relationship between volatility, returns and sentiment. They find that bullish (bearish) changes in sentiment result in downward (upward) adjustments in volatility. Thus bullish markets lead to lower volatility and bearish leads to higher volatility. Following their work we expect market sentiments to be related to tracking errors which we are attempting to explain in this section.

sample period (Table III, panel A and B, 1-minute results), however this trend is not so clear at 5-minutes or higher holding periods. We include seven year dummies representing years 1996 through 2002 (2003 is the base case) in equation 5 to capture the variability across the years.

We estimate equation 5 and its various reduced forms. Table IV presents the estimated coefficients. We find positive, and highly significant coefficient on lnVIX. As mentioned VIX is a benchmark of short term expected volatility in the market (S&P 500). Increasing VIX would suggest more volatile market. To the extent that noise trader activities are by definition uncorrelated, it should be associated positively with volatility. Therefore, the positive coefficient on lnVIX may be interpreted as higher tracking error in presence of more noise traders. The Sadka permanent variable factor (PVF) is a measure of market wide (systematic) adverse selection cost of trading. Alternatively it may be interpreted as an inverse measure of market liquidity (higher adverse selection problem leads to lower liquidity). We find that this variable is significantly negatively associated with tracking error. In other words, periods of low liquidity are associated with lower tracking errors.

Our analysis fails to find any significant relationship between market sentiment and the tracking error. The reduced form model (Table IV, model 4) does find a significant coefficient for the advance decline variable (AD). Although the coefficient becomes insignificant in the expanded model (model 7), the coefficient continues to be positive. A possible explanation for the positive relationship could be that more one sided movement in the market is likely to attract more uninformed trading. Large number of declining stocks could attract large scale selling while market advances might attract large scale buying. The model explains close to 79% of the variability in the tracking error.

3.3 Exploiting the mispricing

Table V presents the ratio of S&P 500 index level/SPDR price (hence referred to as SPDR ratio). As discussed in section 3, this value should theoretically be equal to 10. However, we find evidence of significant intraday deviation between the values of the SPDR and the underlying index. These deviations also drive the SPDR ratios away from 10. Table 5 reports yearly means, medians, and standard deviations for the SPDR ratio.

Examining the means medians and standard deviations calculated in Table V it is interesting to note that neither the mean, nor the median is equal to 10. However, the estimated statistics are for the most part within one standard deviation of 10. Surprisingly, all the means and the medians are found to be slightly lower than 10, which would suggest that the SPDR is trading at a premium. This observation is consistent with the findings of Aber, Li Can (2009). They look at four iShares ETF's and analyzed premium/discount, daily return and tracking error from fund inception up until 14 December 2006. They find that ETFs are more likely to trade at a premium than at a discount to NAV.

Here we attempt to exploit the mispricing by defining and implementing a naïve trading strategy using the information from the SPDR ratios. The SPDRs are by design constructed to provide investors with a security whose initial market value will approximate one-tenth (1/10th) the value of the S&P 500 index. Table V reveals that from January 1, 1996, until December 31, 2003, the ratio of the index value to the per-share price of SPDRs ranged from 9.922 to 9.976. Although these numbers are not very far from the 10:1 ratio, the deviations represent a violation of law of one price present which presents opportunities for arbitrage. A value of greater than 10 denotes either overvalued index or undervalued SPDR. Similarly, value of less than 10 would suggest overvalued SPDR or undervalued index. A possible mode of exploiting this potential mispricing would involve large investors, for example "authorized participants," taking long positions in the underpriced security and short positions in the overpriced security, and unwind those positions by transacting with the ETF-issuing trust.[18]

We implement the above trading strategy at various pre-determined time intervals. For example the 60-second holding period strategy would involve calculating the SPDR ratio at 9:30 (market open) and taking the long position in the overpriced asset and short position in the underpriced asset. After 60 seconds, unwind the existing position and recalculate the SPDR ratio and based on the new ratio take new positions in the two assets. This is repeated through the day until market close, for every day in the sample period. Table VI presents the annualized trading strategy returns for various time intervals and compares them to the annualized risk free and S&P 500 returns. We find that the trading strategy outperformed the S&P 500 in each time interval in each year observed by a large margin. Moreover, the strategy never earned a negative return over an entire year. The strategy's performance is shown to increase as the time interval decreases which is consistent with our hypothesis that the ETF has difficulty tracking the index at small time intervals and therefore greater mispricing. The average return over the entire period for the S&P 500 was 8.45% per year whereas our trading strategy produced average returns of between 47% and 475% per year for one day and one minute transactions respectively.[19] Figure 2 plots the annualized returns of the above trading strategy at all intervals and compares it to the S&P 500 returns. We note that the 60 second time interval far outperformed all other time intervals over the entire period. If we look at the weakest performing time period, the one day

[18] "Authorized Participants" are entities chosen by an exchange-traded fund's (ETF) sponsor to undertake the responsibility of obtaining the underlying assets needed to create an ETF. They are typically large institutional organizations, such as market makers or specialists. This is the primary mechanism by which the ETF and underlying index remain closely tied together

[19] See Figure 2: Annualized strategy returns for a visualization of our strategy as compared to the S&P 500

interval, the cumulative returns of 376 % for the trading strategy far outperform the 68% return of the S&P 500.

