LARGE TRADERS IN FINANCIAL MARKETS

by

Stephen Fagan

A Thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
in the Department
of
Economics

SIMON FRASER UNIVERSITY
Summer 2011

APPROVAL

Name: Stephen Fagan

Degree: Doctor of Philosophy

Title of Thesis: Large Traders in Financial Markets

Examining Committee: Dr. Christoph Luelfesmann
Chair

Dr. Ramazan Gencay, Department of Economics
Senior Supervisor

Dr. Robert Jones, Department of Economics
Supervisor

Dr. Robert Grauer, Beedie School of Business
Supervisor

Dr. Peter Klein, Beedie School of Business
SFU Examiner

Dr. Blake LeBaron, External Examiner,
Brandeis International Business School
Brandeis University

Date Approved: 23 June 2011

Declaration of
Partial Copyright Licence

Abstract

Large traders in financial markets care a lot about the supply of liquidity - factors that allow them to trade quickly without incurring a significant cost - since the size of their trading can cause a large market impact if they demand more liquidity than is currently available.

The first study examines an important and recurring cause of liquidity shocks in futures markets - the accumulation of extreme and opposing positions by hedgers and speculators. These two classes of traders are found to differ in the impact their trades have on market prices and their response to market-relevant news and past returns. As their positions diverge sufficiently, these differences affect the number and heterogeneity of available counterparties, and can cause increased volatility, high bid-ask spreads, and return regularities.

The second study examines the limits of anonymity available to large block traders. In addition to the trade-off between price risk and price impact that block-traders face, this paper introduces an additional, strategic, consideration: by trading with large volume for several periods, block traders risk losing their anonymity and attracting opportunistic strategic traders who want to 'ride' the block traders' market impact. In simulated markets, block traders lose their anonymity under several parameter configurations and strategic traders are able to profitably exploit the knowledge of their presence.

The third study examines the recent 'flash crash' episodes that rose to the public's attention on May 6th, 2010, when within only a few minutes, the Dow Jones Industrial Average experienced its second largest point swing and the biggest one-day point decline in its history. This study examines whether these flash-type events which induce large intraday volatility are becoming more common since the big flash crash of May 6th. Over a 14-month sample period around May 6th, this paper finds that there have been more flash-type events with shorter inter-flash durations since May 6th, however, survival analysis of these events does not establish the significance of these different frequencies.

Acknowledgments

It is a pleasure to thank the many people who made this thesis possible. First, I have had the great fortune and pleasure of having Ramo Gencay as my senior supervisor. Ramo's research into high-frequency finance was the perfect complement to my interest in trading and financial markets. In addition to the theoretical and technical knowledge that he shared with me, I also learned from his example about the importance of integrity and work ethic in producing good research.

I also benefited from the excellent supervision of Robbie Jones and Rob Grauer. Robbie's interest in computational finance and his deep knowledge of the financial markets was a constant source of motivation for me, and I always enjoyed our many interesting conversations. Rob Grauer showed me how questioning the status quo can lead to interesting new research questions. He also made me feel welcome at the business school, and was always willing to help when I had questions or needed advice.

My friends were also invaluable to me for their encouragement, useful suggestions, and sometimes for some enjoyable distractions. In particular, Daniele Signori, Sunny Yang, Soroush Kasami, and Pierre Nguimkeu deserve my thanks. Finally, my partner Carolyn Geh was there for me every day, and her support and unflagging belief in me made the tough days bearable and the good days great. Thank you.

Contents

List of Tables

List of Figures

Chapter 1

Large Traders and Liquidity in Futures Markets

1.1 Introduction

While market liquidity is a multifaceted concept, it generally refers to the ability to easily buy or sell a security without causing a significant change in the market price. An essential requirement of liquidity is the ample availability of counterparties who are willing to sell when others want to buy and who are willing to buy when others want to sell. Finance theory often assumes perfect market liquidity where market participants can trade any amount of a security without affecting the price. This implicitly requires the unlimited presence of counterparties. While this assumption is clearly false, it is usually a reasonable simplification in large and active markets. There are, however, occasions where, even in the largest and most active markets, liquidity dries up because of the scarcity of counterparties. Essentially, these are times when, for the current market price, everyone who wants to be long is already long, or everyone who wants to be short is already short. When this occurs, someone wanting to trade will be unable to do so unless the market price adjusts to induce someone to trade. While the existence of liquidity shocks has been firmly established, understanding how such occurrences develop is still an important and open question.

This paper presents an empirical microstructural analysis of these liquidity-induced price adjustments in futures markets. The market chosen for this study is that for the New York Mercantile Exchange's light sweet crude oil futures contracts. This is the most liquid market

for trading crude oil which is itself the world's most actively traded physical commodity. Because of its liquidity, this market is taken to be very efficient, with the market prices reflecting the value of crude oil. As such, the futures price is used as an international pricing benchmark and an indicator of world energy prices. Despite this overall liquidity, this paper discovers that there are fairly regular occasions where price dynamics, as investigated through returns, realized volatilities and bid-ask spreads, are dominated by the low-liquidity effects resulting from the demographics of available counterparties. These effects are significant and can persist for up to 15 weeks. This paper identifies the underlying mechanism that causes these effects, and isolates its three main components: (1) significant differences between the dominant classes of traders (i.e. hedgers and speculators), (2) the bounded and mean-reverting nature of trader positions, and (3) the various conditions that cause speculators to enter and exit the market.

Futures markets are designed with many unique features that differentiate them from equity and fixed income markets. One of the most prominent differences is the extensive use of futures markets in risk reduction strategies for those involved in businesses related to the underlying commodity. This feature results in a partition of market participants into two classes of traders: hedgers and speculators. There are large and small traders of both classes, however market activity is dominated large speculators and hedgers. For this study, a trader is considered large if they hold more than 350 futures contracts - the reporting threshold set by the Commodity Futures Trading Commission (CFTC). In the crude oil futures market, large hedgers and speculators hold, on average, 84% of the outstanding futures contracts. While futures markets are anonymous in the sense that information on who is involved in any given trade is not publicly available, the CFTC provides a weekly snapshot of the holdings of these classes of traders.

At any given time, the number of traders in the crude oil market, as well as the aggregated positions held by each class of trader, is bounded. These bounds arise from the number of market participants as well as the constraints faced by these participants. Hedgers participate in futures markets to reduce the price risk that they face in their business activities, and so the number of contracts they hold is determined by the size of their business interests. To use more futures contracts than this bound would be to over-hedge and would

increase the price risk they face[1]. Speculators, on the other hand, participate in futures markets to earn profit by accepting additional risk. They too face constraints on their position sizes, but these constraints arise from factors such as available capital, trading strategies, as well as leverage and risk limits. While these differences between hedgers and speculators are well-documented, it is not well understood what, if any, observable market regularities result from these differences. If there are no clear differences between the market impact of these classes of traders, or if there is too much heterogeneity within the classes, then examining opposing positions of hedgers and speculators would not be a useful exercise. This study rejects the hypothesis of no systematic differences between these classes, and establishes reliable differences both in the impact of their trades on market prices and in their response to past market action and market-relevant news.

By examining trade between hedgers and speculators using the net long measure of their positions, large speculators are found to increase prices with their purchases and to engage in trend-following behavior. Large hedgers, on the other hand, decrease prices through their purchases and exhibit contrarian-type behavior. Further, using textual analysis techniques, speculators and hedgers are found to respond differently to market-specific news. These regularities are consistent with the notion that large speculators are informed traders who bring new information to the market through orders that are placed more aggressively than hedgers.

Under liquid market conditions, there are many hedgers and speculators who are not trading near their position limits, and so when a trader wishes to trade, it is highly likely that a counterparty will be available to trade with. Further, the available counterparties on either side of a trade include a mix of both speculators and hedgers. Thus, the systematic differences between the classes of traders effectively gets washed out of the price dynamics. However, when the net long positions of these trader classes significantly diverge, the likelihood of one class of trader dominating the long side of each trade, as the other class dominates the short side, increases. Under such conditions, the differences between trader classes are no longer washed out and begin to significantly impact price movements.

[1]There are many examples of hedgers taking speculative risks. For example, Metallgesellschaft AG, formerly one of Germany's largest industrial conglomerates, lost over 1.4 billion dollars in 1993 through its trading activities in derivatives. [28] and [61] argue that the 1:1 hedge strategy that Metallgesellschaft used was significantly oversized given its underlying oil exposure and increased its oil price risk instead of reduced it. Despite such examples, the distinction between hedgers and speculators in the oil market is empirically significant as is shown in Section II. This distinction is not always apparent, as is seen in Section IV for the Eurodollar market.

Returns, volatilities, and spreads all demonstrate systematic regularities around these times of extreme and opposing positions held by large hedgers and speculators.

Of course, the bounded nature of trader positions prohibits unlimited divergence of trader positions. Consequently, as divergence ends and mean reversion begins, the established concurrent relationship between position accumulation and returns generates sharp changes in price dynamics. The trend following and aggressive order placement of speculators generates persistence in these dynamics for several weeks before and after the divergence of positions reaches its local maximum. For example, when speculators have acquired extremely short net positions and hedgers have acquired extremely long net positions, the price tends to follow a v-shaped reversal pattern, the bid-ask spreads are higher than normal, and the average realized volatility at its 79th percentile. Such liquidity effects are found to be largely determined by the entry and exit of large speculators from the market. In particular, speculators facing increasing losses exit the market using orders that are significantly more aggressive than typical orders, thereby pushing up return volatility. The notion of directional realized volatility is introduced to capture the difference in contributions to volatility from periods of positive and negative returns. Disaggregating the realized volatility in this manner permits the identification of a significant asymmetry in the cause of high volatility before and after these liquidity shocks.

Within the literature, aside from some theoretical studies (e.g. [72]), the study of liquidity has focused mainly on equity [45, 21] and fixed income markets [22, 50, 15]. Trader positions have been related to returns [66] and volatility [73] in futures markets, however, this paper extends the methodology of these studies in three important ways. First, this paper studies futures market microstructure, showing that unique features of these markets can have substantial impacts on market liquidity. Second, the notion that trader positions are bounded is captured by transforming the net long measure into an index that expresses the current position in relation to recent maximum and minimum values. Finally, rather than examining the overall influence of trader positions, this paper focuses on what happens when the position index approaches its bounds.

The paper is organized as follows. Section I discusses the market and CFTC data. Section II examines the empirical regularities that differentiate hedgers from speculators. Section III analyzes ex post market dynamics around extreme holdings of these two classes of traders. Results of robustness checks are described in Section IV and a brief summary concludes the paper in Section V.

1.2 The Market and CFTC Data

The market under study in this paper is the New York Mercantile Exchange's Light Sweet Crude Oil futures market. This futures contract is the most liquid instrument for trading crude oil, which is the world's most actively traded physical commodity. Approximately 500,000 contracts are traded per day in this market, with 1.5 million contracts in open interest. This level of daily trading volume represents almost seven times the daily world production of crude oil[2]. The light sweet variety of crude oil is popular among refineries because of its low sulfur content and its high yield of gasoline, diesel, and other petroleum products. This market is chosen because of its high overall liquidity, as well as the fact that the speculator-hedger distinction is not as fuzzy as in other markets such as financial futures markets where contracts can be used as a speculative and hedging instrument by the same fund manager.

Data from three sources are integrated for the analysis: futures market trade data[3], textual news articles on the crude oil market[4], and Commitments of Traders Report (CoT) data from the Commodity Futures Trading Commission (CFTC). The following sections describe these datasets and the construction of the variables of interest.

1.2.1 Futures Market Data

Futures price data for the NYMEX light sweet crude oil market is collected for the same period for which weekly CoT data is available, namely September 30, 1992 until February 28, 2006 (698 weeks). The data is sampled at the tick level (5,487,792 observations) and aggregated to the scales required for variable construction. Since there is trade in several contracts, each with different expiration dates and market prices, a sensible method of constructing a single price series over this period is needed. Since the front contract - the contract with the closest expiration date - usually has the highest trade volume and open interest, the constructed price series is composed almost entirely of prices from the front contract. As the front contract approaches expiration, most traders begin to close out positions in the front contract and enter position in the first-back-month contract. On

[2]Data on trading activity is from the New York Mercantile Exchange, and data on oil production is from the Energy Information Administration.

[3]The data provider for the crude oil futures market tick data is TickData, www.tickdata.com.

[4]The data provider for the news articles is Dow Jones Factiva, www.factiva.com.

the first day that the daily number of price ticks, a proxy for the volume of trade, of the first-back-month exceeds that of the front contract, the constructed price series uses the prices from the first-back-month contract from that point onwards. The standard backward-adjustment process corrects the series for price jumps that are only due to the switching of contracts.

The returns analyzed in this paper are weekly returns, calculated from prices on the close of each Tuesday - the 'as of' date of the CoT report which is released three days later on Friday. In addition to returns, realized volatility [24, 6] is analyzed as a proxy for the true volatility of this return series. Realized volatility is sampled on a weekly basis and constructed from 5-minute intraday return data. The choice of 5-minute data sampling is consistent with the literature on optimal sampling in the presence of microstructure effects (e.g. [9]). Let there be h five-minute intervals during which the market is open in one week. The current day trading hours for NYMEX crude oil futures is from 10:00 until 14:30 EST. Thus, one trading day has 54 five-minute intervals, and one trading week has $h = 270$ five-minute intervals. Then realized volatility is calculated as

$$RV_t = \sum_{i=1}^{h} \left(r^{(h)}_{t-1+i(\frac{1}{h})} \right)^2 \tag{1.1}$$

where $r^{(h)}_t$ is the five-minute return at time t. Figure 1 presents plots of the price series, the weekly returns, and the realized volatility.

[FIGURE 1 ABOUT HERE]

1.2.2 Directional Realized Volatility

Realized volatility is a measure of the size of price movements, but it says nothing of the direction of those movements. Yet, the ups and downs of prices might not equally contribute to the return volatility in any given period. During a swiftly rising market, one might expect that the numerous buyers would make selling easy, but buying may be difficult in the sense that it would induce significant slippage. Similarly, during a market crash, selling would be difficult, but buying would be easy. To investigate the notion that liquidity shocks may be asymmetric in this sense, the concept of directional realized volatility is introduced.

Directional realized volatility disaggregates the realized volatility into a component that measures the contribution of positive returns to the realized volatility, denoted DRV_t^P, and another component that measures the contribution of negative returns, denoted DRV_t^N. The two components, which are additive variance-preserving transformations of realized volatility, are defined as follows:

$$DRV_t^P = \sum_{i=1}^{h} \left[I\left(r_{t-1+i(1/h)}^{(h)} \geq 0 \right) \cdot r_{t-1+i(1/h)}^{(h)} \right]^2 \tag{1.2}$$

$$DRV_t^N = \sum_{i=1}^{h} \left[I\left(r_{t-1+i(1/h)}^{(h)} < 0 \right) \cdot r_{t-1+i(1/h)}^{(h)} \right]^2 \tag{1.3}$$

where $r_{t-1+i(1/h)}^{(h)}$ is the five-minute return as defined above and $I(\cdot)$ is an indicator function taking the value 1 if the condition in the brackets is satisfied, and 0 otherwise. The positive (negative) directional realized volatility sums the squared five-minute returns that are positive (negative). The directional realized volatility is very similar to the realized semivariance [10], which is a recently proposed measure of donwside market risk.

1.2.3 Estimating the Bid-Ask Spread

While returns and realized volatility are informative when investigating liquidity dynamics, there is a growing literature that has developed many liquidity measures to quantify specific liquidity related factors. One of the most popular liquidity measures is the bid-ask spread [65] which is often used as a liquidity benchmark for evaluating other liquidity measures [35]. As liquidity dries up, trade can only be induced by offering better prices to potential buyers and sellers resulting in a larger bid-ask spread. Conversely, when there are many willing buyers and sellers, the bid-ask spread will be smaller.

A spread proxy can be constructed from transaction data using the tick test [59] to partition the transactions into those likely to have originated from market buy orders and those likely to have originated from market sell orders. A transaction is identified as one likely to have originated as a market buy (sell) order if the previous transaction price was below (above) the current transaction price. If the previous transaction price was the same as the current transaction price, then the previous distinct transaction price is used instead. A critical assumption here is that transactions occur only at the bid or ask prices, thus at

each transaction price is either an observation of the bid or of the ask price. If the bid (ask) is known at time t, to get a proxy of the spread at that time, we must estimate the ask (bid) price at time t. This estimate is taken to be the most recently identified ask (bid) price constrained to be at least one tick above the observed bid price. Figure 2 shows a sample of the bid and ask estimates given the transaction prices.

[FIGURE 2 ABOUT HERE]

This procedure gives an estimated series of bid-ask spreads at the same high frequency as transactions. As pointed out by Lee and Ready (1991), the tick test becomes less precise when the time between transactions increases. Given the high level of activity in the crude oil futures markets, this issue is much less significant then in other less active markets.

With these estimates of the bid-ask spreads, the percentage spread for each transaction is defined to be the ratio of the estimated bid-ask spread to the midpoint between the bid and the ask. These percentage spreads are averaged over each day, t, giving,

$$S_t = \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{Spread_{t,i}}{Midpoint_{t,i}} \tag{1.4}$$

where N_t is the number of transactions on day t, $Spread_{t,i}$ is the ith estimated bid-ask spread, and $Midpoint_{t,i}$ is the midpoint between the ith estimated bid and ask.

1.2.4 Textual News Data

To examine potential differences between hedgers and speculators, their response to news articles related to the crude oil markets is analyzed. Since the electronic news database is relatively sparse before 2001, a more recent sample period, between July 25, 2001 and April 28, 2009, is used for this analysis. Approximately $58,000$ news articles related to the crude oil markets were extracted from the Dow Jones Factiva database. These articles were all from the publications used by the Energy Information Administration (EIA) to form their Annual Oil Market Chronology[5] which lists all of the significant world events that have impacted the crude oil markets. The specific uses of this data are detailed in Section II.

[5] http://www.eia.doe.gov/cabs/AOMC/Sources.html

1.2.5 CFTC Trader Data

The Commodity Futures Trading Commission (CFTC) has been releasing weekly Commitments of Traders (CoT) Reports[6] since September 30, 1992, and at various lower frequencies since 1924. The earliest predecessor of the CoT reports were prepared by the U.S. Department of Agriculture's Grain Futures Administration and were released yearly. Currently, these reports provide a breakdown of the open interest in the American futures markets as of each Tuesday. Open interest refers to the total number of futures contracts that have been entered into and not yet exited through a transaction or delivery. One of the main reasons for collecting this data is to detect and deter attempts at market manipulation (such as the cornering of markets) by large traders. More generally, the CoT reporting system allows the CFTC staff to identify large positions that could pose a threat to orderly trading.

In each CoT report, total open interest is broken down according to the type of trader and by the type of positions (long or short) that each type of trader holds. There are three types of traders identified by the CFTC: reporting commercial, reporting non-commercial, and non-reporting. A trader is classified as a reporting trader if they hold positions in excess of the CFTC-determined threshold of 350 contracts. This threshold varies across time and markets in an effort to capture between 70 and 90 percent of the open interest in the market. While the reporting threshold is quite high, the average reporting traders hold much larger positions. For example, in July of 2007, the average reporting trader held over 9000 contracts. A commercial trader is one who self-identifies as being engaged in business activities hedged by use of the futures and options markets. The non-reportable positions are all positions that are held by traders whose total position size is below the CFTC's threshold. In what follows, reporting commercial traders are considered large hedgers, reporting non-commercial traders are considered large speculators, and non-reporting traders are considered small traders who may be hedging or speculating.

Thus, the disaggregation of open interest performed by the CFTC occurs along three dimensions: reporting versus non-reporting, hedging versus speculating, and long versus short. For noncommercials, the CoT report also identifies how much of their position is held in a spread. The spread number measures the extent to which a speculator holds equal long and short positions. For example, if a speculator was long 100 contracts and short 70 contracts of a different expiration, then the spread contribution will be 70 and the long

[6]The CoT Reports are published on the CFTC website at http://www.cftc.gov/cftc/cftccotreports.htm.

contribution will be 30. The seven components of the total open interest (TOI) that are detailed in the CoT reports are related in the following manner:

$$\overbrace{\overbrace{[S_{Long} + S_{Short} + 2(S_{Spread})]}^{speculators} + \overbrace{[H_{Long} + H_{Short}]}^{hedgers}}^{reporting} + \overbrace{[NR_{Long} + NR_{Short}]}^{nonreporting} = 2(TOI) \tag{1.5}$$

where S, H and NR refer to positions, measured in the number of contracts, held by large speculators, large hedgers and non-reporting traders, respectively, with subscripts (long, short, and spread) indicating the type of positions. On the left side of equation (5), each contract is counted twice since both long and short positions are counted even though a long and a short position constitute only one contract. Consequently, the total open interest is doubled on the right side of the equation.

As can be seen from the sample CoT report[7] in Figure 3, the report also lists the number of traders in each category. Note that the sum of the numbers of traders from each category exceeds the total number of traders due to the fact that a single trader, such as a spreading speculator, can be counted in more than one category. A trader is counted in each category in which the trader holds a position. For example, a hedger who is long December 2008 crude oil while also being short June 2009 crude oil will be counted only once for the number of total traders, but will be counted in both the 'commercial long' and the 'commercial short' categories. Also of note is that the CoT is available in two versions: a futures-only version, and a futures and options version that converts options positions into equivalent number of futures positions. Since the current study is interested in the microstructure of futures markets, the futures-only version is used.