The strategy implemented above is not riskless arbitrage (mispricing might not disappear for potentially long time intervals). Therefore, we need to take a look at some basic risk and return relationship to get a more realistic idea of the performance of the trading strategy. We use Sharpe ratio as the instrument for comparing the risk and return of the trading strategy. Table VII presents the yearly Sharpe ratios obtained for our naive trading strategy across all the years in the sample period. The reward to risk ratio is greater for our naïve strategy compared to the index; in essence the naïve strategy beats the index. The Sharpe ratios of the S&P 500 ranges from -0.057 in 2002 to a maximum of 0.094 in 2003. In contrast the minimum Sharpe ratio for the trading strategy is in 1997 using hourly data the strategy produced a Sharpe ratio of 0.11. The maximum Sharpe ratio obtained occurred in 1998 at the 5 minute interval where the Sharpe ratio is estimated to be 0.63. The average Sharpe ratio of the S&P 500 over the entire sample period (1996-2003) is 0.025 and that for the naïve strategy is found to be 0.4.

We replicate the above trading strategy using SPDR quote midpoints instead of the trade prices. This partially address the issue of spurious returns generated by bid-ask bounces. The resulting returns are presented in table VIII, while the risk return analysis is presented in table IX. The findings are consistent with those obtained using trade prices.

4. Conclusion

This study examines the relationship between uninformed liquidity and intraday market efficiency. The motivation comes from various models of noise trader risk in the market which argue that uninformed trades have the potential to deviate prices away from their fundamentals.

We use the SPDR exchange traded fund and its underlying index, the S&P 500, as the instruments for this investigation, and provide evidence showing that uninformed liquidity can impede price discovery process, thereby making the market relatively inefficient, intraday.

Our main findings are that uninformed trader risk is significant at very short time intervals and it seems to dissipate relatively quickly. The comovement between the ETF and the index (as measured by the R^2 from the regression of SPDR on the index) is nonexistent when measured at 1-second interval. It increases to about 18% at 1-minute interval, 57% at 5-minute and 72% at 10-minutes and 92% in daily returns.[20] The results suggest that the ETF is not good at tracking the underlying index at short intraday time intervals. The tracking error can be mostly explained by two factors. First, short term market volatility: larger the volatility, more the tracking errors. Second, systematic adverse selection problem in the market: higher adverse selection is related to lower liquidity and lower tracking error.

We attempt to explore if the observed tracking error can be economically exploited by investors. We develop a naïve trading strategy which involves calculating the ratio of the S&P 500 index to SPDR at the beginning of every trading day in the sample period. By design each SPDR unit should be 10% of the index. If the ratio is over 10%, it would suggest either overvalued index basket or undervalued SPDR (Ratio below 10% would suggest vice versa). We take a long position in the undervalued asset and a short position in the overvalued asset. At the end of pre-determined time intervals we would close our existing positions and simultaneously open new positions, depending on the calculated ratio at that time. This strategy was able to

[20] Sahalia, Mykland and and Zhang (2005) show that there are instances where the market microstructure noise contained in high frequency financial data can exhibit serial correlation. While this could drive some of our results at ultra high frequency such as one second or one minute it is unlikely to affect lower frequency estimation intervals such as an hour or a day. Moreover, the monotonic pattern displayed in the results across various estimation intervals (one second to one day) reduces the chances of our results being entirely driven by autocorrelated noise.

outperform the S&P 500 by a large margin at all intraday time intervals examined on both a nominal and risk adjusted (Sharpe ratio) basis. One issue to note is that we ignore transaction costs in the calculation of the returns from our naïve strategy. The stated strategy involves considerable trading and may not be a viable option for most retail investors. However, for institutional investors and more specifically authorized participants for the given ETF should be able to exploit the observed mispricing. This is particularly true in the current market where the transaction costs for institutional with a seat on the exchange is all time low and high frequency trading has become the norm in the market. Moreover, we present here a rather naïve strategy for the sake of demonstration. Trading costs can be reduced by simple modifications to this strategy such as, trade only when the index to SPDR ratio flips from less than 10 to greater than 10 and vice versa.

In conclusion, our results suggest that not all liquidity lead to efficient price discovery. Liquidity coming from non-fundamental related trades can impede price discovery and make the market inefficient at intraday levels. These inefficiencies are non-fundamental related and therefore will on most times correct rather quickly. Further research is warranted to explore profitable trading strategies to take advantage of these transitory price inefficiencies in the market.

References

Aber, Jack W., Dan Li, and Luc Can, 2009, Price volatility and tracking ability of ETFs, *Journal of Asset Management* 10, 210-221.

Ackert, Lucy F., and Yisong S. Tian, 2000, Arbitrage and valuation in the market for Standard and Poor's depository receipts, *Financial Management* 29, 71–88.

Ait-Sahalia, Yacine, Per A. Mykland and Lan Zhang., 2010, Ultra high frequency volatility estimation with dependent microstructure noise, *Journal of Econometrics*,160, 160-175.

Amihud, Yakov, and Haim Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223-249.

Amihud, Yakov, Haim Mendelson, and Lasse H. Pedersen, 2005, Liquidity and asset prices, *Foundations and Trends in Finance* 1, 269-364.

Baker, Malcolm, and Jeff Wurgler, 2006, Investor Sentiment and the Cross-Section of Stock Returns, *Journal of Finance* 61, 1645–80.

Barber, Brad, and Terrance Odean, 2000, Trading is hazardous to your health: The common stock investment performance of individual investors, *Journal of Finance* 55, 773-806.

Barber, Brad, Terrance Odean, 2008, All that glitters: the effect of attention on the buying behavior of individual and institutional investors, *Review of Financial Studies* 21, 785–818.

BlackRock, 2011 January, ETF Landscape: Industry Highlights, retrieved from : http://www.blackrockinternational.com/content/groups/internationalsite/documents/literature/etfl_industryhilight_jan11.pdf

Branch, Ben, and Walter Freed, 1977, Bid-asked spreads on the AMEX and the big board, *Journal of Finance* 32, 159–63.

Brennan, Michael J., Tarun Chordia, and Avanidhar Subrahmanyam, 1998, Alternative factor specifications, security characteristics, and the cross-section of expected stock returns, *Journal of Financial Economics* 49, 345-373.

Campbell, John Y., Sanford J. Grossman, and Jiang Wang, 1993, Trading volume and serial correlation in stock returns, *Quarterly Journal of Economics* 108, 905-939.

Chordia, Tarun, Sahn-Wook Huh, and Avanidhar Subrahmanyam, 2007, The cross-section of expected trading activity, *Review of Financial Studies* 20, 709-740.

Copeland, Thomas E., and Dan Galai, 1983, Information effects on the Bid-Ask Spread, *Journal* *of Finance* 38, 1457-1469.

De Long, J. Bradford, Andrei Shleifer, Lawrence. H. Summers, and Robert, J. Waldmann, 1990, Noise Trader Risk in Financial Markets, *Journal of Political Economy* 98, 703–38.

Fama, Eugene, F. 1970. Efficient Capital Markets: A Review of Theory and Empirical Work, *Journal of Finance* 25, 383– 423.

Hendershott, Terrence, Charles M. Jones, and Albert J. Menkveld, 2011, Does algorithmic trading improve liquidity?, *Journal of Finance* 66, 1–33.

Irvine, Paul J., and Jeffrey Pontiff, 2009, Idiosyncratic return volatility, cash flows, and product market competition, *Review of Financial Studies* 22, 1149-1177.

Karpoff, Jonathan M., 1987, The relation between price changes and trading volume: A survey, *Journal of Financial and Quantitative Analysis* 22, 109-126.

Kyle, Albert S., 1985, Continuous Auctions and Insider Trading, *Econometrica* 53, 1315-1335.

Lee, W.Y., Christine X. Jiang, & Daniel C. Indro, 2002, Stock market volatility, excess returns, and the role of investor sentiment, *Journal of Banking and Finance* 26, 2277-2299.

Liu, Weimin, 2006, A liquidity-augmented capital asset pricing model, *Journal of Financial Economics* 82, 631-671.

Morck, Randall, Bernard. Yeung, and Wayne. Yu, 2000, The information content of stock markets: Why do developing markets have synchronous stock price movements?, *Journal of Financial Economics* 58, 215-260.

Rompotis, Gerasimo, G., 2010 Does premium impact Exchange-Traded Funds' returns? Evidence from iShares, *Journal of Asset Management* 11, 298–308.

Sadka, Ronnie., 2006, Momentum and Post-earnings-announcement Drift Anomalies: The Role of Liquidity Risk. *Journal of Financial Economics* 80, 309–49.

Schiller, Robert. J., 1981, Do stock prices move too much to be justified by the subsequent changes in dividends?, *American Economic Review* 71, 421-436.

Shleifer, Andrei and Robert Vishny, 1997, The limits of arbitrage, *Journal of Finance* 52, 35-55.

Tse, Yiuman, and Valeria Martinez, 2007, Price discovery and informational efficiency of international iShares funds, *Global Finance Journal* 18, 1–15.

Wei, Steven X., and Chu Zhang, 2006, Why did individual stocks become more volatile?, *Journal of Business* 79, 259–92.