[FIGURE 3 ABOUT HERE]

Various measures of trader positions are constructed to illuminate the relationships between trader classes and market dynamics. These measures are defined as they are introduced in the next two sections.

[7]See [18] for a history of the CoT reports and a full description of CoT variables.

1.3 Hedgers and Speculators

The large majority of open interest is held by large hedgers and speculators. In the sample under study, an average of 84% of open interest is held by large traders that must report their holdings to the CFTC. Of the reportable holdings, large hedgers dominate the large speculators by holding, on average, 68% of the total open interest compared with 16% for the speculators. Equation (6) demonstrates the calculation of the percent of open interest (POI) held by large speculators. The POI for hedgers is calculated similarly, but without the spread component.

$$POI_t^S = \frac{S_{Long,t} + S_{Short,t} + 2(S_{Spread,t})}{2(TOI_t)} \qquad POI_t^H = \frac{H_{Long,t} + H_{Short,t}}{2(TOI_t)} \tag{1.6}$$

where S refers to positions held by large speculators with the subscripts (long, short, and spread) indicating the type of positions. From the time series of POI presented in Figure 4, the larger size of the positions of hedgers relative to speculators is clear - hedgers hold between 60% and 80% of the outstanding positions, speculators hold between 5% and 35% of the outstanding positions, and small traders hold the remaining positions.

[FIGURE 4 ABOUT HERE]

Another feature of POI evident in Figure 4 is the inverse relationship between movements in positions of hedgers and speculators. Within the sample, the percent of open interest held by large speculators and hedgers, POI_t^S and POI_t^H, have a correlation coefficient of -0.40, indicating that when one trader is entering a position, there is a tendency for a trader of another class to be closing a position. This observation is consistent with the proposition that different trader types provide each other with liquidity.

The main variable of interest in this study is the net long measure of trader positions, denoted NL_t^C for trader class C. The net long measures for speculators and hedgers are calculated by

$$NL_t^S = S_{Long,t} - S_{Short,t} \quad \text{and} \quad NL_t^H = H_{Long,t} - H_{Short,t}. \tag{1.7}$$

These measures indicate the number of contracts held as long positions that are not offset by short positions. A negative net long position indicates that, within the trader class, more contracts are held in short positions than in long positions. A key feature of this measure

is that it emphasizes trade between hedgers and speculators since trade within each class cannot change its net long position.

Summary sample statistics for the net long positions by trader types are presented in Table I. The positive mean, median, and skewness of the net long positions of speculators indicates a tendency of speculators to go long, while hedgers tend to go short. This observation reinforces the proposition that multiple trader types increase market liquidity since if speculators were not in the market, hedgers would have a harder time finding counterparties for their short positions. A second observation can be made from the sample kurtosis measures, both of which indicate thinner than Gaussian tails. If trader positions are bounded due to some budgetary or risk constraints, the distribution of their positions would be expected to be finite or, at most, thin.

[TABLE I ABOUT HERE]

As can be seen in Figure 5, the net long position of large hedgers and speculators tend to move in opposing directions. The strong sample correlation of -0.96 is not unexpected since a net change for one trader type requires another trader type to take an opposing position. A measure of -1 is not seen because of a small amount of trade between large and small traders. What is interesting, however, is that the relationship between small traders and large traders is significantly weaker, with correlations of $\rho(NL^S, NL^{small}) = 0.543$ and $\rho(NL^H, NL^{small}) = -0.755$. Since the small trader class is a mix of small speculators and hedgers, the behaviorial regularities associated with trader classes are weaker, though the signs of the correlations indicate that a position taken by small traders is more likely a speculation than a hedge.

[FIGURE 5 ABOUT HERE]

1.3.1 Trader Positions and Market Prices

It is well-understood that hedgers and speculators face different constraints and participate in futures markets for different purposes. However, it is unclear how these differences impact price dynamics in these markets. This section tests the null hypothesis that, despite their differences, hedgers and speculators are similar in their impact on prices through their

trades, as well as in their response to market action. This hypothesis is rejected and several regularities are found that differentiate hedgers and speculators.

To evaluate any potential differences in the relationship between trader positions and market prices for hedgers and speculators, three models are estimated:

$$NL_t^S = \alpha^1 + \beta_1^1 NL_{t-1}^S + \beta_2^1 r_t + \beta_3^1 r_{t-1} + \varepsilon_t^1 \tag{1.8}$$

$$NL_t^H = \alpha^2 + \beta_1^2 NL_{t-1}^H + \beta_2^2 r_t + \beta_3^2 r_{t-1} + \varepsilon_t^2 \tag{1.9}$$

$$r_t = \alpha^3 + \beta_1^3 r_{t-1} + \beta_2^3 NL_t^S + \beta_3^3 NL_{t-1}^S + \varepsilon_t^3 \tag{1.10}$$

where the error terms are assumed to be conditionally heteroskedastic. In order to account for potential endogeneity of returns and positions, continuous updating efficient generalized method of moments [39] was used to estimate the models. Changes in the number of traders holding long and short positions in each trader class were used as instruments for returns while levels of these variables were used as instruments for positions. The results, presented in Table II, indicate several significant differences between hedgers and speculators. First, by examining the concurrent relationship between trader positions and returns, speculators are found to increase prices with their purchases whereas hedgers tend to decrease prices through their purchases. This difference is consistent with the notion that speculators place more aggressive orders than hedgers. Hedgers may employ less aggressive limit orders in order to reduce their costs of hedging.

[TABLE II ABOUT HERE]

Another difference is found in the relationship between trader positions and lagged returns. The significant positive coefficient of lagged returns in the model of speculative net long positions indicates a tendency for speculators to engage in trend-following behavior. Large hedgers, on the other hand, have a tendency towards following contrarian strategies.

1.3.2 Trader Positions and Market News

A second area in which hedgers and speculators may differ is in their response to news about the crude oil market. In the more traditional equity and bond markets, one might expect that most participants would respond similarly to positive or negative news. However, this may not hold in futures markets where hedgers are not primarily seeking to profit from price

changes. To test the relationship between trader positions and market news, this section employs two methods of textual analysis: (i) a reduced dimensionality regression, and (ii) a high dimensionality binary classification.

These quantitative methods are chosen rather than an event-by-event chronological analysis of relevant world events for several reasons. While these methods do not capture the subtleties of complex political and economic news stories, they are objective, consistent, and reproducible methods that have been successfully used in other settings. Additionally, the regularities studied in the remainder of this paper have strong persistence over time, with some lasting more than 15 weeks. These long periods contain many events that are relevant to oil markets, and isolating specific effects would be difficult. Finally, given that many observations are collected for this study, we can abstract from the peculiarities of a single event and identify the underlying microstructural issues.

The first technique is similar to that used by [69] to reduce the typically high-dimensional character of textual data. They create a measure of news sentiment by identifying the fraction of positive and negative words in market-specific news stories. Positive and negative words are identified using the Harvard-IV-4 classification dictionary. Using the a traditional regression methodology, they find that the information captured by this simple variable has predictive power for earnings and equity returns. For the current study, approximately $58,000$ crude oil related news stories were used to construct following variables,

$$Neg_t = \frac{\#\text{ of negative words in week } t}{\#\text{ of total words in week } t} \qquad \text{and} \qquad neg_t = g\left(\frac{Neg_t - \mu_t^{Neg}}{\sigma_t^{Neg}}\right)$$

where μ_t^{Neg} and σ_t^{Neg} are the mean and standard deviation of Neg for the six months prior to time t, and where g is an exponential moving average[8] with a smoothing parameter of 0.9. Similarly, Pos_t and pos_t are created for positive words.

Using these proxies for the news content, the following model is estimated.

$$NL_t^S = \alpha + \beta_1 neg_t + \beta_2 pos_t + \beta_3 NL_{t-1}^S + \varepsilon_t \tag{1.11}$$

where ε_t is assumed to be a white noise process. Results for this estimation are listed in Table III. Note that NL_t^H was not included because of its strong correlation to NL_t^S and

[8]The results of this section is robust to a wide variety of smoothing operators and parameters.

the potential for multicollinearity. [69] found that for equity markets the negative words contained in news articles contained significant information about firm earnings, beyond what analysts' forecasts and historical accounting data did. Interestingly, for crude oil markets, it is often positive news that will put downward pressure on prices, while bad news drives prices up. After accounting for this difference, similar results are found in crude oil futures markets as speculators significantly decrease their net long positions in the presence of high levels of positive words. Hedgers on the other hand increase their net long positions during such periods.

[TABLE III ABOUT HERE]

Given its large number of words and linguistic relationships, textual data is naturally very highly dimensional. This high dimensionality cannot be maintained in most econometric techniques, however there are some more recent classification and regression frameworks which can accommodate such high dimensionality. The most successful technique in textual studies is the support vector machine (SVM) binary classifier [52, 17]. The basic SVM is a binary classifier that finds a hyperplane with the maximum margin between positive and negative training documents. The three key features of SVMs that reduce the likelihood of over-fitting and make them useful for textual classification are

- Not all training documents are used to train the SVM. Instead, only documents near the classification boarder are used.
- Not all features from the training documents are used, so excessive feature reduction is not needed.
- SMVs can construct irregular boarders between positive and negative training documents.

To confirm the systematic response of trader classes to market news, a binary SVM is used to classify whether the net long positions of trader classes will increase (+1) or decrease (-1). Weekly news is aggregated and the frequency of each unique word (there are $59,323$ unique words in the sample) per week is calculated. These frequency vectors are used as input to train the classifier[9]. Using cross-validation techniques, 300 out-of-sample

[9]The SVM used in this study is available at http://svmlight.joachims.org/.

classifications of whether speculative net long positions increased or decreased in response to market news gave the following results[10]:

$$\begin{aligned}
\text{Accuracy} &= \tfrac{TP+TN}{TP+FP+TN+FN} = 56\%^{**} \\
\text{Precision} &= \tfrac{TP}{TP+FP} = 54\%^{*} \\
&= \text{The proportion of obs. classified as } +1 \text{ that are truely } +1 \\
\text{Recall} &= \tfrac{TP}{TP+FN} = 81\%^{***} \\
&= \text{The proportion of obs. that truely are } +1 \text{ that are classified as } +1
\end{aligned}$$

where TP, FP, TN, and FN denote the number of true positives, false positives, true negatives, and false negatives produced by the classifier, respectively. When this classifier, trained on data of speculative positions, is applied to out-of-sample data on the positions of hedgers, significant misclassifications are observed indicating that hedgers and speculators are responding to news in significantly different ways.

While both of the techniques used in this section capture only a coarse level of linguistic sophistication, they do indicate a clear difference in the response of hedgers and speculators to market news. Results support the notion that speculators respond to news in a similar way to traders in equity markets. That is, with news that puts downward pressure on prices, speculators decrease the net long positions, while hedgers increase their long positions. Combined with the previous results, this section supports the notion that large speculators are informed traders who bring new information to the market through orders that are placed more aggressively than hedgers.

These differences between hedgers and speculators will become particularly pronounced when the net long positions of these two classes of traders diverge and approach their bounds. At such times, the available counterparties for a trade become less heterogeneous, and particular traits of a trader class may temporarily change the price dynamics. The remainder of this paper examines specifically how and why the price dynamics change during these times of extreme and opposing positions.

[10]Note that $*$ indicates $p < 0.10$, $**$ indicates $p < 0.05$, and $***$ indicates $p < 0.01$.

1.4 Extreme Positions and Market Dynamics

To study the market dynamics as trader positions approach their bounds, an event study methodology is used. The first step in this methodology is to define the extreme events that are being studied. To do this, the notion of trader positions, as expressed by the net long measure described in the previous section, must be extended to capture the notion that the positions that trader classes can achieve are bounded. Since the true bounds of positions taken by large hedgers and speculators would be difficult to determine, locally realized limits are used as a proxy for this true limit.

The market under study has been growing over the sample period, so an absolute value of the position bounds for all times is inappropriate. Instead, a local measure, $P_t^C(\tau)$ is employed which uses the maximum and minimum NL_t^C values achieved by trader class C within a moving window of τ periods. In this study, a centered window is used, thereby using information from the recent past and future. Consequently, in this ex post study, we can study return regularities, but no claim of predictability can be made from the results. The advantage of an ex post analysis is that a better proxy for position bounds is permitted. That is, extreme positions are easier to identify when future information is used, whereas in an ex ante setting, a large position may get labeled as extreme and then be dwarfed by an even larger position a few periods later. This trader position measure is expressed as

$$P_t^C(\tau) = \frac{NL_t^C - \min\{NL_{t'}^C\}}{\max\{NL_{t'}^C\} - \min\{NL_{t'}^C\}} \tag{1.12}$$

where $t' \in \{t - (\tau - 1)/2, \cdots, t - 1, t, t + 1, \cdots t + (\tau - 1)/2\}$ for an odd-valued τ. In an ex ante form, using only past information, this variable is similar to the market sentiment indicator of [16]. However, rather than relying on notions of sentiment, the current study argues that the resulting dynamics are the consequence of structural constraints imposed by the bounded size of the market which induces mean-reverting behavior when these bounds are approached.

For the current study, τ is taken to be around nine months (39 weeks), but the results are robust across many window sizes as is shown in Section IV. Clearly, values of $P_t^C(\tau)$ take values between 0 and 1, with values achieving 1 when net long positions exceed all values within the nine month centered moving window. Figure 6 plots this position index for large speculators and hedgers.

[FIGURE 6 ABOUT HERE]

The events of interest occur when one trader class is extremely long and the other class is extremely short. To identify the times of these events, two indicator variables are constructed. SL_t identifies times when speculators (S) are extremely long (L) and hedgers are extremely short, and HL_t identifies times when hedgers (H) are extremely long (L) and speculators are extremely short. These variables are defined to be

$$SL_t = \begin{cases} 1 & \text{if} \quad P_t^S(\tau) - P_t^H(\tau) = \max[P_{t'}^S(\tau) - P_{t'}^H(\tau) : t' \in [t - \frac{(\tau-1)}{2}, t + \frac{(\tau-1)}{2}]] \\ 0 & \text{otherwise} \end{cases} \tag{1.13}$$

$$HL_t = \begin{cases} 1 & \text{if} \quad P_t^S(\tau) - P_t^H(\tau) = \min[P_{t'}^S(\tau) - P_{t'}^H(\tau) : t' \in [t - \frac{(\tau-1)}{2}, t + \frac{(\tau-1)}{2}]] \\ 0 & \text{otherwise} \end{cases} \tag{1.14}$$

An SL-type event occurs when the difference between the position indices of speculators and hedgers is larger than at any other time within a centered window of length τ. For example, if $P_t^S(\tau) = 1$ and $P_t^H(\tau) = 0$, then speculators are extremely long and hedgers are extremely short, and so $SL_t = 1$ captures the extreme opposition of the positions of these two classes of traders. Within the sample period there are 20 events of each type.

Such extreme positions do not develop without reason and are clearly dependent on world events related to the crude oil markets. However, it is not within the scope of this paper to identify the specific causes of these extreme positions. That said, Section II.A identified divergent position accumulation by hedgers and speculators in response to market related news. Therefore, if news variables have temporal persistence, then extreme positions would regularly result from the chance occurrence of many news innovations of the same type (i.e. either a 'good' or 'bad' cluster of news events). Such persistence in news variables is in fact observed, with significant first order autocorrelations of $\rho_1(Neg_t) = 0.54$, $\rho_1(neg_t) = 0.97$, $\rho_1(Pos_t) = 0.45$, and $\rho_1(pos_t) = 0.96$.

1.4.1 Extreme Positions and the Number of Traders

The argument that bounded trader positions can induced liquidity shocks is based on the limited heterogeneity of counterparties when extreme positions are taken. Of course, the number of traders of each class in the market is another determinant of the heterogeneity of

available counterparties. As is shown below, the number of traders in the market and the acquisition of extreme positions is not independent. In the case of HL-type events, both factors move in conjunction and result in strong liquidity induced dynamics. In the case of SL-type events, the two factors more in opposition to each other thereby decreasing the strength of the liquidity-induced dynamics.

Up until this point in the analysis, no consideration has been given to the demographics of the groups of large hedgers and speculators. The Commitments of Traders Reports list the number of large traders in each of these groups, thereby allowing an analysis of their composition around the times that they acquire extreme positions. For the class of SL-type events, where speculators are extremely long and hedgers extremely short, the left panel of Figure 7 shows that, on average, traders are entering the market before the event date and then leave the market afterwards. Conversely, the right panel of the same figure shows that for the class of HL-type events, where hedgers are extremely long and speculators extremely short, traders tend to leave the market before the event date and then return afterwards.

[FIGURE 7 ABOUT HERE]

Two questions arise from the observation that traders leave or enter the market before and after extreme positions are taken: Who is it that is leaving the market, and why are they leaving? A partial answer to the first question is seen in Figure 8 which shows that the majority of the traders that systematically leave and enter the the market around the event dates are speculators, and not hedgers. The fact that hedgers do not enter or exit the market in a regular fashion around these extreme position events is due to the fact that their positions are not dependent on the profitability of their trades, but are rather determined by their business interests.

[FIGURE 8 ABOUT HERE]

The number of speculators can be further disaggregated into the number of speculators who are long and the number that are short. Figure 9 shows these numbers, thereby giving a fuller answer to the question of who is leaving the market and why they are leaving. The right panel of the figure indicates that virtually all of the traders that are leaving the market before the HL-type events are speculators that were long. Similarly, the majority of speculators who enter the market leading up to an SL-type event are also those speculators that were long. In the following sections the set of speculators that enter and exit long

positions are shown to be significant determinants of the market dynamics around these extreme events.

[FIGURE 9 ABOUT HERE]

1.4.2 Extreme Positions and Returns

It was shown in Section II that a strong relationship exists between trader positions and market returns. Purchases by speculators, and sales by hedgers, tend to push prices higher. Sales by speculators, and purchases by hedgers, tend to push prices lower. This relationship implies that extreme events are often preceded by periods of trending prices. Once trader positions are near their upper or lower bounds, they inevitably move back towards their average since they cannot become much more extreme than they currently are. When trader positions revert away from their extremes, the price trend often reverses.

In the case of SL-type events, where speculators are extremely long and hedgers extremely short, there is a significant increase in prices that begins around seven weeks before the extreme event resulting, on average, in an 8.8% cumulative abnormal return by the event date, where the abnormal return is defined to be the return minus the sample mean weekly return of $\overline{r} = 0.36\%$. More formally, the cumulative abnormal return between times $t - \tau$ and t is

$$CAR(t-\tau, t) = \prod_{i=0}^{\tau}(1 + r_{t-i} - \overline{r}) - 1. \tag{1.15}$$

After the event, the upward trend is reversed and prices tend to fall for the next eight weeks. Thus, SL-type events tend to coincide with local price maximums. In the case of HL-type events, where speculators are extremely short and hedgers extremely long, there is a significant decrease in prices that begins around six weeks before the extreme event resulting, on average, in a -9.3% cumulative abnormal return by the event date. The price trend reverses after the event date and tends to rise for the next seven weeks. Thus, HL-type events tend to coincide with local price minimums. Both of these cases are seen in Figure 10.

[FIGURE 10 ABOUT HERE]

To evaluate the effect of trader positions on returns for the week before and after those positions become extreme, the following model is estimated

$$r_t = \alpha + \beta_1 SL_t + \beta_2 SL_{t-1} + \gamma_1 HL_t + \gamma_2 HL_{t-1} + \varepsilon_t \tag{1.16}$$

where ε_t is assumed to be a white noise process, SL_t is an indicator variable for times when large speculators are extremely long and large hedgers are extremely short, and HL_t is an indicator variable for times when large speculators are extremely short and large hedgers are extremely long. The result of this estimation is found in Table IV. All of the estimated variable coefficients except for γ_1 are significant. Since the return, r_t, measures the price change between times t and $t-1$, β_1 is the change in price leading up to an SL-type event occurring at time t. Similarly, β_2 is the change in price that occurs when the SL-event occurred at time $t-1$, that is, the change in price after such an event. Estimates of β_1 and β_2 are positive and negative, respectively, indicating that prices tend to rise during the week before an SL-type event and fall the following week. Conversely, estimates of γ_1 and γ_2 are negative and positive, respectively, indicating that prices tend to fall during the week before an HL-type event and rise the following week. The low R^2 of this and other regressions in the paper are low, as expected, since there are only 20 events of each type. Clearly, there is a lot more going on in the return dynamics than just those resulting from extreme trader positions. However, when such events occur, the effects are significant, both economically and statistically. These results are consistent with the price trends and reversals around extreme events seen in Figure 10.

[TABLE IV ABOUT HERE]

This relationship between positions and returns can be explained through the bounded and mean reverting nature of extreme positions, but there is also an influence from the entry and exit of speculators around these event dates. Recall from the right panel of Figure 9 that virtually all of the traders that are leaving the market before the HL-type event are speculators who are long. Given the returns during this time, these traders were long in a falling market, so it is likely that they were leaving the market to avoid further losses. At the same time leading up to the event date, the number of speculators that are short is actually increasing. These traders who are selling short in a falling market are realizing significant profits, and immediately after the event date they begin to take their profits by

closing their short positions. This is seen in Figure 9 where the number of speculators who are short after the HL-type event dates steadily falls. Also after the event date, the traders who were long and had left the market begin to re-enter the market establishing new long positions. Thus, there are regularities in the entry and exit of traders around the event dates that are consistent with speculative profit-taking and with capital constraints forcing traders to exit losing trades. A similar argument can be made to explain the regularities around SL-type events in the left panel of Figure 9.