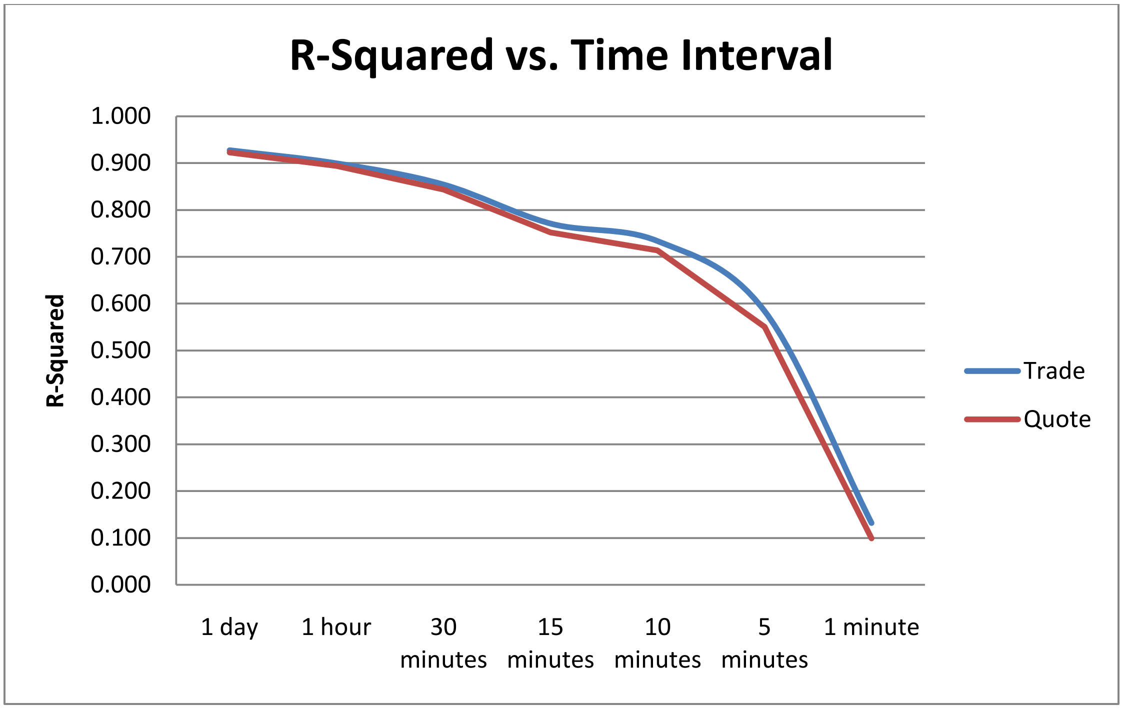

Figure 1: The figure above plots the R^2 values obtained for regressions using different time intervals.