To check whether these regularities in the entry and exit of speculators plays a role in the microstructure of returns the following model is estimated,

$$r_t = \alpha + \beta_1 \triangle T^{spec}_{long,t} + \beta_2 \triangle T^{spec}_{short,t} + \varepsilon_t \tag{1.17}$$

where $\triangle T^{spec}_{long,t}$ is the change since the previous period in the number of speculators who are long, $\triangle T^{spec}_{short,t}$ is the change in the number of speculators who are short, and ε_t is a white noise process. The estimation results in Table V indicate that an increase in the number of long speculators is associated with a significantly higher return.

[TABLE V ABOUT HERE]

An interesting observation arises when market demographics are examined around the price reversals associated with SL and HL-type events. If a speculative trader were aware of the date of a significant reversal of price movements, then they would take positions that would profit from this price action. An informed trader[11] would therefore want to be long on an HL-event date and short on an SL-event date. However, these events are essentially defined to be times when large speculators are on the wrong side of the market. On average, at SL-event dates, only 36% of large speculators are short, and only 38% are long on HL-event dates. Thus, an essential feature of these liquidity shocks is that a significant number of traders be uninformed in their timing.

Figure 9 shows that the majority of new speculative bets are made in the 'correct' direction with respect to the price trends both before and after the event dates. The longer the trend has been in effect, the more speculators trade in that direction. However, some

[11]Since a hedger's trading behavior is determined by business interests and risk management strategies, their actions would be independent of knowledge of these event dates. In this sense, hedgers act as if they are uninformed traders.

traders enter the trend too late as is indicated by the spike at time zero. Since only aggregate trader numbers are reported in the CoT reports, it is impossible to tell when a specific trader both enters and exits the market. Consequently, the distribution of profits within a class of traders is unknown.

With the majority of large speculators being on the wrong side of the market at SL and HL-event dates, some conclusions about uninformed traders can be drawn. First, it would be inappropriate to model these uninformed traders as being equally likely to take long or short positions, since that would imply that on average the majority of speculators would be on the right side of the market[12]. Second, uninformed trades appear to be dependent on past trades since they appear to want to follow a trend that is about to end. Finally, the number of speculators trading in the direction of a trend tends to increase with the duration of the trend. This seems to indicate that there are varying degrees to which traders are informed, with well informed traders entering a trend early, other informed traders entering as the trend develops, and the uninformed traders entering too late as the trend ends.

1.4.3 Extreme Positions and Realized Volatility

So far in the analysis, significant differences in the behavior of hedgers and speculators, as well as the bounded nature of trader positions, have been found to drive return regularities around dates of extreme and opposing trader positions. These dates also coincide with systematic entries and exits of large speculators from the market, primarily by those holding long positions. In both SL and HL-type events, the changes in the number of speculators who are short moves in opposition to the changes in the number of speculators who are long, however the magnitude of the latter changes are more than twice as large. Since speculators facing losses may leave a market through orders placed more aggressively than other orders, there is reason to suspect that return volatility may differ between SL-type and HL-type events.

As was shown earlier in Figure 10, HL-type events are characterized by first falling and then rising prices. As prices initially fall, a significant number of speculators holding long

[12]If uninformed traders are equally likely to take long or short positions, then in expectation half of the uninformed traders are on the right side of the market. Since informed traders would be on the right side of the market as well, then overall there are more traders on the right side of the market than on the wrong side.

positions leave the market (see the right panel of Figure 9). These speculators were long in a falling market, and hence exited the market while they were facing increasing losses. Under such circumstances, orders to exit a losing trade may well be placed more aggressively than would ordinarily be observed. Such aggressive orders would be observed as an increase in the return volatility. After the price trend reverses at the event date, and speculators return to the market, volatility is expected to return to normal levels.

Around SL-type events, a similar argument would suggest that speculators holding long positions who leave the market after the event date would also drive up the return volatility. To examine the conjecture that speculators facing losses leave the market using aggressive orders thereby causing high return volatilities, this section examines the behavior of the realized volatility around both types of extreme position events. Realized volatilities are sampled weekly and constructed with five-minute returns, as described in Section I. Figure 11 plots the average realized volatility around the event dates. There is a clear difference between the two classes of extreme events. The average realized volatility before and immediately after type-SL extreme positions, where speculators are long and hedgers are short, is significantly lower than the overall sample average[13]. The sharp increase in volatility after SL-event dates is consistent with the 'leverage effect' where volatility is rising in falling markets ([7]). In contrast, it is significantly higher before and just after HL-type extreme positions, where speculators are short and hedgers are long. In this case, the average realized volatility on the event date is at the 79th percentile of the entire sample and is even higher for several weeks before the event date. Both panels of Figure 11 are consistent with the notion that traders facing losses can significantly increase return volatility as they exit the market.

[FIGURE 11 ABOUT HERE]

To quantify the increase in volatility around HL-type events, the following model is estimated:

$$\ln(RV_t) = \alpha + \beta \ln(RV_{t-1}) + \gamma HL_t^* + \varepsilon_t \tag{1.18}$$

where RV_t is the realized volatility at time t, and ε_t is assumed to be a white noise process.

[13]For Figure 11 and the following estimations, a single outlier realized volatility value was removed from the series. The outlier was on 24 March 1998 and was over 19 standard deviations away from the sample mean. The value was replaced with an average of the preceding and following values. After examining the high-frequency price data, the outlier appears to have been caused by an incorrect price entry.

HL_t^* is an indicator variable for whether a HL-type extreme event occurs at time t or during the next five weeks following t. The five week window captures the fact that volatility builds, and in fact peaks, during the several weeks leading up the the event date. This new indicator is defined more formally as

$$HL_t^* = \begin{cases} 1 & \text{if} \quad HL_t \vee HL_{t+1} \vee HL_{t+2} \vee HL_{t+3} \vee HL_{t+4} \vee HL_{t+5} \\ 0 & \text{otherwise} \end{cases} \tag{1.19}$$

where $\vee$ is the logical 'OR' operator. The results of this estimation are found in Table VI. The estimated coefficient of HL_t^* indicates realized volatility is expected to rise by 6.6% five weeks prior to an HL-type event and remain higher until the event date. Since HL_t^* dates occur in groups of six consecutive weeks, and since the model (19) accounts for previous weeks, the realized volatility at the event date will tend to be 17% higher than it was six weeks prior to the event date, indicating that HL-type extreme events can be an economically important cause of high-volatility periods.

[TABLE VI ABOUT HERE]

The different volatility effects that are observed before and after HL and SL-type events, as well as the trend reversals seen at these times, it is likely that realized volatilities are not symmetrically composed of positive and negative returns. Figure 12 plots the average positive and negative directional realized volatilities for the weeks around HL-type events, when hedgers are extremely long and speculators extremely short. It is clear that the realized volatilities before and after the event date are composed of two nonsymmetric components. Before the event, the majority of the volatility comes from negative returns as is seen from the increase in DRV_t^N before the event date in Figure 12(a). After the event date, DRV_t^N falls sharply, but DRV_t^P, in Figure 12(b), peaks thereby keeping the overall realized volatility higher than average. This asymmetric composition of return volatility indicates that the potential ill-effects of liquidity shocks do not affect all market participants equally. High negative directional volatility would be bad for sellers, while high positive directional volatility would be bad for buyers. Further, a significant difference between the negative and positive volatilities indicates trending prices.

[FIGURE 12 ABOUT HERE]

1.4.4 Extreme Positions and Bid-Ask Spreads

The significance of these dynamic regularities around extreme position dates can be further verified using the prominent percentage bid-ask spread measure of liquidity. Figure 13 plots the average daily percentage bid-ask spread around both types of event dates. The pattern of larger than average spreads around HL-type events and lower than average spreads around SL-type events are similar to the patterns of realized volatility around these dates. With fewer speculators in the market at HL-type event dates, trade can only be induced by offering better prices to potential buyers and sellers resulting in a larger bid-ask spreads. On SL-type event dates, there are more speculators in the market and finding willing counterparties is easier, resulting in smaller bid-ask spreads. These findings indicate that, despite the fact that speculators can raise return volatility as they close losing positions, their increase presence in the market during SL-type events improves liquidity. While this study is concerned with the movement of large speculators into and out of a single market, if these movements actually constitute movements from or into other futures markets, then in addition to impacting liquidity within the crude oil futures market, these movements of traders may also account for some of the time-varying cross-market liquidity variation [68] in futures markets.

[FIGURE 13 ABOUT HERE]

To quantify the effects of these events on the bid-ask spread, the following model is estimated:

$$S_t = \alpha + \beta(HL_{t-1} + HL_t + HL_{t+1}) + \gamma(SL_{t-1} + SL_t + SL_{t+1}) + \varepsilon_t \tag{1.20}$$

where S_t is the estimated daily percentage bid-ask spread described in Section I, and ε_t is assumed to be a white noise process. Since extreme events cannot occur in consecutive weeks, $(HL_{t-1} + HL_t + HL_{t+1})$ and $(SL_{t-1} + SL_t + SL_{t+1})$ are indicator functions taking the value 1 when the time t is within one week of an extreme event and taking the value 0 otherwise. Table VII contains the estimation results where we see that both β and γ are significant. The coefficient β has the expected positive sign since the spreads are higher on HL-type event dates when there are fewer speculators in the markets, and γ has the expected negative sign since there are more speculators in the market during SL-type events resulting in smaller spreads.

[TABLE VII ABOUT HERE]

1.5 Robustness Checks

The results of this study depend largely on the identification of dates during which hedgers and speculators are holding extreme and opposing positions. The extremity of trader positions is defined by their size relative to positions held within a centered window of width τ weeks. Thus, τ is a critical parameter in the study and the results should be robust to changes in its value.

Return and realized volatility regularities are central results of this study, and models (17) and (19) were used to quantify these regularities. The estimation of these models used a centered moving window of $\tau = 39$ weeks. The results were also checked for window sizes $\tau = 31, 33, 35, 37, 41, 43, 45$, and 47 weeks. The coefficient estimates and their p-values for models (16) and (18) for each of these window lengths are listed in Table VIII. For robustness to hold, the sign, magnitude, and significance of these estimates should be reasonably close to each other and to the values reported in this paper for $\tau = 39$. Of the 45 coefficient estimates for the crude oil market in Table VIII, all have the same sign as those listed in the paper, all are of the same order of magnitude as those listed in the paper, and all but four are significant or not, as listed in the paper.

[TABLE VIII ABOUT HERE]

Similar results should be found in other futures markets where the distinction between hedgers and speculators is clear. The lower portion of Table VIII lists the estimates of models (16) and (18) for three additional markets: Soybeans from the Chicago Board of Trade, Live Cattle and 3-Month Eurodollars, both from the Chicago Mercantile Exchange. All estimated coefficients from the model of returns have the same sign as those of crude oil, though several are not significant. The model for Eurodollars returns has no significant estimates, and this is likely because of the fuzzy distinction between hedgers and speculators in financial futures markets. When a trader uses a market for both hedging and speculating, as is common in interest rate and equity markets, they are identified as hedgers in the Commitments of Traders data even if their dominant activity is speculative. This reporting standard causes the identifiable systematic differences between hedgers and speculators to disappear. In agricultural markets, on the other hand, the classes of hedgers and speculators are much more distinct, and consequently more estimates are significant.

Cross-sectional differences across markets are more pronounced in the volatility regularities around event dates. For example, as seen in the top panels of Figure 14, there is a clear rising of volatility around SL-type events and a clear falling of volatility around HL-type events in the Soybean market. This is the opposite of the effect found in crude oil futures, and is due to the confounding effects of the differing systematic movements of speculators into and out of the market around these dates.

[Figure 14 ABOUT HERE]

There is a large literature on the market equilibrium approach to futures pricing [33, 30] which use fundamental market factors to model prices. For crude oil, prominant examples of such factors include convenience yields, crude oil inventories, and crude oil production measures. Data for crude oil inventories, production and spot prices are available from the Energy Information Administration[14] (E.I.A.), and 1-month Treasury bill rates are available from Kenneth French's website[15]. While several event studies of these factors did not reveal any systematic regularities around extreme position accumulations, a full incorporation of these variables is difficult due to different release dates and frequency of observations. Given the current focus on the microstructural mechanism of liquidity shocks, investigating the relationship between extreme positions and market fundamentals is left for future study.

In summary, the central results of this study are robust within the crude oil futures market across several neighboring parameter values, and supporting evidence is found in agricultural futures markets where the hedger-speculator distinction is clear. There appear to be some cross-sectional differences between markets that induce differing regularities around event dates. Since these results depend on reliable differences in the behavior of hedgers and speculators, supporting results are not found in financial futures markets where the distinction between hedgers and speculators is not always clear. Finally, there may be fundamental factors at play which drive position accumulation, however the relationship between trader positions and these factors is not simple and will be pursued in a future study.

[14] http://www.eia.doe.gov/

[15] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/

1.6 Conclusion

Liquidity is a complex notion that captures the impact and costs of trading in a market. In liquid markets, trades are executed with low costs and result in little impact on the market price. Yet even in large and active markets, there are periods of illiquidity characterized by high volatility and bid-ask spreads, and even directional return regularities. This paper presents an empirical microstructural study of such liquidity-induced dynamics in the New York Mercantile Exchange's crude oil futures market, a very large and active market which is usually considered a very liquid market.

The paper's results suggest that during times when large hedgers and speculators acquire large and opposing positions, with one group going extremely long and the other going extremely short, liquidity-induced dynamics temporarily dominate price action. These dynamics can include price trend reversals, high volatility, and high bid-ask spreads. The underlying mechanism that cause these effects has three components: (1) significant differences between hedgers and speculators, (2) the bounded and mean-reverting nature of trader positions, and (3) the various conditions that cause speculators to enter and exit the market.

Large hedgers and speculators are the dominant participants in futures markets, and each group have different constraints as well as different reasons for participating in the market. Their impact on prices through their trades are also found to be significantly different. Large speculators are found to increase prices with their purchases, to engage in trend-following behavior, and to react to market-related news. Large hedgers, on the other hand, decrease prices through their purchases and exhibit contrarian-type behavior.

Speculators and hedgers are also found to differ in their response to news about the crude oil market. In two distinct analyses using over 58,000 news stories relating to crude oil markets, we found that news variables have temporal persistence and that traders systematically accumulate positions (with positions of hedgers and speculators diverging) in response to news. Combined with the previous results, these results support the notion that large speculators are informed traders who bring new information to the market through orders that are placed more aggressively than hedgers.

While positions of hedgers and speculators can diverge in response to news, there are bounds on how far they may diverge. When the positions held by hedgers and speculators

diverge and approach their limits, one class of traders begins to dominate the set of available counterparties for purchases, and the other class dominates that for sales. When this happens, the differences between trader classes, which are usually 'washed out' when the set of counterparties is well mixed, begin to have a pronounced impact on price changes. Prices trend strongly until position bounds are approached, and then reverse as positions revert towards normal levels.

The impact on liquidity during these episodes is found to be largely determined by the entry and exit of large speculators from the market. In particular, speculators facing increasing losses exit the market using orders that are significantly more aggressive than typical orders, thereby pushing up return volatility. On the other hand, speculators entering the market tend to improve market liquidity.

The asymmetry of liquidity shocks is also investigated. Both before and after speculators are extremely short and hedgers extremely long, realized volatility is significantly higher than its average. However, while these positions are being accumulated the volatility is dominated by negative returns. As these positions are being unwound, the volatility is dominated by positive returns. These effects are captured by a directional realized volatility measure and indicate that liquidity shocks can affect different market participants in opposing ways. That is, an illiquid market for a seller can be a liquid market for a buyer, and liquidity measures should be able to capture such asymmetries.

Trader Class	Min	Max	Median	Mean	Std Dev	Skewness	Kurtosis
Speculators	$-71,928$	$88,712$	$8,361$	$8,471$	$30,544$	0.06207	2.777
Hedgers	$-103,854$	$94,868$	$-8,103$	$-7,804$	$39,144$	-0.05284	2.617

Table 1.1: Summary Sample Statistics for the Net Long Measure of Position Holdings for Large Hedgers and Speculators. Statistics are calculated from 698 weekly observations from September 30, 1992 until February 28, 2006, as reported in the CFTC Commitments of Traders Reports.

$$NL_t^S = \alpha^1 + \beta_1^1 NL_{t-1}^S + \beta_2^1 r_t + \beta_3^1 r_{t-1} + \varepsilon_t^1$$

	Intercept	NL_{t-1}^S	r_t	r_{t-1}
Coefficient	-1388.58	0.89	$417,665$	$90,670$
t-stat	-2.20	44.20	14.75	5.94
p-value	0.03	0.00	0.00	0.00
R^2	0.91			

$$NL_t^H = \alpha^2 + \beta_1^2 NL_{t-1}^H + \beta_2^2 r_t + \beta_3^2 r_{t-1} + \varepsilon_t^2$$

	Intercept	NL_{t-1}^H	r_t	r_{t-1}
Coefficient	1991.80	0.88	$-549,829$	$-124,626$
t-stat	2.55	46.25	-14.60	-6.55
p-value	0.01	0.00	0.00	0.00
R^2	0.91			

$$r_t = \alpha^3 + \beta_1^3 r_{t-1} + \beta_2^3 NL_t^S + \beta_3^3 NL_{t-1}^S + \varepsilon_t^3$$

	Intercept	r_{t-1}	NL_t^S	NL_{t-1}^S
Coefficient	0.0020	-0.23	0.0000026	-0.0000023
t-stat	1.39	-5.89	12.57	-11.45
p-value	0.17	0.00	0.00	0.00
R^2	0.35			

Table 1.2: Trader Positions and Returns. Estimation is through continuous updating GMM. Statistics are calculated from 698 weekly observations from September 30, 1992 until February 28, 2006. Trader positions are reported in the CFTC Commitments of Traders Reports and the weekly returns are calculated from tick-level price data from TickData, www.tickdata.com.

$$NL_t^S = \alpha + \beta_1 neg_t + \beta_2 pos_t + \beta_3 NL_{t-1}^S + \varepsilon_t$$

	Intercept	neg_t	pos_t	NL_{t-1}^S
Coefficient	27,524	312,887	−1,263,992	0.90
Std Error	20,187	410,409	551,890	0.02
t-stat	1.36	0.76	−2.29	44.08
p-value	0.17	0.45	0.02	0.00
R^2	0.84			

Table 1.3: Impact of Positive and Negative Words on Trader Positions. Statistics are calculated from 405 weekly observations from July 25, 2001 until April 28, 2009. The *pos* and *neg* variables represent the frequency of positive and negative words in news articles from each week. Trader positions are reported in the CFTC Commitments of Traders Reports.

$$r_t = \alpha + \beta_1 SL_t + \beta_2 SL_{t-1} + \gamma_1 HL_t + \gamma_2 HL_{t-1} + \varepsilon_t$$

	Intercept	SL_t	SL_{t-1}	HL_t	HL_{t-1}
Coefficient	0.000	0.019	−0.031	−0.018	0.024
Std Error	0.002	0.007	0.007	0.012	0.011
t-stat	0.095	2.530	−4.331	−1.465	2.130
p-value	0.924	0.012	0.000	0.143	0.034
R^2	0.025				

Table 1.4: Regression Model of the Effect of Current and Past Extreme Trader Positions on Weekly Returns. The variable SL_t is an indicator variable for times when large speculators are extremely long and large hedgers are extremely short, and similarly HL_t is an indicator variable for times when large speculators are extremely short and large hedgers are extremely long. Heteroskedasticity robust standard errors and p-values are reported. The error term is assumed to be from a white noise process. Statistics are calculated from 698 weekly observations from September 30, 1992 until February 28, 2006. Trader positions are reported in the CFTC Commitments of Traders Reports and the weekly returns are calculated from tick-level price data from TickData, www.tickdata.com.

$$r_t = \alpha + \beta_1 \triangle T_{long,t}^{spec} + \beta_2 \triangle T_{short,t}^{spec} + \varepsilon_t$$

	Intercept	$\triangle T_{long,t}^{spec}$	$\triangle T_{short,t}^{spec}$
Coefficient	0.000	0.003	−0.001
Std Error	0.002	0.000	0.000
t-stat	−0.080	10.771	−1.524
p-value	0.937	0.000	0.128
R^2	0.173		

Table 1.5: Regression Model of the Effect of Changes in the Number of Speculators on Weekly Returns. The variable $\triangle T_{short,t}^{spec}$ is the change in the number of speculators who are short, and $\triangle T_{long,t}^{spec}$ is the change in the number of speculators who are long. Heteroskedasticity robust standard errors and p-values are reported. The error term is assumed to be from a white noise process. Statistics are calculated from 698 weekly observations from September 30, 1992 until February 28, 2006. Trader positions are reported in the CFTC Commitments of Traders Reports and the weekly returns are calculated from tick-level price data from TickData, www.tickdata.com.

$$\ln(RV_t) = \alpha + \beta \ln(RV_{t-1}) + \gamma HL_t^* + \varepsilon_t$$

	Intercept	$\ln(RV_{t-1})$	HL_t^*
Coefficient	−2.254	0.629	0.066
Std Error	0.712	0.116	0.027
t-stat	−3.167	5.434	2.464
p-value	0.002	0.000	0.014
R^2	0.465		

Table 1.6: Regression Model of the Effect of Extreme Trader Positions on Weekly Realized Volatility. RV_t is the realized volatility at time t, HL_t^* is an indicator variable for whether a HL-type event occurs within the next five weeks, and ε_t is assumed to be a white noise process. HL_t-type events occur when large speculators are extremely short and large hedgers are extremely long. Heteroskedasticity robust standard errors and p-values are reported. The error term is assumed to be from a white noise process. Statistics are calculated from 698 weekly observations from September 30, 1992 until February 28, 2006. Trader positions are reported in the CFTC Commitments of Traders Reports and the weekly returns are calculated from tick-level price data from TickData, www.tickdata.com.

$$S_t = \alpha + \beta(HL_{t-1} + HL_t + HL_{t+1}) + \gamma(SL_{t-1} + SL_t + SL_{t+1}) + \varepsilon_t$$

	Intercept	β	γ
Coefficient	0.112	0.012	−0.009
Std Error	0.001	0.004	0.004
t-stat	96.05	3.258	−2.351
p-value	0.000	0.001	0.025
R^2	0.025		

Table 1.7: Regression Model of the Effect of Extreme Trader Positions on Average Daily Percentage Bid-Ask Spreads. The variable SL_t is an indicator variable for times when large speculators are extremely long and large hedgers are extremely short, and similarly HL_t is an indicator variable for times when large speculators are extremely short and large hedgers are extremely long. Statistics are calculated from 698 weekly observations from September 30, 1992 until February 28, 2006. Trader positions are reported in the CFTC Commitments of Traders Reports and the weekly returns are calculated from tick-level price data from TickData, www.tickdata.com.

$$r_t = \alpha + \beta_1 SL_t + \beta_2 SL_{t-1} + \gamma_1 HL_t + \gamma_2 HL_{t-1} + \varepsilon_t$$

$$\ln(RV_t) = \alpha + \beta_3 \ln(RV_{t-1}) + \delta HL_t^* + \varepsilon_t$$

τ	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\gamma}_1$	$\widehat{\gamma}_2$	$\widehat{\delta}$
31	0.025	−0.034	−0.015	0.020	0.117
	(0.000)	(0.002)	(0.155)	(0.038)	(0.010)
33	0.026	−0.024	−0.016	0.029	0.118
	(0.000)	(0.002)	(0.099)	(0.001)	(0.011)
35	0.026	−0.024	−0.017	0.023	0.093
	(0.000)	(0.002)	(0.547)	(0.036)	(0.057)
37	0.021	−0.028	−0.012	0.021	0.122
	(0.007)	(0.000)	(0.370)	(0.064)	(0.029)
39	0.019	−0.031	−0.018	0.024	0.132
	(0.012)	(0.000)	(0.143)	(0.034)	(0.014)
41	0.021	−0.027	−0.016	0.024	0.126
	(0.007)	(0.001)	(0.246)	(0.043)	(0.020)
43	0.026	−0.031	−0.015	0.023	0.141
	(0.000)	(0.000)	(0.296)	(0.050)	(0.012)
45	0.026	−0.034	−0.008	0.018	0.140
	(0.000)	(0.000)	(0.627)	(0.132)	(0.020)
47	0.026	−0.034	−0.011	0.017	0.146
	(0.001)	(0.000)	(0.486)	(0.182)	(0.018)
Soybeans	0.004	−0.013	−0.012	0.011	−0.102
$\tau = 39$	(0.658)	(0.076)	(0.017)	(0.050)	(0.033)
Live Cattle	0.008	−0.013	−0.002	0.007	0.137
$\tau = 39$	(0.013)	(0.000)	(0.795)	(0.260)	(0.002)
Eurodollar	0.000	−0.000	−0.000	0.000	0.106
$\tau = 39$	(0.119)	(0.600)	(0.768)	(0.680)	(0.109)

Table 1.8: Robustness Check on the Size of the Centered Window Size and on Other Markets. A potentially critical parameter in the determination of extreme event dates is the size of the moving window τ. Throughout the paper $\tau = 39$ weeks is used. This table shows that the sign, magnitude, and significance of the critical regression coefficients are reasonably close for 8 other window lengths. Statistics are calculated from 698 weekly observations from September 30, 1992 until February 28, 2006. Trader positions are reported in the CFTC Commitments of Traders Reports and the weekly returns are calculated from tick-level price data from TickData, www.tickdata.com.

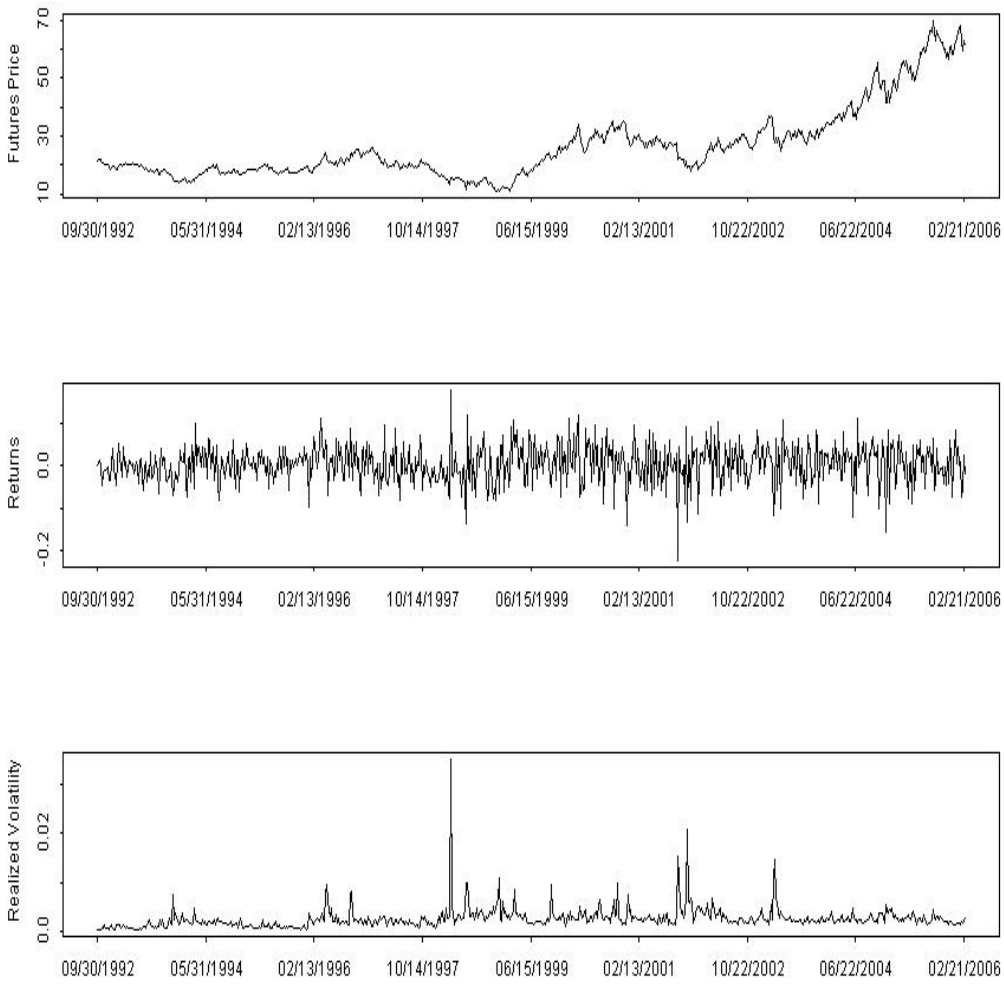


Figure 1.1: Weekly Prices (top frame), Returns (middle frame), and Realized Volatility (bottom frame) for NYMEX Light Sweet Crude Oil Futures. Prices are the closing price on each Tuesday. Returns are based on an unleveraged investment of the value of the underlying commodity at the futures price of the previous Tuesday. Realized volatility is also sampled each Tuesday, and is constructed by summing the squared five-minute returns over the previous week. The sample period is from September 30, 1992 to February 28, 2006 with 698 observations. Source: TickData, www.tickdata.com.

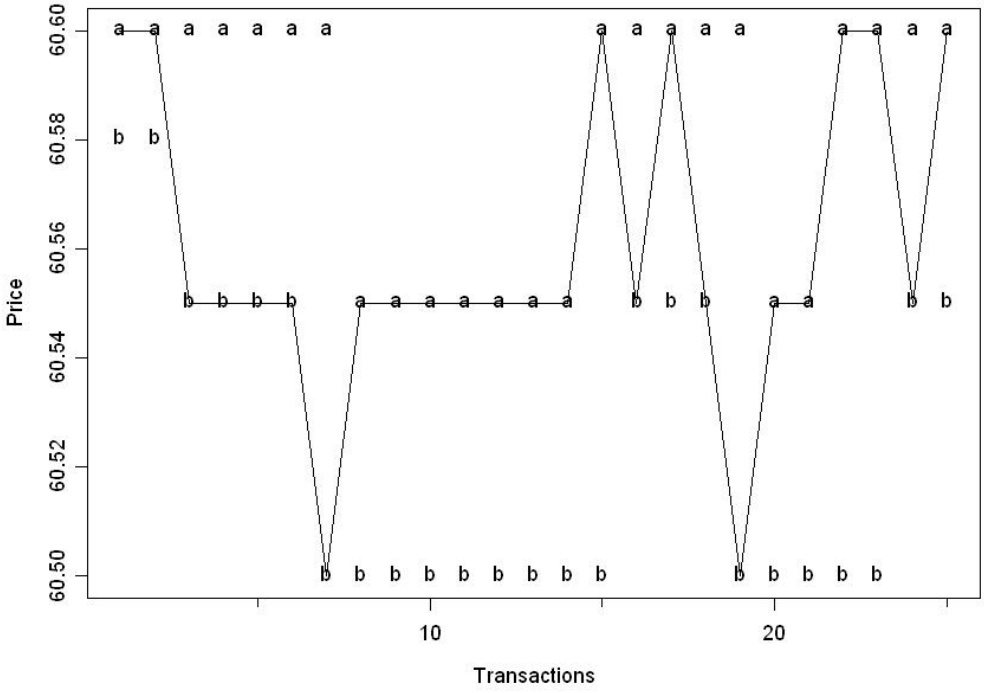


Figure 1.2: Transaction Prices with Estimated Bids and Asks. Transaction prices are shown by the solid line, with the estimated bids for each transaction at the 'b' markers and the estimated asks at the 'a' markers. The estimated bid-ask spread is $a_t - b_t$. Source: TickData, www.tickdata.com.

CRUDE OIL, LIGHT SWEET - NEW YORK MERCANTILE EXCHANGE Code-067651
FUTURES ONLY POSITIONS AS OF 07/17/07

NON-COMMERCIAL			COMMERCIAL		TOTAL		NONREPORTABLE POSITIONS	
LONG	SHORT	SPREADS	LONG	SHORT	LONG	SHORT	LONG	SHORT
(CONTRACTS OF 1,000 BARRELS)					OPEN INTEREST:		1,549,425	
COMMITMENTS								
246,844	137,421	310,801	911,229	1015492	1468874	1463714	80,551	85,711
CHANGES FROM 07/10/07 (CHANGE IN OPEN INTEREST:					3,053)			
-2,201	663	12,566	-10,804	-18,584	-439	-5,355	3,492	8,408
PERCENT OF OPEN INTEREST FOR EACH CATEGORY OF TRADERS								
15.9	8.9	20.1	58.8	65.5	94.8	94.5	5.2	5.5
NUMBER OF TRADERS IN EACH CATEGORY (TOTAL TRADERS:					324)			
107	94	127	91	102	268	269		

Figure 1.3: Example of a CFTC Commitments of Traders (CoT) Report. Released every Friday, the CoT report summarizes the positions of various classes of traders as of the previous Tuesday. The commercial class refers to large hedgers, and the non-commercial class refers to large speculators. Non-reportable positions are those held by small traders. The Commitments of Traders Reports are published every Friday on the CFTC website at http://www.cftc.gov/cftc/cftccotreports.htm.

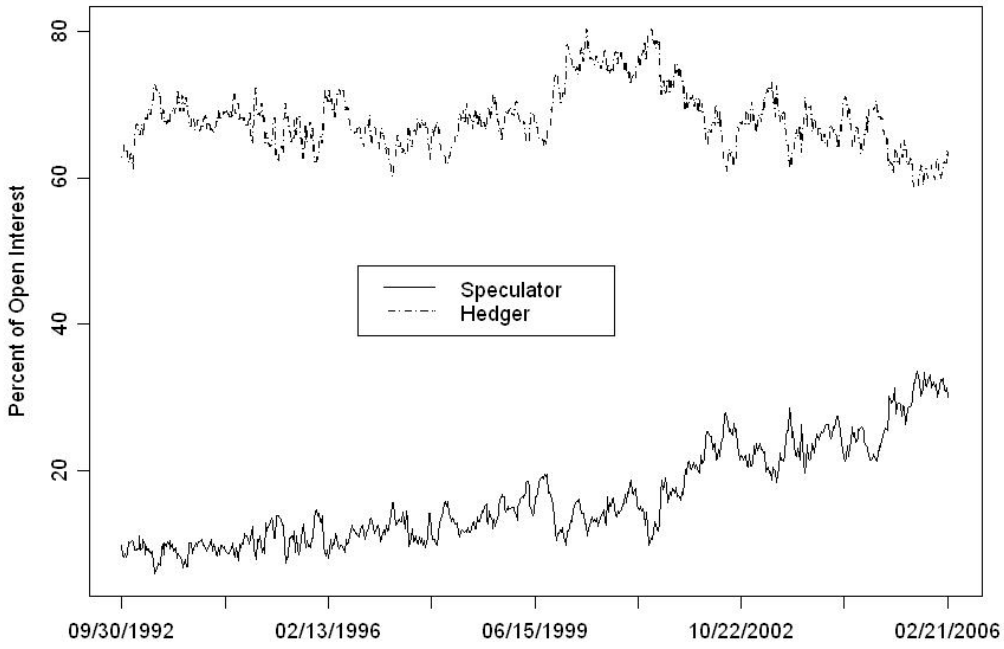


Figure 1.4: Percent of Open Interest for Large Hedgers and Speculators. Information is from 698 weekly observations from September 30, 1992 until February 28, 2006. Trader positions are reported in the CFTC Commitments of Traders Reports.

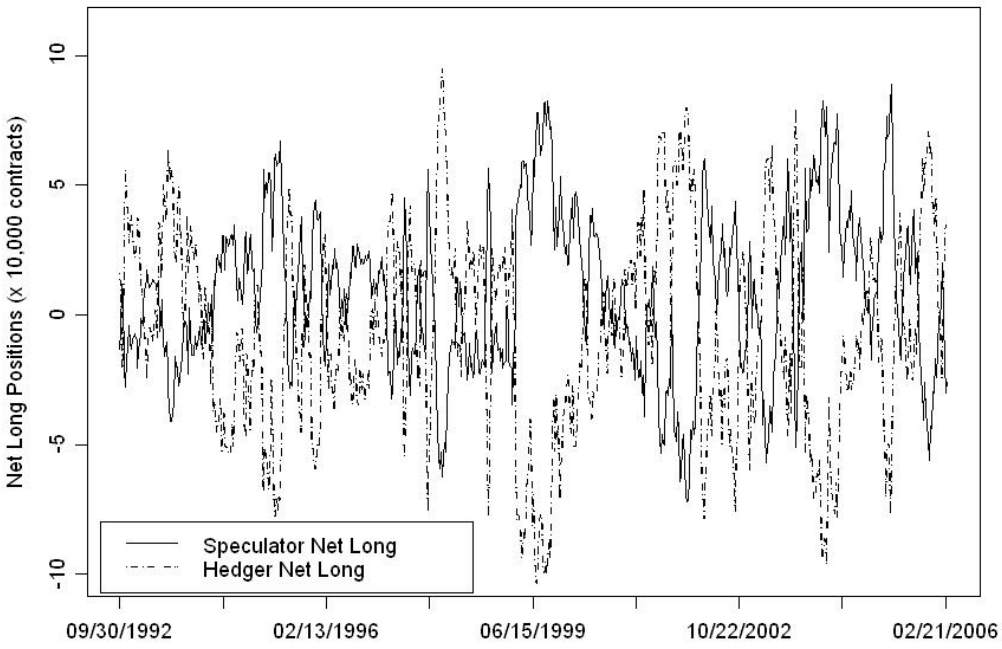


Figure 1.5: Net Long Positions of Large Hedgers and Speculators. The net long positions are the number of long positions minus the number of short positions within the trader class. Values are calculated from 698 weekly observations from September 30, 1992 until February 28, 2006, as reported in the CFTC Commitments of Traders Reports.

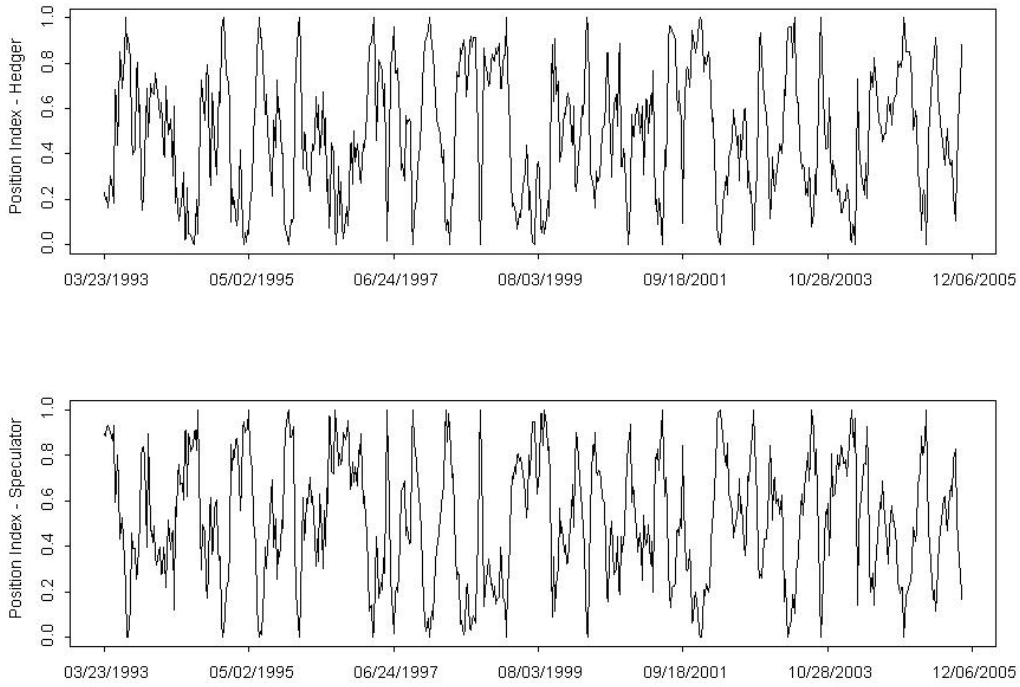


Figure 1.6: Trader Position Index, $P_t^C(\tau)$, for Hedgers and Speculators with $\tau = 39$ Weeks. The position index is a measure of how long a trader class relative to the maximum and minimum net long positions of that class over the centered window of length τ weeks. A value of 1 indicates that a new τ-period net long maximum has been reached. A value of 0 indicates that a new τ-period net long minimum (a net short maximum) has been reached. The upper panel is the position index for large hedgers, and the lower panel is the position index for large speculators. Values are calculated from 698 weekly observations from September 30, 1992 until February 28, 2006, as reported in the CFTC Commitments of Traders Reports.

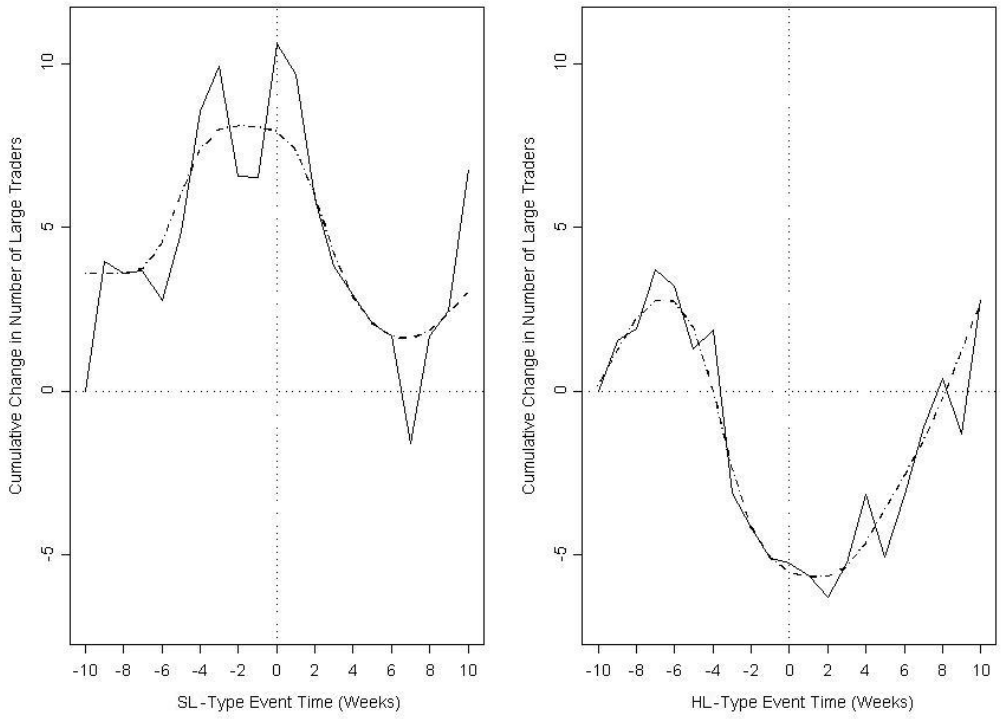


Figure 1.7: Averaged Cumulative Change in the Number of Large Traders Around Dates when Speculators and Hedgers Have Extreme and Opposing Positions. SL-type events are when speculators are extremely long and hedgers extremely short. HL-type events are when speculators are extremely short and hedgers extremely long. The dashed line is a smoothed version of the series, computed by using a method of running medians known as 4(3RSR)2H with twicing. Times relative to the event date, measured in weeks, are on the horizontal axis from 10 weeks before traders take an extreme position until 10 weeks after. The left panel shows that traders are entering the market before an SL-type event, and then leave the market afterwards. The right panel shows that traders are leaving the market before and slightly after an HL-type event, and then begin entering again. The statistics are calculated from 698 weekly observations from September 30, 1992 until February 28, 2006. The number of traders in the market is reported in the CFTC Commitments of Traders Reports.