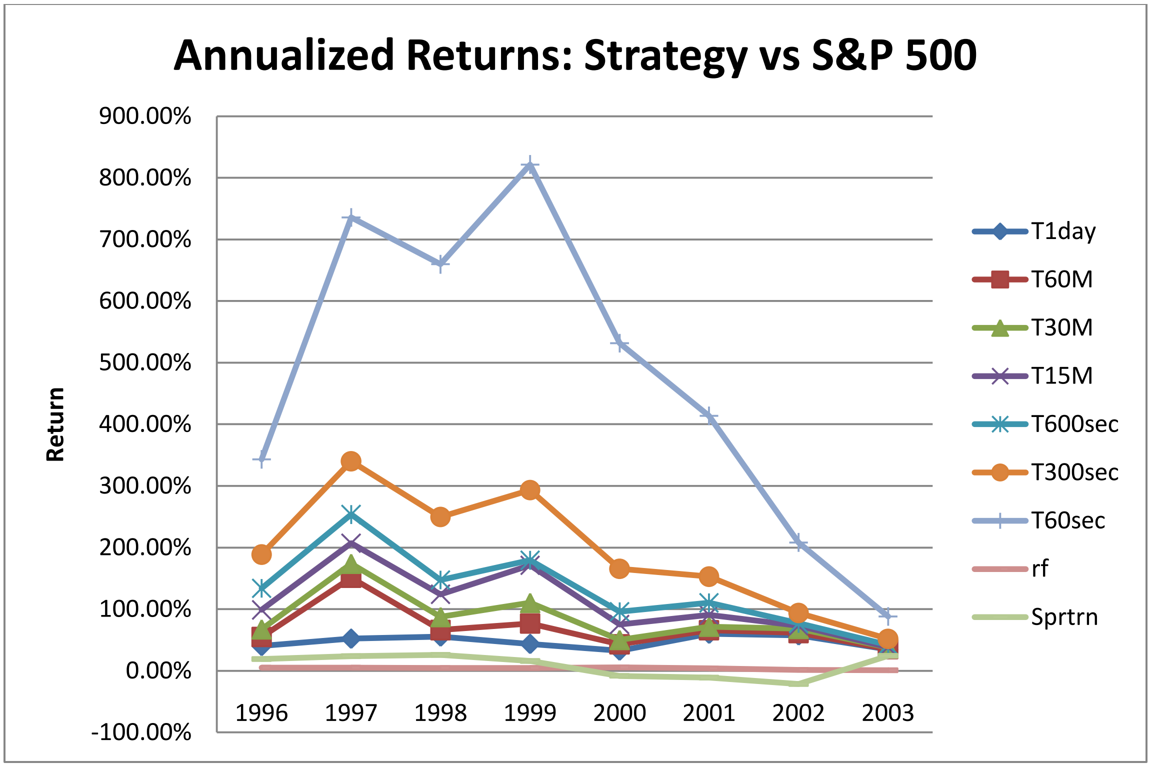

Figure 2 above demonstrates the performance of our simple trading strategy vs the performance of the S&P 500 between 1996 and 2003. T1day represents daily trades. T60M represents trades every hour, T30M represents trades every 30 minutes, T15M represents trades every 15 minutes, T600sec represents trades every 10 minutes, T300sec represents trades every 5 minutes, T60 sec represents trades every 60 seconds, rf is the short term risk free rate. Sprtrn is the return on the S&P 500.

Table I
Descriptive Statistics

This table presents descriptive statistics for all data between 1996 and 2003. The table provides the average share price and the shares oustanding for S&P 500 components and SPDR. Trades presents the average number of trades per year.

S&P 500	mean	median	min	max	st dev	25%	75%
Price	43.00	38.81	0.36	659.00	25.57	26.13	54.88
Shares Out. (000's)	409,976	187,159	1,889	11,144,681	762,441	102,823	387,467
Trades	275,682,940	285,333,132	75,619,680	464,679,437	164,585,178	107,205,133	461,169,796

Panel B

SPDR	mean	median	min	max	st dev	25%	75%
Price	106.91	107.35	59.97	153.56	24.27	89.78	127.51
Shares Out. (000's)	131,204	27,109	9,200	465,295	144,471	27,109	149,422
Trades	1,676,999	415,837	41,695	5,963,984	2,346,309	191,853	3,940,621

Table II:
Correlation Estimates

This table provides Pearson correlation statistics as well as Spearman rank correlations for the variables used to explain tracking error. The right triangular (above the diagonal) matrix presents Pearson's correlation coefficients between various explanatory variables used in our regression analysis (Eq. 5). The left triangular (below the diagonal) matrix presents Spearman's Rank correlation coefficients. PC is the put call ratio from the CBOE, lnVIX is the natural log of the closing VIX price, PVF is the Sadka Permanent Variable Factor, AD is the maximum of the advancers or decliners divided by the total number of issues traded, SI is a sentiment variable from Baker and Wurgler (2006). D96 is a dummy for 1996, D97 is a dummy for 1997, D98 is a dummy for 1998, D99 is a dummy for 1999, D00 is a dummy for 2000, D01 is a dummy for 2001, D02 is a dummy for 2002.

	lnVIX	PVF	PC	AD	SI
lnVIX		-.378**	0.170	.479**	0.077
PVF	-.316**		-0.003	-0.0628	-0.176
PC	0.086	0.043		.400**	-.268**
AD	.374**	-0.026	.455**		-0.134
SI	0.0696	-0.175	-.292**	-.210*	

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table III: Comovement of SPDR and S&P 500

R^2 values obtained from regressing SPDR returns on S&P 500 returns for all time periods examined (using trades and quote midpoints).

$$r_{SPDR,t} = \alpha + \beta \times r_{index,t} + \varepsilon_t$$

The table below presents R^2 values for all time periods studied (1, 5, 10, 30, 60 minutes and 1 day) split by year, by month and by day of the week for SPDR trades and quote midpoints. $r_{SPDR,t}$ is the continuously compounded return on SPDR at time t and $r_{index,t}$ is the continuously compounded return on the index at time t.

Panel A: SPDR trades R^2 values by year

Year	1 min	5 min	10 min	15 min	30 min	hour	Day
1996	0.022	0.264	0.507	0.676	0.770	0.840	0.937
1997	0.037	0.428	0.606	0.676	0.728	0.774	0.782
1998	0.064	0.586	0.753	0.813	0.892	0.944	0.972
1999	0.089	0.488	0.728	0.609	0.853	0.883	0.906
2000	0.103	0.659	0.812	0.855	0.908	0.940	0.966
2001	0.226	0.646	0.698	0.766	0.852	0.928	0.956
2002	0.508	0.867	0.913	0.936	0.949	0.977	0.988
2003	0.423	0.612	0.733	0.752	0.764	0.789	0.831
Average	0.184	0.569	0.719	0.760	0.839	0.884	0.917