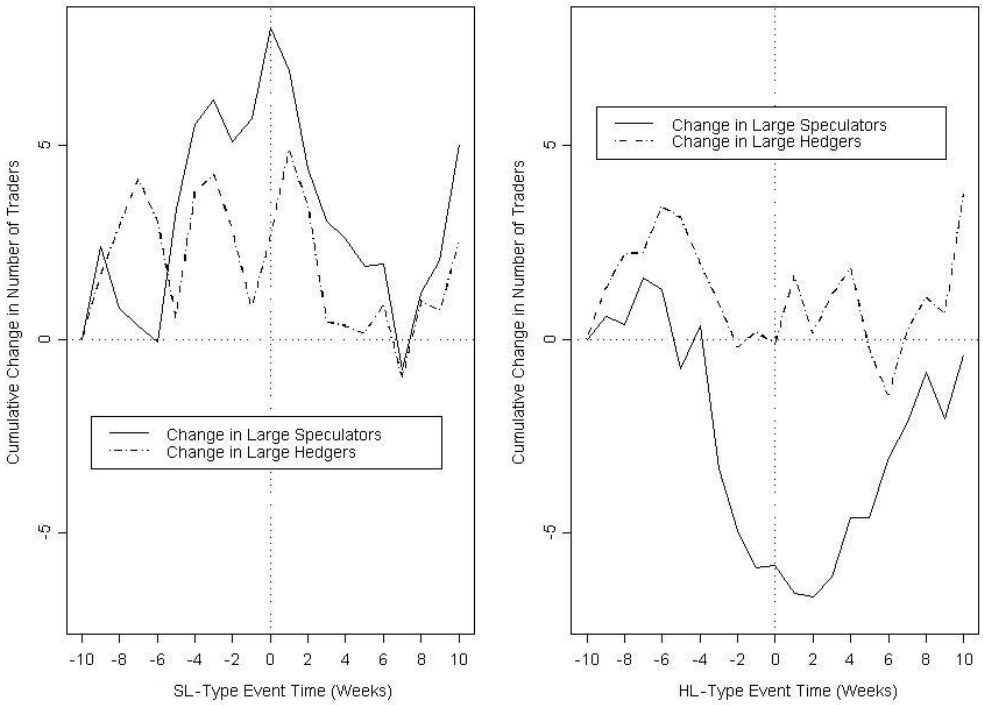


Figure 1.8: Averaged Cumulative Change in the Number of Speculators and Hedgers Around Dates when Speculators and Hedgers have Extreme and Opposing Positions. SL-type events are when speculators are extremely long and hedgers extremely short. HL-type events are when speculators are extremely short and hedgers extremely long. The solid and dashed lines represent the cumulative change in the number of speculators and hedgers, respectively. Times relative to the event date, measured in weeks, are on the horizontal axis from 10 weeks before traders take an extreme position until 10 weeks after. The left panel shows that traders are entering the market before an SL-type event, and then leave the market afterwards. The right panel shows that traders are leaving the market before and slightly after an HL-type event, and then begin entering again. The statistics are calculated from 698 weekly observations from September 30, 1992 until February 28, 2006. The number of traders in the market is reported in the CFTC Commitments of Traders Reports.

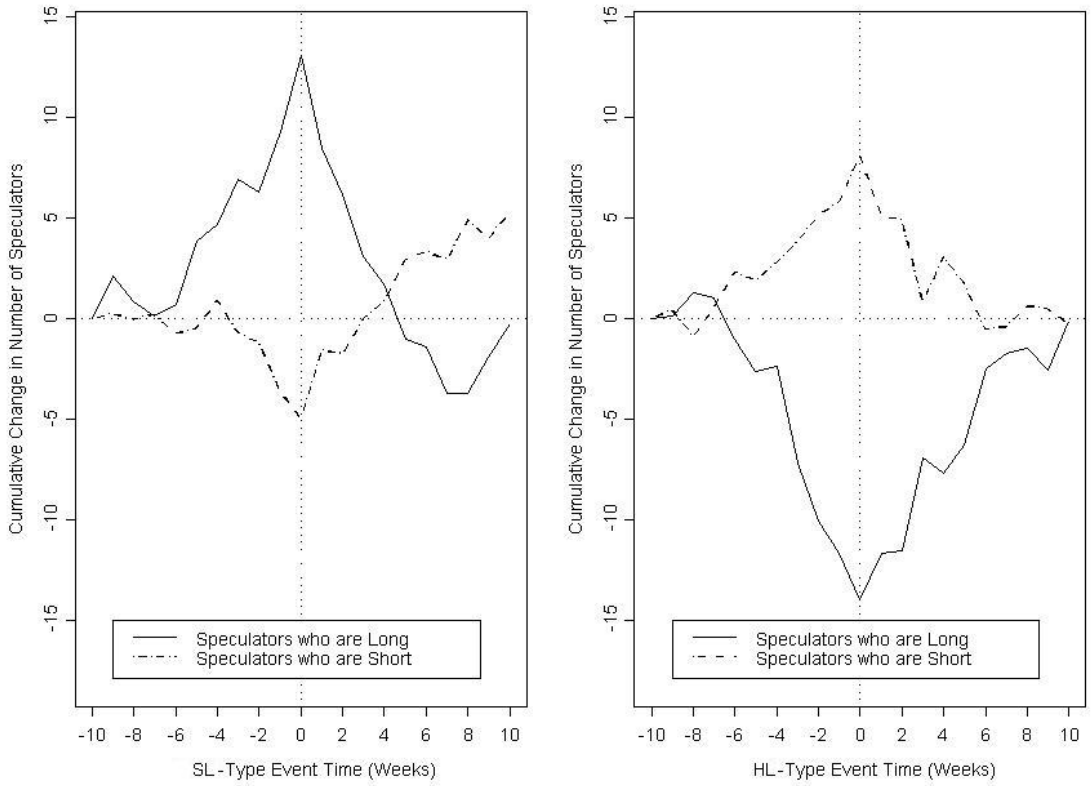


Figure 1.9: Averaged Cumulative Change in the Number of Speculators who are Long Versus who are Short Around Dates when Speculators and Hedgers have Extreme and Opposing Positions. SL-type events are when speculators are extremely long and hedgers extremely short. HL-type events are when speculators are extremely short and hedgers extremely long. The solid and dashed lines represent the cumulative change in the number of speculators long and short, respectively. Times relative to the event date, measured in weeks, are on the horizontal axis from 10 weeks before traders take an extreme position until 10 weeks after. The dotted vertical line indicates the event time. The averaged cumulative change in the total number of large hedgers. The statistics are calculated from 698 weekly observations from September 30, 1992 until February 28, 2006. The number of traders in the market is reported in the CFTC Commitments of Traders Reports.

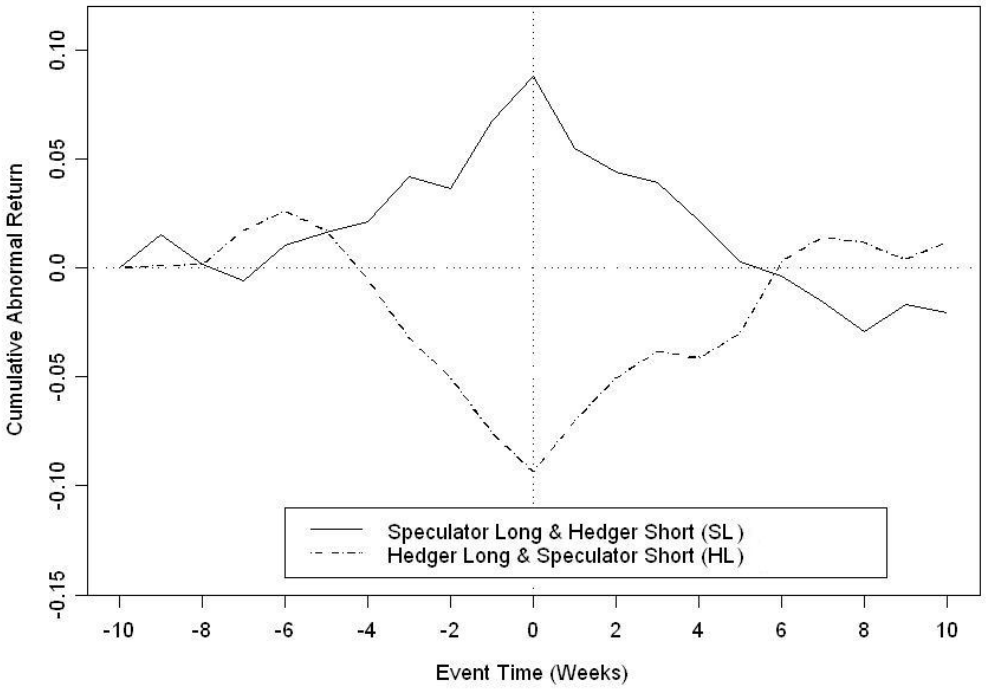


Figure 1.10: Averaged Cumulative Abnormal Returns. Returns are assumed to have a constant mean. Time relative to the event date, measured in weeks, are on the horizontal axis. Events at time zero (indicated by a vertical dotted line) are extreme opposing positions by large hedgers and speculators. When speculators are long and hedgers are short (solid line), the event date coincides with a local price top. When speculators are short and hedgers long, the event date coincides with a local price bottom. Statistics are calculated from 698 weekly observations from September 30, 1992 until February 28, 2006. Within the sample period, there are 20 extreme events of each type for a total of 40 extreme events. Trader positions are reported in the CFTC Commitments of Traders Reports and the weekly returns are calculated from tick-level price data from TickData, www.tickdata.com.

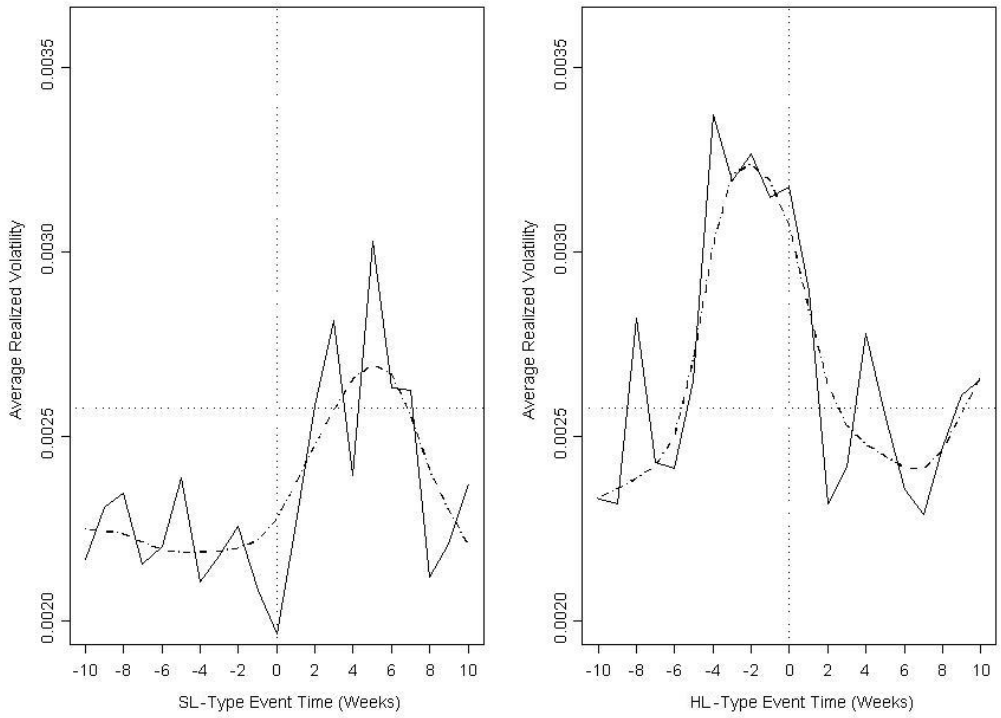
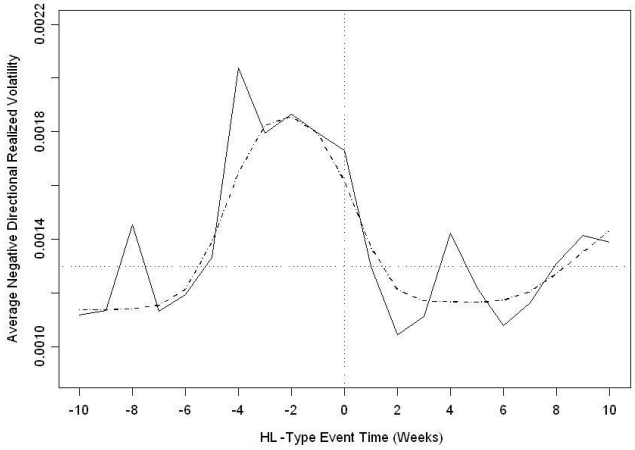
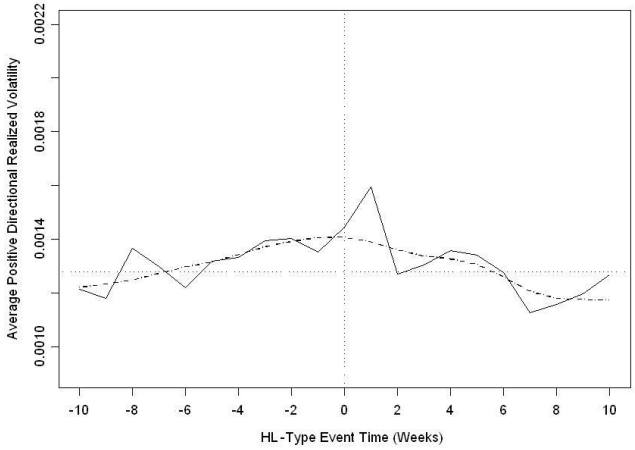


Figure 1.11: Average Realized Volatility Around Type SL-Type Event Dates and Type HL-Type Event Dates. SL-type events are when speculators are extremely long and hedgers extremely short. HL-type events are when speculators are extremely short and hedgers extremely long. The dashed lines are a smoothed versions of the average realized volatility computed by using a method of running medians known as 4(3RSR)2H with twicing. The dotted horizontal line is the average realized volatility over the entire sample. Times relative to the event date, measured in weeks, are on the horizontal axis from 10 weeks before traders take an extreme position until 10 weeks after. Statistics are calculated from 698 weekly observations from September 30, 1992 until February 28, 2006. Trader positions are reported in the CFTC Commitments of Traders Reports and the weekly returns are calculated from tick-level price data from TickData, www.tickdata.com.



(a)



(b)

Figure 1.12: Average Directional Realized Volatility Around Type HL-Type Event Dates. HL-type events are when speculators are extremely short and hedgers extremely long. Panel (a) plots the average negative directional realized volatility, and panel (b) plots the positive directional realized volatility. The dashed lines are a smoothed versions of the average realized volatility computed by using a method of running medians known as 4(3RSR)2H with twicing. The dotted horizontal lines in the respective panels are the average negative and positive directional realized volatility over the entire sample. Times relative to the event date, measured in weeks, are on the horizontal axis from 10 weeks before traders take an extreme position until 10 weeks after. Note that the spike in panel (a) at time -1 week is spurious and is caused by a single large value. Statistics are calculated from 698 weekly observations from September 30, 1992 until February 28, 2006. Trader positions are reported in the CFTC Commitments of Traders Reports and the weekly returns are calculated from tick-level price data from TickData, www.tickdata.com.

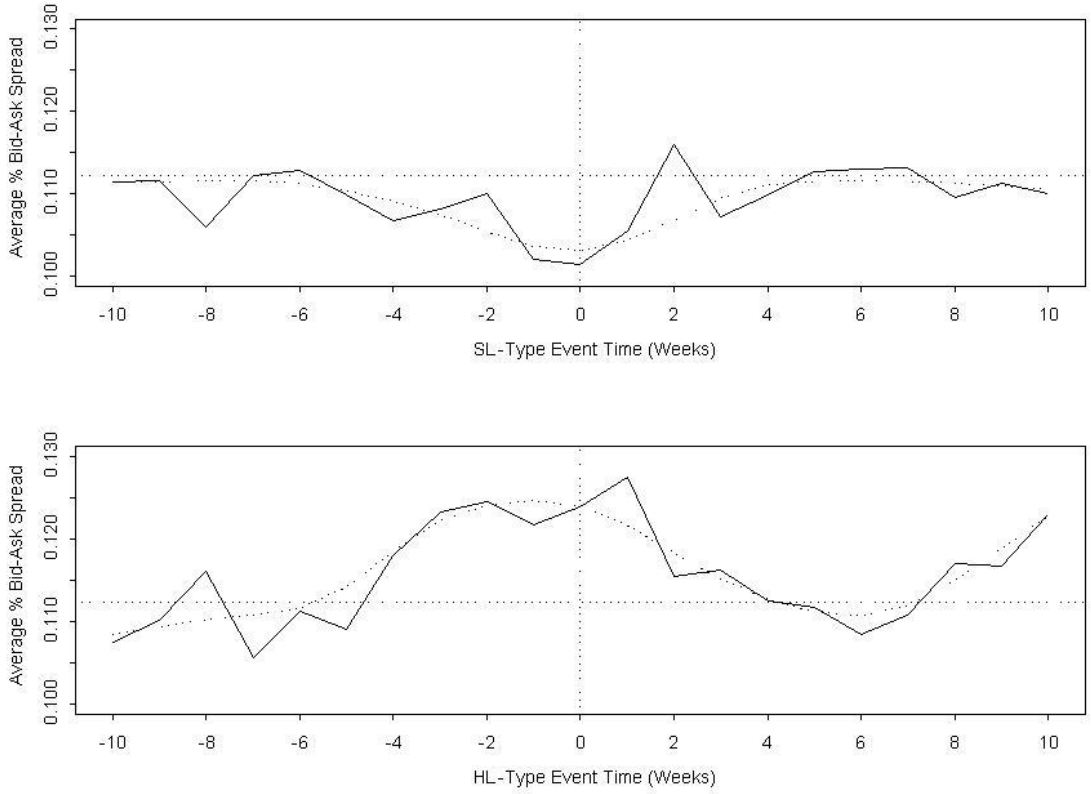


Figure 1.13: Averaged Daily Percentage Bid-Ask Spread Around Event Dates. Spreads were estimated from transaction price data. Time relative to the event date, measured in weeks, are on the horizontal axis. Events at time zero (indicated by a vertical dotted line) are extreme opposing positions by large hedgers and speculators. The dotted horizontal line indicates the average daily percentage bid-ask spread from all observations. The dashed lines are a smoothed versions of the average daily percentage bid-ask spread computed by using a method of running medians known as 4(3RSR)2H with twicing. Statistics are calculated from 698 weekly observations from September 30, 1992 until February 28, 2006. Trader positions are reported in the CFTC Commitments of Traders Reports and the weekly returns are calculated from tick-level price data from TickData, www.tickdata.com.

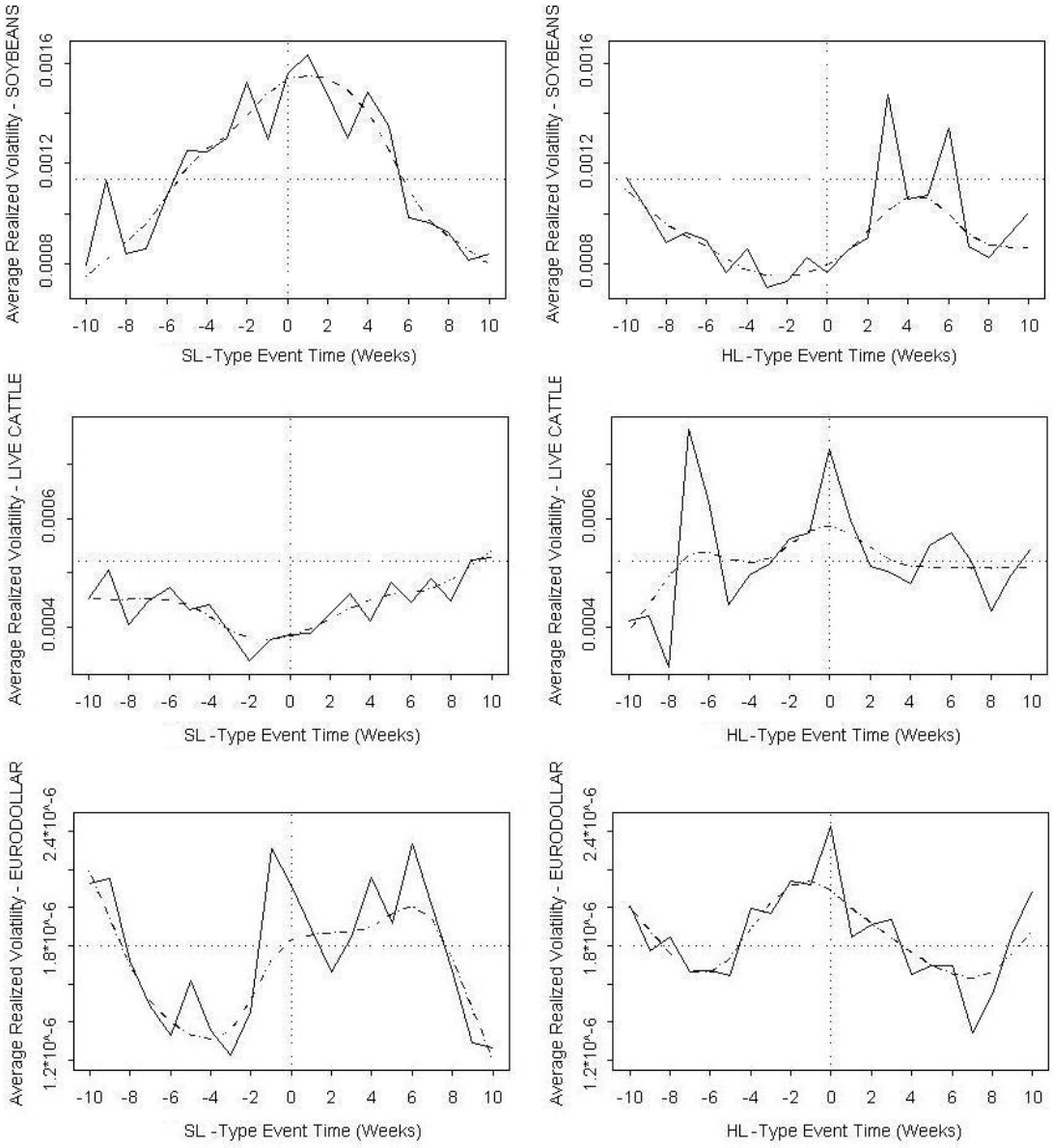


Figure 1.14: Average Realized Volatility Around Type SL-Type Event Dates and Type HL-Type Event Dates for Three Markets. SL-type events are when speculators are extremely long and hedgers extremely short. HL-type events are when speculators are extremely short and hedgers extremely long. The dashed lines are a smoothed versions of the average realized volatility computed by using a method of running medians known as 4(3RSR)2H with twicing. The dotted horizontal line is the average realized volatility over the entire sample. Times relative to the event date, measured in weeks, are on the horizontal axis from 10 weeks before traders take an extreme position until 10 weeks after. Statistics are calculated from weekly observations from September 30, 1992 until February 28, 2006. Trader positions are reported in the CFTC Commitments of Traders Reports and the weekly returns are calculated from tick-level price data from TickData, www.tickdata.com.

Chapter 2

The Limits of Anonymity: Identifying Large Block Trades in Real-Time

2.1 Introduction

Financial markets are not perfectly liquid trading environments, and as a result, trading is costly. An important component of the cost of trading is referred to as market impact - the impact of the traders own actions on the market price. For a small trader, the market impact may only be the movement between the best bid and ask prices, but for a trader needing to buy or sell a large number of shares (referred to as a block) the market impact can be the most important component of the cost of trading [5]. Because of this practical importance, there has been considerable academic and industrial research into understanding market impact and how to trade optimally in its presence. This paper adds to this literature by addressing the question: can market participants detect the market impact of a block trader and infer his or her presence even when transactions are anonymously matched? Under certain size and market characteristics, the answer to this question is yes, and this has strategic implications for optimal trading which have thus far not been considered in the published literature.

The market impact of a block trade is a function of the size of the trade (or the rate of trading) and various market attributes. If a large block trade were to be executed very

quickly, it would 'consume' liquidity faster than it was replaced and prices would have to move in favor of potential counterparties, and against the block trader. A vivid example of a 'liquidity shock' caused by a trade too large for the market to accommodate was the 'flash crash' of May 6th, 2010, where, within only minutes, the Dow Jones stock market index experienced and recovered from its biggest intraday point decline in its history. The SEC and CFTC found this crash to have originated from a single algorithmic trader selling $75,000$ E-Mini stock index futures contracts too quickly and without regard for market impact.

To avoid causing excessive market impact, block traders typically spread their trades over time to allow liquidity to be replenished at a rate close to its rate of consumption from the block trade executions. Slow execution, however, is also costly since it increases the uncertainty of the average execution price for the block. The block trader must decide on the optimal sequence of trades by trading off his concern for market impact and price uncertainty.

For an illustrative and motivating example, we can look to the French stock market in January of 2008. Since exchange-based trading is anonymous, identifying block traders typically requires access to private brokerage data (e.g. see [5]). However, on January 19th of 2008, Societe Generale discovered that one of their traders, Jerome Kerviel, had accumulated massive unauthorized equity positions worth 49.9 billion Euros, an amount that was larger than the bank's total market capitalization. Over the three days of January 21 - 23, Societe Generale sold off Kerviel's positions thereby giving us an excellent example of a large block trade in action. Public disclosure of the block trade was made on January 24th, the day after the trade was completed. One of main components of the French CAC 40 index is Total SA, a large French multinational oil company. Since this company was almost surely a part of Kerviel's portfolio, we can examine its stock for characteristics of a large block trade during that period.

[FIGURE 1 ABOUT HERE]

Figure 1 shows the behavior of Total SA's stock between January 15th and the 25th, and indicates Societe General's sell-off period. This figure demonstrates several common features of block trades. First, in the top panel, we see a fairly steady fall in the stock price during the block sale period. This is consistent with market impact from a large sale causing downward price pressure. The bank's liquidation of Kerviel's portfolio reportedly cost 4.9

billion Euros. Second, in the middle panel, we see that the filtered absolute hourly price range jumps to a higher level during the block trade and then reverts back to it's earlier level. This is consistent with higher volatility during a block trade due to wider spreads and order imbalances. Third, in the bottom panel, we see that the deseasonalized hourly volume is consistently above typical volumes throughout the block sale, consistent with the block trader adding to liquidity demand. Finally, we see that these effects in price, volatility, and volume all revert, or decay, after the block trade is completed.

Given that these characteristics develop and persist in real-time during block trades, this paper investigates whether another market participant with no insider information about block trades (i.e. only using public market data) can accurately infer when a block trade is actively present in the market. To formally test this hypothesis, a state-switching methodology is used to identify times when a block trader is present, and its performance on simulated data is measured using precision and recall measures from the classification literature. The answer to this question of whether such methodologies work is, in certain environments, yes, and so a second question is asked: Could this market participant who is capable of inferring block trading strategically exploit this information to extract profit? This hypothesis is tested using a simple trading strategy based on the classifier's output. Such a capacity to earn profits from the market impact of block trading will impose additional costs to block traders, and would make optimal block trading a strategic interaction.

This investigation proceeds by first developing a formal framework for simulating a market environment with block trading. Within this environment, a Markov state switching model is proposed as an option for identifying the times when a block trader is present. This choice was inspired by the work of James Hamilton [36] on business cycles. This model is then implemented on the simulated data over various block trading parameters to determine under what conditions block traders can be identified using such approaches. A simple heuristic is introduced to distinguish block purchases from block sales, and a trading strategy based on this classification system is then tested. The paper ends with a discussion of the results and their implications for trader anonymity and optimal trading. Throughout the paper, the relevant literature streams are reviewed to provide context and additional information.

2.2 The Simulation Environment

While the Societe Generale block trade lasted for three days because of its size, most block trades have a much shorter duration. In the empirical study by [5], their sample of block trades of at least 0.25% of average daily volume took an average of 0.39 days to complete (minimum = 0.00, median = 0.32, and max = 1.01). With core trading hours between 9:30am and 4:00pm on the New York Stock Exchange, most of these block trades were completed in just over two hours[1]. Although many of these trades would be too small to interest a strategic trader, these sample statistics demonstrate the need for very high-frequency sampling in order to generate useful estimates of important quantities without depending on too long of a temporal history. A long temporal history will bias estimates shortly after the entry or exit of a block trader since true values will jump to a new level but estimates will be based on data generated before the jump.

For this simulation study, high frequency data will be generated from a model of prices with a market impact component. A very general market impact model is presented by [32] and is adapted for the current study. The stock price P_t at time t is given by

$$P_t = P_0 + \int_0^t f(\dot{x}_s)G(t-s)ds + \int_0^t \sigma\sqrt{P_t}dZ_s$$

where $\dot{x}_s$ is the rate of trading ($\dot{x}_s > 0$ for a purchase and $\dot{x}_s < 0$ for a sale) used by the block trader at time $s < t$, $f(\dot{x}_s)$ is the instantaneous market impact of this trading, $G(t-s)$ is a decay factor, and Z_s is a standard Brownian motion. Notice that as long as $f(0) = 0$ the market impact component only plays a role when the block trader is actively present in the market. The other component of the price is a random walk component with volatility $\sigma\sqrt{P_t}$ that increases and decreases with P_t.

Of course, trading is a discrete process, and this continuous time process can be viewed as the limit of the following discrete time process:

$$P_t = P_0 + \sum_{i<t} f(\delta x_i)G(t-i) + \varepsilon_t$$

where δx_i is the amount of stock traded by the block trader in the ith time interval, f is the

[1]While most block trades appear to be completed intraday, some large block trades involve thousands of individual transactions and can take weeks to complete [20, 71]

discrete market impact function (with apologies for the abuse of notation), G is the decay factor, and ε_t is a zero-mean i.i.d. Gaussian innovation with variance $\sigma^2 P_t$.

This general framework encompasses several specific functional forms for the market impact and decay functions that have been proposed in the literature. For example, [5] define the temporary market impact as $f(v) = \eta\sigma v^\beta$ where v is a measure of the rate of trading, σ is volatility, and where β and η are estimated empirically ($\beta \approx 0.6$ and $\eta \approx 0.14$). In their model, temporary market impact decays instantaneously and so G is simply an indicator function with the value 1 when $t - s = 0$ and is equal to 0 otherwise. This temporary component of market impact would not leave any opportunity for a strategic trader, however the authors also include a permanent market impact component that does not decay at all.

Another specific model of market impact that can be captured by this general framework includes that by [63]. They define market impact to be proportional to the trading rate, $f(v) = \alpha v$, and to decay exponentially, $G(t - s) = e^{-\rho(t-s)}$. The model of [14] also fits into this framework with a market impact that is concave in the trading rate, $f(v) = \alpha \log(v)$, and a power-law decay function, $G(t - s) = \alpha(l_0 + t - s)^{-\gamma}$. [32] points out that the market impact function, f, and the decay function, G, should be related to avoid statistical arbitrage in 'round trip' trading. That is, we should not unconditionally expect to earn positive profits simply by buying and selling an equal number of shares, regardless of the rates at which we trade.

For this study the popular power-law functional form is used for both the market impact and decay functions. There has long been empirical evidence that the market impact of small transaction volumes (relative to the overall volume of the particular market) is a concave function of the trading rate [42, 46, 53]. The square-root and logarithmic functions have been used, but a general market impact function that can capture differences across markets and aggregation levels is the following power-law formulation [14]:

$$f(v) = \frac{\alpha v^\psi}{\lambda}$$

where λ is a liquidity parameter which can be dependant on the market capitalization of the stock. The exponent ψ can also be a function of the rate of trade, but for aggregated transaction data, ψ is commonly found to be about 0.5 for various markets and aggregation levels [70, 49, 31]. [5] estimates the exponent for temporary market impact to be $\psi = 0.6$

with a coefficient $\alpha = 0.142$ and since they use U.S. data, as this study uses, their estimates are used in the following simulations.

The power-law formulation of market impact is a simplification of the true function which is dependent on the available liquidity in the order book and the particular way that liquidity is replenished as it is consumed. The power-law function will clearly fail to describe the market impact of very large orders for which market impact should become convex in the trading rate. The market impact would theoretically be unbounded if a market order was placed for more that the total liquidity being offered. This aspect of the market impact function is not considered in this study.

As with the market impact function, the decay function used in this study has a power-law specification.

$$G(t-s) = (t-s)^{-\gamma}$$

This specification was argued for by [14] and they estimate that $\gamma = 0.4$. Given the chosen specification of market impact, the choice of decay function is more restricted in order to avoid building statistical arbitrage opportunities into the simulation environment. [32] provides a lemma describing the 'dynamic-no-arbitrage' condition for small trading rates under power-law market impact and decay functions with exponents ψ and $-\gamma$, respectively. This condition requires that $\alpha + \gamma \geq 1$. The estimates of [5] and [14] of $\psi = 0.6$ and $\gamma = 0.4$ satisfy this condition and are used in this study.

Given the cost of trading, captured by the market impact function, and the uncertainty of the transaction price of future orders, captured by the volatility of the market, a block trader has two problems to solve: (1) how to schedule their trading over time, and (2) what type of trades to use. The second question asks whether the trader will use market or limit orders, and if limit orders, what limit price should be used. This question of order placement has been studied theoretically by [8] and [40] among others, and has been empirically studied by [41], [60] and [43]. The simple market impact model used in this study does not differentiate order types, and so this component of block trading is ignored in the simulation.

While we can ignore order placement, the trading schedule of the block trader must be specified in order to complete the simulations. Optimal block trading has been extensively studied, with one of the classic references being [11]. More recent and sophisticated

approaches have been developed by [3], [1], [4], [2] and [13]. The general approach is to trade off market risk (i.e. the risk of prices changing before the block trade is complete) and market impact (i.e. the implicit cost of trading quickly) by minimizing a mean-variance criteria. Usually, this will result in a convex trading trajectory (the broken line in Figure 2) where the trade rate is initially high, but decelerates throughout the trade. To simplify the simulation environment, this paper uses a simple constant rate trading schedule (the solid line in Figure 2) throughout the block trades.

[FIGURE 2 ABOUT HERE]

Given the above discussion, we can now summarize the simulation environment used in this study. The discrete price process used is:

$$P_t = P_0 + \sum_{i<t} sign(\delta x_i)\alpha\sigma \left(\frac{|\delta x_i|}{\overline{V}_i}\right)^{\psi} \frac{1}{(t-i)^{\gamma}} + \varepsilon_t$$

with market impact parameters $\alpha = 0.142$ and $\psi = 0.6$, market impact decay parameter $\gamma = 0.4$, and where ε_t is a zero-mean i.i.d. Gaussian innovation with variance $\sigma^2 P_t$. Artificial days have 1000 periods and the daily volatility is set to 0.01. The block trade is carried out at a constant rate δx_t per period and the expected volume per period is constant at $\overline{V}_t$. The trade rate and expected volume are constant, but are subscripted with t to remain general enough for future use with real market data. Since only a single asset is being simulated, no liquidity factors, λ, are included in the model.

[FIGURE 3 ABOUT HERE]

Figure 3 shows an example of this simulation. In this example, a block purchase representing 10% of the typical trade rate occurs between periods 400 and 600. The top panel shows the stock price between periods 0 and 1000 under the block trade (solid line) and under the counterfactual where the block trade never occurred. The middle panel shows the extracted cumulative market impact where we see the monotonically rising market impact component of the price during the block trade, followed by a period of decay after the block trade is complete. The bottom panel shows the demeaned trade volume throughout this period, with a slightly higher trade rate during the block trade. This simulation implicitly

assumes that there are no sizeable strategic traders identifying the block trader and 'piggy-backing' their trades. The simulation also does not explicitly model any order-flow outside of the block trade.

In order to train and test classification schemes, a sequence of such simulations will be constructed. Simulations and testing will be done under various block trading rates and durations to determine how these factors influence the performance of the identification algorithms.

The algorithms will take as input, variables that seem to be useful in distinguishing block-trading episodes from non-block-trading episodes. From looking at the Total SA example (Figure 1) and the simulation example (Figure 3), volume is the clear candidate since it jumps to a new level during a block trade. Volatility could also be indicative of a block trade. Returns will be important for distinguishing purchases from sales, but this is likely to be the most difficult component of the exercise since return predictability is notoriously difficult. The specific use of these variables will be further described in the following sections where the identification algorithms are described in detail.

2.3 A Markov State Switching Model

Whether or not a strategic trader can identify the presence of a large block trader hinges on how block trading impacts the market. The hypothesis being investigated is that markets behave in measurably different ways when a block is being traded compared to when no blocks are being traded. In the simulation environment, this difference has been built into the environment by adding a market impact component to the price model. So we know the difference is there, so the question is really whether we can detect it in real-time. To formalize this idea, we can think of the market as being in one of two (or more) states: (1) a block trader is present, and (2) no block trader is present. The market has different dynamics in each of these states, but the difficulty is that we cannot directly observe which state the market is in using public information. As a result, the state of the market, as well as the state-dependent parameters, must be inferred from the observable data.

In the econometric literature, one of the first studies to model two separate regimes for a single economic time series was by [64]. The use of an unobserved Markov process to model regime changes was introduced by [34]. [36] popularized the use of discrete Markov processes when he proposed a simple Markov state switching autoregressive model, which

he subsequently generalized [37] and used to model U.S. real Gross National Product. In his study, one state corresponded to periods of economic expansion, and the other corresponded to periods of recession. It is this type of Markov model that is used in this section to model financial time series in the presence and absence of large block traders.

There is substantial heterogeneity in the size, duration, and strategy of large block trades, so to assign a small number of states to capture the differences induced by block traders requires a degree of simplification. There are several choices for the number of states, including:

- Two States: A block trader are either present or absent.
- Three States: A block traders is buying, or selling, or absent.
- Five States: A block trader is buying, or is selling, or the market impact of a prior block purchase is decaying, or the market impact of a prior block sale is decaying, or there is no substantial influence of block trading in the market.
- K States: Block traders of multiple size categories may be influencing the market.

This section develops the most basic model with only two states in order to generate a base case for block trading identification that can be extended by future research. This allows us to measure the potential identification of block trading, but not to distinguish between block purchases and sales. Since a strategic trader will need to know whether a block trader is buying or selling, a simple heuristic is added to the the Markov model. A further simplification comes in the Markov model by using only the volume series to identify block trading, and using the return series after a block trade has been identified to distinguish purchases from sales. The choice to begin model development using volume data rather than return data is made because changes in volume are a stronger and less noisy signal. Both of these components are described below.

A discrete state Markov process (often called a Markov chain) classifies the state of the world, S_t, (often called regimes) at any time t as one of a set of discrete types. In the current model there are two possible states,

$$S_t \in \{1, 2\}.$$

One state, $S_t = 0$, corresponds to normal market conditions when there are no block traders present, and the other, $S_t = 1$, corresponds to market conditions when a block trader is

present. These different states are associated with different dynamics which are modeled as having different parameter values or functional relationships. Transitions from one state to another is governed by the Markov property (i.e. that the next state depends only on the current state and not on the past). State transition probabilities are given by

$$\begin{aligned}
&\Pr(S_t = 0|S_{t-1} = 0) = P_{00} \geq 0 \\
&\Pr(S_t = 0|S_{t-1} = 1) = P_{01} \geq 0 \\
&\Pr(S_t = 1|S_{t-1} = 0) = P_{10} \geq 0 \\
&\Pr(S_t = 1|S_{t-1} = 1) = P_{11} \geq 0
\end{aligned}$$

with the usual probabilities requirement that

$$\sum_{j=1}^{2} \Pr(S_t = j|S_{t-1} = i) = 1 \tag{2.1}$$

for $i = 1, 2$. These probabilities are usually collected in a state transition matrix,

$$P = \begin{bmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{bmatrix}$$

Within this two-state environment, volume (V_t) is modeled using a state-dependent intercept model:

$$\begin{aligned}
&V_t = \alpha_{S_t} + \varepsilon_t, \quad t = 1, 2, \ldots, T, \\
&\varepsilon_t \sim N(0, \sigma^2_{S_t}), \\
&\alpha_{S_t} = \alpha_0(1 - S_t) + \alpha_1 S_t, \\
&\sigma^2_{S_t} = \sigma^2_0(1 - S_t) + \sigma^2_1 S_t, \\
&S_t = 0 \text{ or } 1.
\end{aligned}$$

Volume was transformed by the natural logarithm for the estimation to avoid having to constrain it to be greater than zero and to improve the symmetry of the distribution. Other constrained parameters were transformed with monotonic functions for similar reasons. Since each $\sigma^2_{S_t}$ is constrained to be within $(0, \text{inf})$, we can estimate $y = \log \sigma^2_{S_t}$ and then get our wanted parameter back using the transformation $\sigma^2_{S_t} = e^y$. Similarly, the transition probabilities are constrained to be in $(0, 1)$, so the unconstrained transformation $y_{ij} = \log[P_{ij}/(1 - P_{ij})]$ was estimated, and our desired probabilities were restored using $P_{ij} = 1/(1 + e^{-y_{ij}})$. Finally, in order to ensure that $\alpha_0 \leq \alpha_1$, so that we know that state

1 is the block trading state, the equation $\alpha_1 = \alpha_0 + e^{\beta_1}$ was used and $\beta_1 = \log(\alpha_1 - \alpha_0)$ was estimated. In the results throughout the paper, all transformations have been reversed after the estimation and the results are reported in their intuitive form.

If the S_t, $t = 1, 2, \ldots, T$, were observable (so we knew a priori when the block trader was present), then the above would simply be a dummy variable model. Unfortunately for the strategic trader, the states are not directly observable. By assuming that the states follow a (first-order) Markov switching process, we can estimate the parameters of the model as well as the likelihood of each state $\Pr[S_t = i|\Psi_{t-1}]$ for $i = 0, 1$ and each time t.