Panel B: SPDR quote midpoints R^2 values by year

Year	1 min	5 min	10 min	15 min	30 min	hour	Day
1996	0.035	0.381	0.643	0.796	0.861	0.902	0.952
1997	0.040	0.499	0.666	0.723	0.772	0.798	0.779
1998	0.051	0.606	0.789	0.840	0.914	0.954	0.976
1999	0.032	0.315	0.540	0.497	0.744	0.841	0.902
2000	0.076	0.587	0.764	0.824	0.890	0.928	0.965
2001	0.162	0.597	0.667	0.731	0.834	0.918	0.932
2002	0.393	0.827	0.889	0.909	0.930	0.965	0.986
2003	0.298	0.545	0.675	0.712	0.723	0.757	0.814
Average	0.136	0.545	0.704	0.754	0.833	0.883	0.913

Panel C: SPDR trades R^2 values by month

Month	1 min	5 min	10 min	15 min	30 min	hour	Day
January	0.107	0.540	0.680	0.771	0.824	0.861	0.914
February	0.094	0.528	0.739	0.836	0.894	0.932	0.962
March	0.135	0.646	0.781	0.875	0.921	0.951	0.983
April	0.133	0.622	0.742	0.786	0.829	0.903	0.952
May	0.085	0.402	0.688	0.521	0.878	0.931	0.980
June	0.072	0.562	0.734	0.790	0.869	0.930	0.954
July	0.194	0.634	0.773	0.760	0.838	0.876	0.919
August	0.170	0.678	0.799	0.854	0.895	0.940	0.976
September	0.134	0.507	0.618	0.715	0.788	0.821	0.855
October	0.216	0.646	0.754	0.805	0.845	0.860	0.826
November	0.189	0.664	0.776	0.800	0.844	0.946	0.973
December	0.103	0.629	0.796	0.815	0.898	0.933	0.954
Average	0.136	0.588	0.740	0.777	0.860	0.907	0.937

Panel D: SPDR quote midpoints R^2 values by month

Month	1 min	5 min	10 min	15 min	30 min	hour	Day
January	0.065	0.476	0.628	0.737	0.803	0.857	0.910
February	0.078	0.537	0.770	0.852	0.905	0.936	0.969
March	0.112	0.621	0.785	0.861	0.912	0.953	0.983
April	0.137	0.684	0.782	0.820	0.866	0.923	0.951
May	0.083	0.404	0.715	0.528	0.899	0.941	0.982
June	0.064	0.578	0.747	0.794	0.883	0.936	0.951
July	0.145	0.595	0.740	0.733	0.812	0.872	0.899
August	0.101	0.591	0.746	0.820	0.888	0.930	0.967
September	0.072	0.423	0.543	0.653	0.724	0.774	0.839
October	0.160	0.596	0.741	0.776	0.831	0.846	0.819
November	0.135	0.609	0.755	0.770	0.843	0.944	0.973
December	0.089	0.618	0.788	0.818	0.892	0.939	0.964
Average	0.103	0.561	0.728	0.763	0.855	0.904	0.934

Panel E: SPDR trades R^2 values by day of the week

Month	1 min	5 min	10 min	15 min	30 min	hour	Day
Monday	0.139	0.560	0.706	0.786	0.867	0.926	0.957
Tuesday	0.131	0.534	0.738	0.652	0.818	0.858	0.888
Wednesday	0.137	0.618	0.729	0.815	0.872	0.902	0.903
Thursday	0.129	0.594	0.752	0.813	0.860	0.910	0.951
Friday	0.126	0.615	0.742	0.804	0.856	0.907	0.941
Average	0.133	0.584	0.734	0.774	0.855	0.901	0.928

Panel F: SPDR quote midpoints R^2 values by day of the week

Month	1 min	5 min	10 min	15 min	30 min	hour	Day
Monday	0.102	0.560	0.752	0.786	0.867	0.927	0.955
Tuesday	0.104	0.534	0.729	0.652	0.818	0.852	0.883
Wednesday	0.103	0.618	0.705	0.815	0.872	0.906	0.899
Thursday	0.098	0.594	0.738	0.813	0.860	0.901	0.938
Friday	0.093	0.615	0.706	0.804	0.856	0.886	0.938
Average	0.100	0.584	0.726	0.774	0.855	0.895	0.923

Panel G: SPDR trades and quotes R^2 values for entire period

	1 min	5 min	10 min	15 min	30 min	hour	Day
Trade	0.132	0.584	0.734	0.771	0.854	0.900	0.928
Quote	0.099	0.550	0.714	0.752	0.843	0.894	0.922

Table IV
Explaining variability in monthly tracking error

$$TE_i = \sum_t \quad {}_i \Big|$$

$$TE_i = \alpha + \beta_1 \times \ln VIX_i + \beta_2 \times PVF_i + \beta_3 \times PC_i + \beta_4 \times AD_i + \beta_5 \times SI_i + \beta_{6L\ 12} \times D_{96L\ 02} + \nu_i$$

PC is the put call ratio from the CBOE, lnVIX is the natural log of the closing VIX price, PVF is the Sadka Permanent Variable Factor, AD is the maximum of the advancers or decliners divided by the total number of issues traded, SI is a sentiment variable from Baker and Wurgler (2006). D96 is a dummy for 1996, D97 is a dummy for 1997, D98 is a dummy for 1998, D99 is a dummy for 1999, D00 is a dummy for 2000, D01 is a dummy for 2001, D02 is a dummy for 2002.