The econometric details of this maximum likelihood estimation can be found in [55] and a discussion of its implementation can be found in [57] and [74]. The estimation process is iterative, and directly programming it in S-Plus or other interpreted languages can result in slow and inefficient computation. Fast optimization code written in C for this kind of iterative estimation has been included in the S-Plus package S+FinMetrics/SsfPack. This package can be adopted for Markov switching models, however the models must be formatted as state space models. A brief description of state space models and how to write the above model in state space form has been included in Appendix I at the end of this paper.

2.3.1 Estimation and Classification Results

Two behaviorial variables can impact the identifiability of a block trader: (1) how fast they trade relative to typical trade volume, and (2) how long they continue to trade at that rate (i.e. the duration of the block trade). In this section, these two factors are varied to see how easily the Markov-based identification scheme can find the block traders under varying conditions.

While classification performance will be described for all variable pairs, full estimation results are only given for a single case. This case will illustrate several stylized facts of the estimated models. This base case uses the simulated environment described in the previous section with block trades lasting 0.20 days each and have a high trade rate of 30% of normal trade volume per period. This long duration, high rate example should be easy for the model to capture. Following this example, trade durations and rates are lowered to find how classification performance falls with these factors.

[TABLE I ABOUT HERE]

Table I shows the estimation results of the two-state Markov switching model run on

simulated data for this base case. The estimated intercepts of 99 and 129 are very close to the true values of 100 and 130, respectively. [55] derives a simple equation to infer the duration of each state from the estimated transition probabilities. The average duration of a block trade is:

$$\frac{1}{1-P_{11}} = 189.3 \approx 0.19 days$$

An alternative measure of the persistence of the block trading state are the ergodic probabilities for a stationary Markov chain presented by [38]. Let $\pi_t = (P_0, P_1)$ be the probabilities of being in each state at time t. Then given the state transition matrix P, the probability of being in each state at time $t+1$ is $\pi_{t+1} = P'\pi_t$. The ergodic probabilities, π, are those that satisfy $\pi = P'\pi$. From the estimated transition probabilities,

$$\pi = (0.803, 0.197),$$

which are very close to the true unconditional probabilities of being in a normal market state (80%) and being in a block-trading state (20%).

Using a filtering algorithm developed by [54], SsfPack approximates the probabilities $\Pr(S_t = j|\Gamma_t)$ of being in a state conditional on all information available at time t, denoted Γ_t. These probabilities for the block trading state, $S_t = 1$, are shown in Figure 4. In this example, the block trades occur every day between the 400th and 600th period. A close up of the time around the transition period from normal trading to block trading at time 400 is shown in the bottom panel. Within only a few periods after the block trade begins, the probability of a block trade is near 100% indicating that a block trade of this size should be easy to spot.

[FIGURE 4 ABOUT HERE]

A simple probability threshold rule is used to classify times as being in state 0 or 1. If $\Pr(S_t = 1|\Gamma_t) > \rho^*$ then time t is classified as a time when a block trade is likely to be occurring (i.e. state 1), otherwise it is classified as a time without a block trade (i.e. state 0). For this study, $\rho^* = 0.50$ was chosen, however, depending on the costs of misclassifications, the threshold probability can be adjusted.

With this classification rule, we generate a sequence of predicted states which we can then compare with the true states using classification performance measures. Appendix II at the end of this paper provides a brief introduction of classification and provides formal definitions

of the performance measures that are used in this paper. To evaluate the binary classification performance of the Markov model with the threshold probability classification rule, the accuracy, precision, and recall measures are reported. Accuracy is a popular classification measure defined as the proportion of correct classifications. This measure has a severe problem for cases where the class of interest, like block trading times, is very infrequent since a useless rule of always predicting the frequent class will have a high accuracy, but be useless in practice.

To overcome this problem with accuracy, two other classification performance measures, precision and recall, have been developed and are reported in this paper. Precision, with respect to the block trading state, is the proportion of periods predicted to be block trading periods which truly belong to that class. So a classifier with a high precision will have a low proportion of 'false positives.' The other measure, recall, is the proportion of periods that are truly block trading periods which are classified as belonging to that class. So a classifier with a high recall will have a low proportion of 'false negatives.' These two performance measures, along with accuracy, are reported for 20 different pairs of block trading rates and durations in Table II.

[TABLE II ABOUT HERE]

Table II shows that as expected, accuracy results are consistently high, even when precision and recall measures are low. Also, as expected, precision and recall fall with decreases in either the block trading rate or duration. It is interesting that even with a very short duration of 0.02 days, a high trading rate easily identifies a block trader. However, even long duration block trades are difficult to identify if their trade rate is low enough.

It is possible that a strategic trader could take advantage of the knowledge that a block trader is trading without knowledge of whether the block is a purchase or a sale. This could be done through convergence or pairs trading strategies, or by trading volatility in the options market. However, the primary question of this paper is whether a strategic trader can directly take advantage of the market impact of a block trade, and for this it is important to know whether a block is being purchased or sold.

There are many options in constructing a heuristic to distinguish a purchase from a sale. Signed order flow could be estimated in order to determine whether the excess volume is likely from purchases or sales. Since the simulated data was of a market price without a bid-ask spread component, this strategy cannot be used here. Another option is to use recent

return data and rely on the cumulative market impact to reveal whether it is positive (for purchases) or negative (for sales). This option has its weaknesses for pure classification since, for example, a block purchase could be occurring during a falling market where negative price innovations are dominating positive market impact, thereby causing the strategic trader to classify the period as a block sale. This problem will become more pronounced for smaller block trades. Despite this drawback, this is the option chosen for this paper.

To implement this heuristic based on a window of past returns, the natural choice for the beginning of the window is the beginning of the block trade. Since the strategic trader only has an estimate of the start of the block trade, and the bottom panel of Figure 4 shows that this estimated start time of a block trade lags the true start by a few periods. To account for this lag and provide a stronger return signal, this study uses the price from five periods before the first block trading signal to calculate cumulative returns. If the cumulative return based on this price is positive for a time that is identified as having an active block trade, the rule for distinguishing block purchases from sales is to label the period a purchase. If the cumulative return is negative, then the period is labeled a sale. In order to avoid resetting the reference price too often when a time in the middle of a block trade is misclassified (see, for example, the middle panel of Figure 4), the probabilities of states are filtered with an exponential moving average for the purposes of finding the first time of the block trade.

This signal weakens over time since the price impact is decaying. Figure 5 shows the decaying cumulative market impact for a 30% and a 10% block trade. The block purchase in this figure begins at time zero, and time at time t the cumulative market impact is calculated as:

$$MI_t(\frac{\delta x_i}{V_i}) = \sum_{i<t} sign(\delta x_i)\alpha\sigma \left(\frac{|\delta x_i|}{V_i}\right)^{\psi} \frac{1}{(t-i)^{\gamma}}$$

Although the price process used in the simulation has price dependent volatility, for short horizons without any block trades, the process can be approximated by a Wiener process for which variance increases linearly with the length of the time interval. That is,

$$Var(w_t - w_0) = \sum_{i=1}^{T} \Delta t = T\Delta t = t$$

for a Wiener process $\{w_t\}$. This means that the percentiles for price innovations scale linearly with time. The straight broken lines in Figure 5 show the 70th and 90th percentiles for the cumulative price innovations. We can see that as time passes, the variance of the

underlying price process eventually dominates the decaying market impact.

[FIGURE 5 ABOUT HERE]

Used in conjunction with the Markov model and binary classification rule described above, this heuristic for distinguishing purchases from sales provides a three-state classifier. The accuracy, precision and recall performance of this classifier is reported in Table III. Since purchases and sales are equally present in the simulation sample, the reported statistics are averaged with equal weights for purchases and sales.

[TABLE III ABOUT HERE]

The three-state classification problem is more difficult than the binary case since it involves identifying temporary trends in non-stationary prices. As a result, the classification performance measures are worse compared to those in Table II. When the trade rate is high, classification performance is still good even for short duration block trades. However, since the market impact impact of low trading rates is a smaller component of price changes relative to the exogenous price innovations, performance falls quickly as trading rate decreases.

In addition to classification performance, Table III lists trading performance from a long/short trading strategy based on the three-state classifications. If a period was classified as a block-purchase time, then a long position was taken in the next period. Similarly, if a period was classified as a block-sale time, then a short position was taken in the next period. When no block trading was predicted, then no position was taken in the following period. Figure 6 shows the cumulative return for periods when a long or short position was taken for the long-duration (0.20 days), high-trade-rate (30% of typical volume) case. This strategy does fairly well compared with the passive case of holding a long-only position for the same periods. No trading costs were accounted for in these calculations.

[FIGURE 6 ABOUT HERE]

In order to compare the trading performance across parameter pairs, a Sharpe multiplier is listed in Table III. The Sharpe multiplier is defined as:

$$SM(\{r_t^{TS}\}) = \frac{SR(\{r_t^{TS}\})}{|SR(\{r_t^{Long}\})|}$$

where $\{r_t^{TS}\}$ is the sequence of returns from the trading strategy, and $\{r_t^{Long}\}$ is the sequence of returns from the long-only strategy for the same times, and $SR(\{r_t\}) = E[r_t]/\sigma_{r_t}$ is the Sharpe ratio. This measure is a conservative measure of the performance of the trading strategy since it compares it against the best of either the pure long-only or pure short-only strategies.

Interestingly, the trading strategy does not do the best for the longest duration and highest trading rate parameter pair, even though this case produces the best buy/sell precision and recall statistics. This is due to the fact that the measure is constructed for periods when the strategy takes a position, and shorter duration block trades have a higher average cumulative market impact since there is less time for decay to occur. However, the shortest duration and highest trading rate parameter pair does not produce good trading results either. This is because of the sharp decay of cumulative market impact after the block trade ends, and for short duration block trades, the classifier spends a longer portion of it's invested time producing false positive signals after the block trade has ended.

2.4 Limits of Anonymity

The results of the previous section indicate that it is indeed possible for a strategic trader to identify the presence of a block trader in the market. Of course, these results are for the simplified simulation environment and require further research to test their validity in real markets. Nevertheless, the results do shed light on the limits of anonymity in financial markets. Block traders lose their anonymity by trading in large volumes for extended periods.

Block traders may not be concerned about their loss of anonymity unless they face higher trading costs as a result. This paper introduced a hypothetical strategic trader who was on the look out for block traders with the hope of 'riding' the market impact of their block trades. Of course, if this strategic trader was also trading in significant volumes, he would himself have market impact which would likely increase the cost to the block trader. Looking back at the motivating example of the block trade of Societe Generale (Figure 1), while their block trade ended at the close of January 23rd, the trade volume for January 24th was also high, indicating that the activity of Societe Generale may have induced others to trade who were then exiting their position after they realized that the block trade was finished.

The ability to identify a block trader can be profitable for a strategic trader, as we saw in Figure 6. However, classifiers with lower precision and recall for identifying block purchases and sales can generate trading strategies that perform quite poorly. Low precision and recall performance have quite different impacts on a trading strategy. A low recall classifier will have a high proportion of 'false negatives', that is, it will predict no block trader when there is in fact one present. These false negatives will be most prevalent at the beginning of a block trade where the classifier must use lagged data that is from before the block trade. On the other hand, a low precision classifier will have a high proportion of 'false positives' where it predicts a block trade when there is not one occurring. This is likely to be prevalent after a block trade has just ended, and the classifier again lags the true state.

To understand the impact of these classification lags, and the importance of precision and recall, we can imagine a classifier that has an equal number of false negatives (FNs) and false positives (FPs), and where all of the false negatives occur at the beginning of a block trade in the form of a classification lag, and all of the false positives occur just after the block trade has entered. This situation is illustrated in Figure 7.

[FIGURE 7 ABOUT HERE]

The true and predicted probabilities of block trading are displayed in the top panel, and the bottom panel shows how much of the market impact a strategic trader can capture using the lagging classifier. It is clear that as the lags increase, a smaller and smaller portion of the market impact is captured. Eventually, at a lag of about 50% of the block trade duration, the purchase and sale price received by the strategic trader will be equal and he will earn zero expected profit. Any increase in the lag after this will result in expected losses. This relationship between lags and expected profit by the strategic trader is displayed in Figure 8.

[FIGURE 8 ABOUT HERE]

Given a block trader's time horizon and block size, he should be able to decide whether he will lose his anonymity and attract strategic traders. If this is the case, and he may want to trade strategically rather than continuously as most optimal trading strategies recommend.

2.5 Conclusion

In a simulated market environment, a hypothetical strategic trader attempted to use a simple Markov state switching model to identify the presence of block trades of various size and duration. For sufficiently large and extended block trades, the strategic trader was able to identify the block trader sufficiently well so that he could earn positive profits from a simple long-short trading strategy that attempted to 'ride' the market impact of the block trades. The block-trading strategies, as well as the identification schemes used in this paper have been fairly simple, and certainly more sophisticated methods are employed by real agency and proprietary trading algorithms. However, as a baseline case, the identifiability of block trades from publicly available data is certainly a possibility, and this adds an important strategic consideration to block trading that is typically ignored in the optimal trading literature.

2.6 Appendix I: State Space Models

The unknown parameters of the Markov state switching models used in this paper are estimated by maximum likelihood estimation in S-Plus. Since the estimation process is iterative, directly programming it in S-Plus can result in slow and inefficient computation. The S-Plus package S+FinMetrics/SsfPack has estimation procedures using fast C code for optimization that can be adopted for Markov switching models, however the models must be formatted as state space models.

A state space model is a model of the form:

$$\begin{aligned} \alpha_{t+1} &= d_t + T_t \cdot \alpha_t + H_t \cdot \eta_t \\ y_t &= c_t + Z_t \cdot \alpha_t + G_t \cdot \varepsilon_t \end{aligned}$$

where $\alpha_1 \sim N(a, P)$, $\eta_t \sim iidN(0, I_r)$, $\varepsilon_t \sim iidN(0, I_N)$, and $E[\varepsilon_t \eta_t'] = 0$. The first equation is called the transition equation and describes the motion of the state vector α_t. The second equation is called the measurement equation and describes the dynamics of the observed variables y_t. This general format can capture many of the time series models used in economics including ARMA models, time-varying regression models, unobserved component models, stochastic volatility models, and non-parametric models, among others. Detailed descriptions of state space methods can be found in [55] and [26], and descriptions of their implementation with SsfPack can be found in [57] and [74].

SsfPack requires state space models to be described in a compact form:

$$\begin{pmatrix} \alpha_{t+1} \\ y_t \end{pmatrix} = \delta_t + \Phi_t \cdot \alpha_t + u_t$$

where $u_t \sim iidN(0, \Omega_t)$ and

$$\delta_t = \begin{pmatrix} d_t \\ c_t \end{pmatrix}, \Phi_t = \begin{pmatrix} T_t \\ Z_t \end{pmatrix}, u_t = \begin{pmatrix} H_t \eta_t \\ G_t \varepsilon_t \end{pmatrix}, \Omega_t = \begin{pmatrix} H_t H_t' & 0 \\ 0 & G_t G_t' \end{pmatrix}$$

To represent a Markov state switching model, some of the parameters are state-dependent and need to be specified: $\delta_t = \delta_{S_t}$, $\Phi_t = \Phi_{S_t}$, $\Omega_t = \Omega_{S_t}$. For the simple, two-state intercept models used for modeling the volume trade, much of the state space model structure is not

needed. The model can be specified using:

$$\delta_{S_t} = \begin{cases} \binom{0}{c_1} & \text{if} \quad S_t = 1 \\ \binom{0}{c_2} & \text{if} \quad S_t = 2 \end{cases}, \qquad \Phi_{S_t} = \binom{0}{0}, \qquad \Omega_{S_t} = \begin{cases} \begin{pmatrix} 0 & 0 \\ 0 & \sigma_{\varepsilon_1} \end{pmatrix} & \text{if} \quad S_t = 1 \\ \begin{pmatrix} 0 & 0 \\ 0 & \sigma_{\varepsilon_2} \end{pmatrix} & \text{if} \quad S_t = 2 \end{cases}$$

with a two-by-two transition matrix and some initial value parameters.

2.7 Appendix II: Classification Performance Measures

In the simulations created for this study, we know the exact times when simulated block trading is occurring which allows us to determine how well the block-trading classifiers perform. The purpose of the identification scheme (i.e. the Markov state switching model) is to identify when this block trading is occurring and when it is not. In the computing science literature, this is referred to as a clustering problem (also referred to as an unsupervised learning task) since we want to group (or cluster) each trading period into sets corresponding to block-trading periods or normal-trading periods. Clustering problems are distinct from classification problems (also referred to as supervised learning) since the true classes are not available for learning or estimation under clustering, but are for classification. In this paper, we have aspects of both clustering and classification: we do not use the true classes for learning or estimation, like a clustering task, but we do know the true classes, like a classification task. So although we are using clustering methods, we can use classification performance measures after the fact to determine how well the clustering methods work. While the notion of clustering is developed throughout this paper, this appendix reviews the formal definition of a classification problem and several useful performance measures that are used in this paper. Classification is a very general task, but it is presented in this appendix in the context of our trading problem.

Classification is the task of assigning each period in our time series $\mathcal{T} = \{t_1, ..., t_{|\mathcal{T}|}\}$ to predefined classes (or states) $\mathcal{C} = \{c_1, ..., c_{|\mathcal{C}|}\}$. Thus, a classifier is a function

$$\mathcal{F} : \mathcal{T} \times \mathcal{C} \to \{T, F\}$$

where $\mathcal{F}(t_i, c_j) = T$ if and only if period t_i belongs to class c_j. Under this general specification, a period may simultaneously belong to multiple classes, as may be appropriate under certain definitions of the states. Such a problem is usually broken down into $|\mathcal{C}|$ simpler binary classification problems. That is, for each class $c_i \in \mathcal{C}$, we define a binary classifier

$$\mathcal{F}_i : \mathcal{T} \to \{T, F\}$$

where $\mathcal{F}_i(t_j) = T$ if and only if period t_j belongs to class c_i. When exactly one class can be assigned to each document, as in the case of 'BLOCK TRADE' and 'NO BLOCK TRADE',

the classifier will be of the form

$$\mathcal{F} : \mathcal{T} \to \mathcal{C}$$

As before, depending on the classification algorithm chosen, this type of problem may be reduced to a set of binary classifiers with rules for dealing with multiple class assignments.

If $\mathcal{F}$ is the correct or authoritative classifier, then we wish to approximate this function with $\widehat{\mathcal{F}}$. Approximating classifiers has been extensively studied in the field of machine learning [62]. The specific classification scheme of $\widehat{\mathcal{F}}$ is determined by a set of training documents for which the correct classifications are known. These training documents can be classified by a domain expert according to their specific properties, or they can be classified according to some specific data which is aligned with the time series. As discussed above, the task of this paper is not a classification task, but rater a clustering task. However, since we do know the true classes for our simulated date, we can use the evaluation measures that are traditionally applied to classifiers.

Recall that the classifier $\mathcal{F}$ maps $\mathcal{T} \times \mathcal{C}$ into the $\{T, F\}$ such that $\mathcal{F}(t_i, c_j) = T$ if and only if period t_i belongs to class c_j. In this case, we call t_i a positive example of class c_j. When $\mathcal{F}(t_i, c_j) = F$ then period t_i is not a member of the class c_j and so we call t_i a negative example of c_j. To capture the correctness of classifications from a trained classifier $\widehat{\mathcal{F}}$, we introduce the following four basic evaluation functions:

$$TP_{\widehat{\mathcal{F}}}(t_i, c_j) = \begin{cases} 1 & \text{if} \quad \widehat{\mathcal{F}}(t_i, c_j) = T \ \text{ and } \ \mathcal{F}(t_i, c_j) = T \\ 0 & \text{otherwise} \end{cases}$$

$$TN_{\widehat{\mathcal{F}}}(t_i, c_j) = \begin{cases} 1 & \text{if} \quad \widehat{\mathcal{F}}(t_i, c_j) = F \ \text{ and } \ \mathcal{F}(t_i, c_j) = F \\ 0 & \text{otherwise} \end{cases}$$

$$FP_{\widehat{\mathcal{F}}}(t_i, c_j) = \begin{cases} 1 & \text{if} \quad \widehat{\mathcal{F}}(t_i, c_j) = T \ \text{ and } \ \mathcal{F}(t_i, c_j) = F \\ 0 & \text{otherwise} \end{cases}$$

$$FN_{\widehat{\mathcal{F}}}(t_i, c_j) = \begin{cases} 1 & \text{if} \quad \widehat{\mathcal{F}}(t_i, c_j) = F \ \text{ and } \ \mathcal{F}(t_i, c_j) = T \\ 0 & \text{otherwise} \end{cases}$$

Thus, every classification by $\widehat{\mathcal{F}}$ will either be a true positive (TP), a true negative (TN), a false positive (FP), or a false negative (FN). The TPs and TNs indicate correct classifications, and the FPs and FNs indicate incorrect classifications.