Variable	1	2	3	4	5	6	7
	0.511					0.403	0.394
lnVIX	(7.81)					(6.10)	(4.93)
		-0.341				-0.211	-0.202
PVF		(-5.9)				(-4.0)	(-3.6)
			0.192				-0.109
PC			(1.98)				(-1.2)
				0.468			0.049
AD				(3.91)			(0.43)
					-0.023		
SI					(-0.1)		
	0.406	0.174	0.237	0.445	0.215	0.345	0.244
D96	(5.71)	(2.38)	(2.75)	(4.12)	(1.81)	(5.11)	(2.76)
	0.584	0.552	0.681	0.968	0.629	0.551	0.562
D97	(8.79)	(7.49)	(7.50)	(8.09)	(4.77)	(8.90)	(5.01)
	0.607	0.627	0.778	0.950	0.725	0.575	0.573
D98	(8.95)	(8.44)	(8.61)	(9.51)	(6.23)	(9.13)	(5.81)
	0.578	0.625	0.799	0.952	0.692	0.564	0.534
D99	(8.53)	(8.51)	(7.77)	(9.06)	(6.26)	(9.02)	(4.65)
	0.710	0.673	0.907	0.944	0.794	0.660	0.616
D00	(10.6)	(8.98)	(8.46)	(10.3)	(5.02)	(10.4)	(6.00)
	0.261	0.342	0.438	0.475	0.426	0.254	0.240
D01	(3.79)	(4.65)	(5.04)	(5.60)	(1.95)	(4.00)	(3.09)
	0.122	0.232	0.274	0.308	0.296	0.122	0.128
D02	(1.74)	(3.15)	(3.22)	(3.70)	(2.67)	(1.90)	(1.81)
Adjusted R^2	0.761	0.712	0.612	0.649	0.594	0.797	0.787

Table V:
SPDR Ratios

$$IS_t = \frac{Index_t}{SPDR_t}$$

This table reports the ratio of S&P 500 Index level/ SPDR price. By construction the ratio should be equal to 10, however it is allowed to deviate from 10. The mean median and standard deviation of the ratios for each time period examined are reported. 1 Min represent 1 minute intervals, 5 min represents 5 minute intervals etc.

Year		minute			5 min			10 min			15 min	
	Mean	Median	St.dev	Mean	Median	St.dev	Mean	Median	St.dev	Mean	Median	St.dev
1996	9.968	9.973	0.035	9.968	9.973	0.035	9.968	9.973	0.035	9.968	9.973	0.035
1997	9.967	9.975	0.043	9.967	9.975	0.043	9.968	9.975	0.067	9.967	9.975	0.043
1998	9.971	9.976	0.034	9.970	9.976	0.035	9.970	9.976	0.035	9.970	9.976	0.035
1999	9.968	9.973	0.045	9.968	9.972	0.045	9.968	9.972	0.048	9.968	9.973	0.048
2000	9.973	9.977	0.063	9.973	9.977	0.063	9.973	9.978	0.063	9.973	9.977	0.063
2001	9.974	9.974	0.032	9.974	9.974	0.033	9.974	9.974	0.034	9.974	9.974	0.035
2002	9.967	9.969	0.078	9.966	9.969	0.077	9.967	9.969	0.078	9.966	9.969	0.076
2003	9.958	9.957	0.032	9.958	9.957	0.032	9.922	9.957	0.054	9.957	9.957	0.032

SPDR Ratios continued

Year		30 min			hour			Day	
	Mean	Median	St.dev	Mean	Median	St.dev	Mean	Median	St.dev
1996	9.968	9.973	0.035	9.968	9.973	0.036	9.970	9.972	0.043
1997	9.967	9.975	0.044	9.968	9.975	0.045	9.970	9.973	0.051
1998	9.970	9.976	0.036	9.971	9.975	0.037	9.973	9.976	0.045
1999	9.968	9.972	0.054	9.968	9.972	0.051	9.968	9.972	0.047
2000	9.973	9.978	0.064	9.973	9.978	0.064	9.973	9.977	0.067
2001	9.974	9.974	0.037	9.974	9.974	0.041	9.976	9.975	0.059
2002	9.967	9.969	0.080	9.967	9.969	0.081	9.971	9.970	0.106
2003	9.957	9.957	0.033	9.957	9.956	0.035	9.956	9.954	0.043

Table VI:
S&P 500 returns vs trading strategy with SPDR trades

The table below presents returns for the simple trading strategy using SPDR trades as well as the S&P 500 and the risk free rate. T1day represents daily trading, T60M represents trading every hour, T30M represents trading every 30 minutes and so on. Sptrn is the total return on the S&P 500, rf is the risk free rate.

Year	T1day	T60M	T30M	T15M	T600sec	T300sec	T60sec	rf	Sprtrn
1996	40.27%	54.76%	67.45%	98.84%	133.87%	188.67%	343.04%	5.07%	19.16%
1997	52.15%	151.14%	173.97%	207.17%	253.59%	339.86%	735.73%	5.08%	23.56%
1998	55.62%	66.11%	87.35%	124.20%	147.39%	249.60%	659.68%	4.72%	25.71%
1999	43.66%	76.78%	110.42%	171.22%	179.30%	293.17%	821.61%	4.53%	15.80%
2000	32.77%	42.87%	50.20%	74.98%	95.99%	165.41%	531.67%	5.68%	-8.23%
2001	60.21%	65.78%	71.75%	90.55%	110.27%	152.81%	413.85%	3.73%	-11.09%
2002	57.60%	61.45%	68.14%	73.06%	77.03%	93.71%	208.31%	1.56%	-21.57%
2003	34.66%	35.05%	37.94%	38.16%	41.73%	51.90%	88.22%	0.94%	24.30%
Average	47.12%	69.24%	83.40%	109.77%	129.90%	191.89%	475.26%	3.91%	8.45%

Table VII:
Sharpe ratios calculated from simple trading strategy using trade data

The Sharpe ratios below are calculated by taking excess return of the trading strategy divided by the standard deviation of the returns for each time interval analyzed using SPDR trade data.

Year	Return on S&P 500	minute	5 min	10 min	15 min	30 min	hour	Day
1996	0.075	0.478	0.512	0.517	0.517	0.474	0.405	0.337
1997	0.067	0.452	0.245	0.186	0.152	0.127	0.110	0.447
1998	0.065	0.615	0.630	0.617	0.554	0.564	0.505	0.467
1999	0.040	0.142	0.135	0.136	0.128	0.124	0.167	0.454
2000	-0.039	0.559	0.541	0.560	0.498	0.505	0.466	0.384
2001	-0.044	0.514	0.543	0.472	0.518	0.484	0.456	0.425
2002	-0.057	0.422	0.525	0.591	0.611	0.604	0.603	0.571
2003	0.094	0.368	0.308	0.555	0.561	0.573	0.608	0.615
Average	0.025	0.444	0.430	0.454	0.442	0.432	0.415	0.463

Table VIII

S&P 500 returns vs trading strategy with SPDR quote midpoints

The table below presents returns for the simple trading strategy using SPDR quote midpoints as well as the S&P 500 and the risk free rate. Q1day represents daily trading, Q60M represents trading every hour, Q30M represents trading every 30 minutes and so on. Sptrn is the total return on the S&P 500, rf is the risk free rate.

Year	Q1day	Q60M	Q30M	Q15M	Q600sec	Q300sec	Q60sec	rf	Sprtrn
1996	48.49%	53.34%	61.58%	75.26%	92.81%	108.87%	173.11%	5.07%	19.16%
1997	48.20%	139.67%	152.76%	174.30%	195.28%	244.93%	435.75%	5.08%	23.56%
1998	47.18%	54.74%	66.89%	90.46%	108.79%	163.78%	419.05%	4.72%	25.71%
1999	35.57%	80.21%	133.95%	218.18%	277.53%	456.46%	1655.41%	4.53%	15.80%
2000	29.08%	41.36%	55.80%	86.17%	116.35%	211.84%	818.01%	5.68%	-8.23%
2001	62.33%	67.92%	76.54%	95.94%	124.32%	171.19%	529.34%	3.73%	-11.09%
2002	60.55%	65.16%	69.66%	78.43%	86.37%	114.41%	277.51%	1.56%	-21.57%
2003	34.56%	33.75%	39.47%	41.27%	41.48%	58.04%	113.43%	0.94%	24.30%
Average	45.75%	67.02%	82.08%	107.50%	130.37%	191.19%	552.70%	3.91%	8.45%

Table IX:
Sharpe ratios calculated from simple trading strategy using quote midpoints

The Sharpe ratios below are calculated by taking excess return of the trading strategy divided by the standard deviation of the returns for each time interval analyzed using SPDR quote midpoints.

Year	Return on S&P 500	minute	5 min	10 min	15 min	30 min	hour	Day
1996	0.075	0.471	0.460	0.435	0.436	0.419	0.383	0.354
1997	0.067	0.302	0.178	0.143	0.127	0.111	0.101	0.427
1998	0.065	0.508	0.559	0.506	0.530	0.459	0.410	0.376
1999	0.040	0.142	0.207	0.207	0.163	0.150	0.174	0.349
2000	-0.039	0.443	0.455	0.449	0.473	0.465	0.419	0.329
2001	-0.044	0.548	0.582	0.513	0.542	0.473	0.416	0.448
2002	-0.057	0.448	0.507	0.584	0.637	0.610	0.623	0.608
2003	0.094	0.262	0.302	0.521	0.545	0.551	0.589	0.600
Average	0.025	0.390	0.406	0.420	0.432	0.405	0.389	0.436