A simple, and sometimes overused, performance measure of classifiers is accuracy (A) measured as the proportion of correct classifications.

$$A = \frac{\sum_{i=1}^{|\mathcal{T}|}\sum_{j=1}^{|\mathcal{C}|}\left(TP_{\widehat{\mathcal{F}}}(t_i,c_j) + TN_{\widehat{\mathcal{F}}}(t_i,c_j)\right)}{\sum_{i=1}^{|\mathcal{T}|}\sum_{j=1}^{|\mathcal{C}|}\left(TP_{\widehat{\mathcal{F}}}(t_i,c_j) + TN_{\widehat{\mathcal{F}}}(t_i,c_j) + FP_{\widehat{\mathcal{F}}}(t_i,c_j) + FN_{\widehat{\mathcal{F}}}(t_i,c_j)\right)}$$

The converse of accuracy is the error rate (E) measured as

$$E = 1 - A$$

Accuracy, and error, are often useful performance measures, but they do not always capture the intended notion of correctness. For example, when trying to classify rare events, positive and negative examples will be strongly imbalanced with far more negative than positive examples. In this case, a universal rejector (i.e. $\widehat{\mathcal{F}}(t_i,c_j) = F,\ \forall t_i,c_j$) will have a high accuracy while being of no practical use.

To address such concerns, two other performance measures have become popular: precision and recall. Precision, with respect to a class c_j, is the proportion of periods assigned to class c_j which actually belong to that class.

$$P_{\widehat{\mathcal{F}}}(c_j) = \frac{\sum_{i=1}^{|\mathcal{T}|} TP_{\widehat{\mathcal{F}}}(t_i,c_j)}{\sum_{i=1}^{|\mathcal{T}|}\left(TP_{\widehat{\mathcal{F}}}(t_i,c_j) + FP_{\widehat{\mathcal{F}}}(t_i,c_j)\right)}$$

So, given a particular class, precision is the ratio of correct positive classifications to the total number of positive classifications. Precision can be aggregated across classes in two ways. First, microaveraged precision averages the precision of $\widehat{\mathcal{F}}$ for each class, weighted by the number of positive examples.

$$P_{\widehat{\mathcal{F}}}^{Micro} = \frac{\sum_{j=1}^{|\mathcal{C}|}\sum_{i=1}^{|\mathcal{T}|} TP_{\widehat{\mathcal{F}}}(t_i,c_j)}{\sum_{j=1}^{|\mathcal{C}|}\sum_{i=1}^{|\mathcal{T}|}\left(TP_{\widehat{\mathcal{F}}}(t_i,c_j) + FP_{\widehat{\mathcal{F}}}(t_i,c_j)\right)}$$

Alternatively, macroaveraged precision averages the precision for each class with equal weights.

$$P_{\widehat{\mathcal{F}}}^{Macro} = \frac{\sum_{j=1}^{|\mathcal{C}|} P_{\widehat{\mathcal{F}}}(c_j)}{|\mathcal{C}|}$$

The second performance measure is recall. Recall, for a given class c_j, is the proportion of periods that truly belong to c_j which are classified as belonging to that class by $\widehat{\mathcal{F}}$.

$$R_{\widehat{\mathcal{F}}}(c_j) = \frac{\sum_{i=1}^{|\mathcal{T}|} TP_{\widehat{\mathcal{F}}}(t_i, c_j)}{\sum_{i=1}^{|\mathcal{T}|} \left(TP_{\widehat{\mathcal{F}}}(t_i, c_j) + FN_{\widehat{\mathcal{F}}}(t_i, c_j)\right)}$$

So, given a particular class, precision is the ratio of the number of correct positive classifications by the total number of truly positive class examples. As with precisions, we can use microaveraging to define a measure of recall across all classes.

$$R_{\widehat{\mathcal{F}}}^{Micro} = \frac{\sum_{j=1}^{|\mathcal{C}|} \sum_{i=1}^{|\mathcal{T}|} TP_{\widehat{\mathcal{F}}}(t_i, c_j)}{\sum_{j=1}^{|\mathcal{C}|} \sum_{i=1}^{|\mathcal{T}|} \left(TP_{\widehat{\mathcal{F}}}(t_i, c_j) + FN_{\widehat{\mathcal{F}}}(t_i, c_j)\right)}$$

Alternatively, we may define a macroaverage measure of recall.

$$R_{\widehat{\mathcal{F}}}^{Macro} = \frac{\sum_{j=1}^{|\mathcal{C}|} R_{\widehat{\mathcal{F}}}(c_j)}{|\mathcal{C}|}$$

Most classifiers can be set up to tradeoff precision for recall, or vice-versa. Consequently, it is useful to present a combined measure of performance, the F_1 score, which is the harmonic mean of precision and recall:

$$F_1 = \frac{2 \times P \times R}{P + R}$$

where P and R are either micro- or macroaveraged.

$$V_t = \alpha_0(1 - S_t) + \alpha_1 S_t + \varepsilon_t, \;\; \varepsilon_t \sim N(0, \sigma_0^2(1 - S_t) + \sigma_1^2 S_t)$$

Model Parameter	Estimated Value	p-Value
α_0	99.43	0.000
α_1	129.45	0.000
σ_0^2	0.0101	0.000
σ_1^2	0.0059	0.000
P_{12}	0.0013	0.000
P_{21}	0.0053	0.000
$P_{11} = 1 - P_{12}$	0.9987	0.000
$P_{22} = 1 - P_{21}$	0.9947	0.000

Table 2.1: Estimates of a Markov state switching model. Data is from simulated volume data with 1000 periods per day for 20 days, with a block trade each day lasting 200 periods with a trade rate of 30% of normal volume. True values for α_0 and α_1 are 100 and 130, respectively. The model was estimated on volume data that had been transformed by the natural logarithm function, but intercept estimates are presented after being untransformed by the exponential function. Variance estimates are for the log-transformed data. P_{ij} is the probability of switching from state i to state j.

Block Trade Duration	Classification Measure	Block Trade Rate 0.30	0.20	0.10	0.05	0.02
	Accuracy	0.998	0.994	0.978	0.945	0.825
0.20	Precision	0.996	0.989	0.962	0.885	0.565
	Recall	0.993	0.980	0.926	0.831	0.539
	Accuracy	0.997	0.995	0.978	0.943	0.894
0.10	Precision	0.984	0.976	0.895	0.824	0.439
	Recall	0.984	0.970	0.889	0.552	0.227
	Accuracy	0.997	0.995	0.982	0.956	n/a
0.05	Precision	0.979	0.971	0.868	0.664	n/a
	Recall	0.970	0.932	0.748	0.275	n/a
	Accuracy	0.998	0.994	0.985	n/a	n/a
0.025	Precision	0.968	0.900	0.712	n/a	n/a
	Recall	0.920	0.808	0.420	n/a	n/a

Table 2.2: Binary classification performance. The block trade duration is the fraction of a trading day that the block trade takes to execute. Block trade rate is the percentage of the expected volume that the block trade adds to the market volume. Accuracy, Precision and Recall are classification performance measures. Accuracy is measured as the proportion of correct classifications. Precision is the proportion of periods predicted to be block trading periods in which a block trade is truly executing. Recall is the proportion of periods that are truly block trading periods which are classified as belonging to that class. Entries marked with an 'n/a' are not available because the estimator would not converge for simulations with block trades of low trading rate and short duration.

Block Trade Duration	Performance Measure	Block Trade Rate 0.30	0.20	0.10	0.05	0.02
	Accuracy	0.995	0.978	0.948	0.911	0.851
0.20	Precision	0.974	0.895	0.750	0.558	0.238
	Recall	0.971	0.886	0.721	0.524	0.232
	Sharpe Multiplier	6.81	3.66	0.08	−0.09	−5.02
	Accuracy	0.992	0.986	0.964	0.948	0.936
0.10	Precision	0.926	0.864	0.641	0.466	0.224
	Recall	0.923	0.857	0.636	0.311	0.120
	Sharpe Multiplier	13.51	4.61	1.11	−1.02	−5.03
	Accuracy	0.994	0.987	0.977	0.972	n/a
0.05	Precision	0.877	0.752	0.558	0.356	n/a
	Recall	0.869	0.717	0.475	0.164	n/a
	Sharpe Multiplier	4.41	3.51	−1.13	−0.96	n/a
	Accuracy	0.995	0.994	0.990	n/a	n/a
0.025	Precision	0.777	0.725	0.480	n/a	n/a
	Recall	0.738	0.650	0.283	n/a	n/a
	Sharpe Multiplier	−0.08	−16.27	−6.52	n/a	n/a

Table 2.3: Three-state classification and trading performance. The block trade duration is the fraction of a trading day that the block trade takes to execute. Block trade rate is the percentage of the expected volume that the block trade adds to the market volume. Accuracy, Precision and Recall are classification performance measures. Accuracy is measured as the proportion of correct classifications. Precision is the proportion of periods predicted to be block trading periods in which a block trade is truly executing. Recall is the proportion of periods that are truly block trading periods which are classified as belonging to that class. The Sharpe multiplier is an investment performance measure that compares a strategic trader's performance resulting from identifying block trades to the better of a long-only or short-only strategy. Entries marked with an 'n/a' are not available because the estimator would not converge for simulations with block trades of low trading rate and short duration.

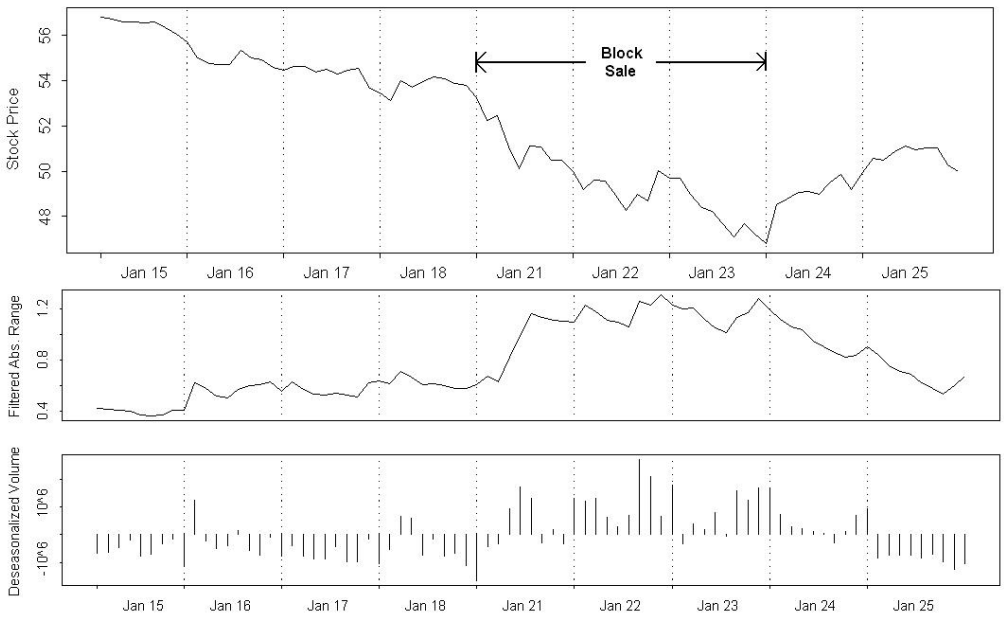


Figure 2.1: Example of a block trade. Over the three days of January 21 - 23, 2008, Societe Generale had to sell the the large positions (approximate value of 49.9 billion Euros) accumulated by rogue trader Jerome Kerviel. The figure shows data on Total SA stock which is a significant part of the CAC 40 index of French stocks. The top panel shows the stock price, the middle panel shows the filtered absolute hourly range (a proxy for volatility), and the bottom panel shows deseasonalized hourly volume. Source: Bloomberg.

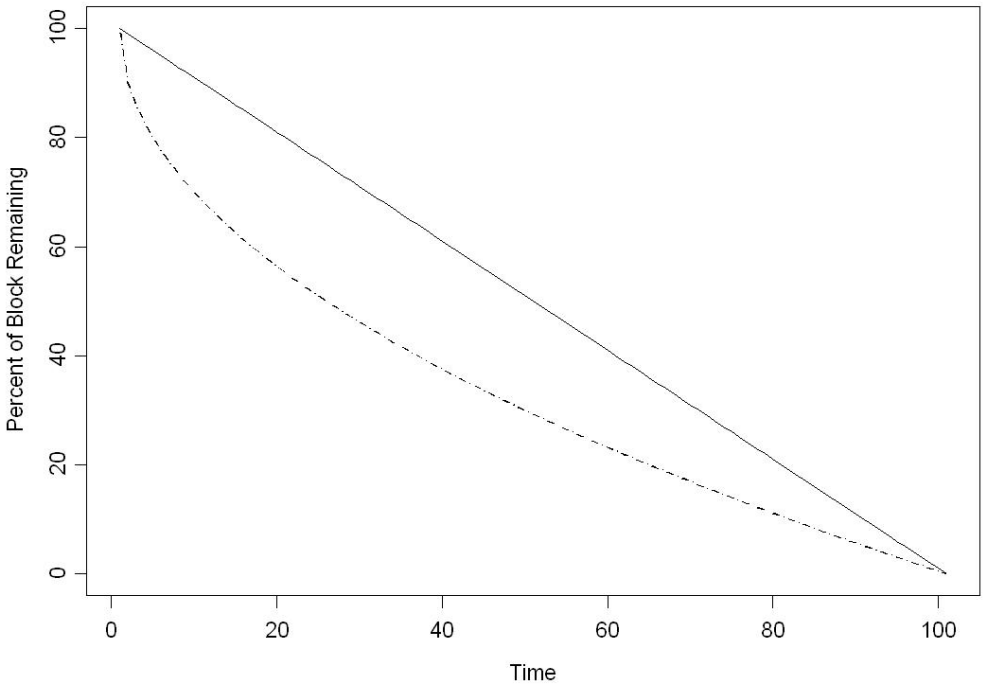


Figure 2.2: Two block trading trajectories. The solid line represents a constant trading rate throughout the block trade. The broken line represents a block trade that is more aggressively traded early in the trade but with a decelerating trading rate.

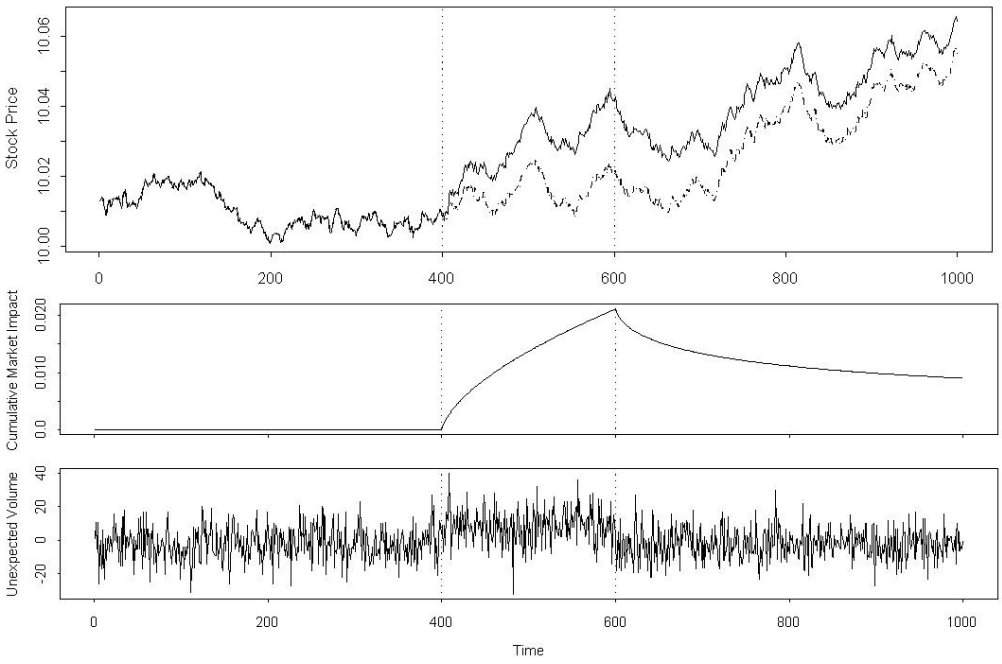


Figure 2.3: Example of simulated block trade. In this example, a block purchase is executed between times 400 and 600, during which time the block trade adds 10% to the expected volume. The top panel shows the stock price with the block trade present (solid line) and absent (broken line), the middle panel shows the cumulative market impact of the block trade, and the bottom panel shows the demeaned transaction volume.

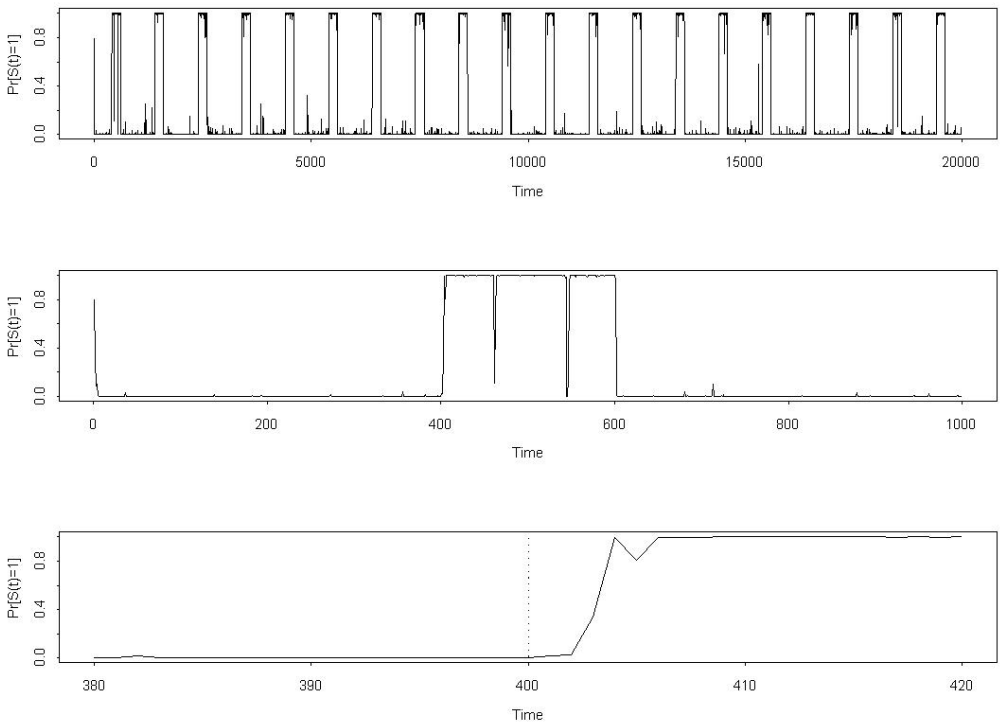


Figure 2.4: Estimated probabilities of block trading. Each panel shows the estimated probability of block trading ($S_t = 1$) from the Markov state switching model. The top panel shows the full sample of 20 days with a 0.20-day-long block trade occurring each day, the middle panel shows the first day of the sample when a block trade occurs between times 400 and 600, and the bottom panel shows the transition period around time 400.

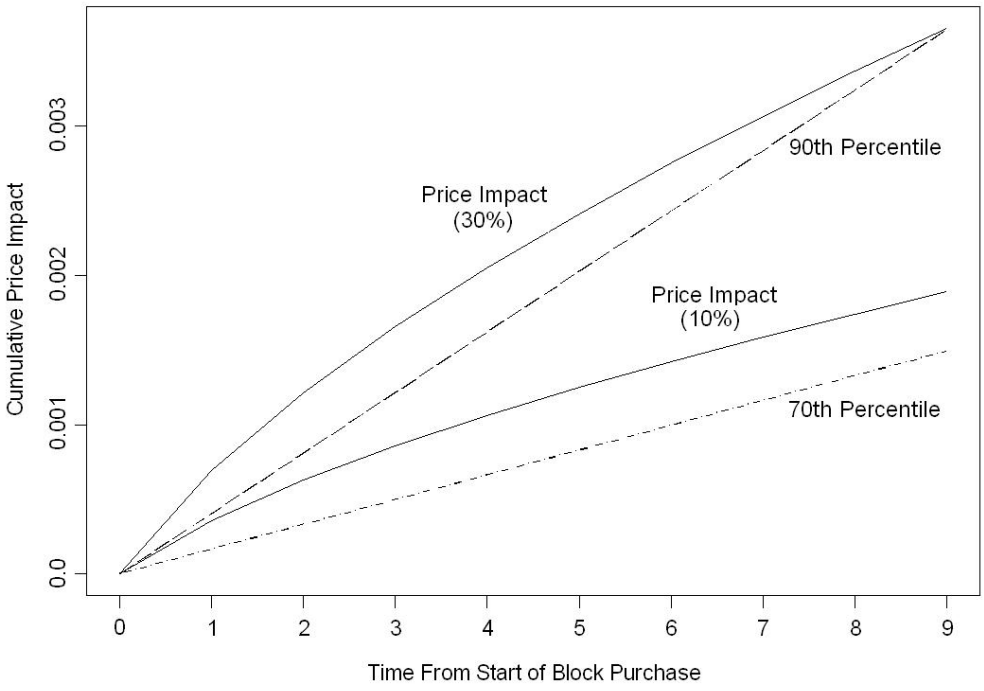


Figure 2.5: Cumulative market impact and price innovation percentiles after a block trade begins. The cumulative market impact figures are for block trades of 30% and 10% of typical volumes. Percentiles of cumulative price innovations increase linearly with the time interval.

Figure 2.6: Simulated trading returns. Simulations are for an environment with a block trader trading at a rate of 20% of typical volume and for a duration of 0.20 days for 20 days. Block trades alternate between days that contain a block purchase and days that contain a block sale. The cumulative returns are shown only for days that the strategic trader is invested. The upper line is the cumulative noncompounded returns for being long when a block seller is predicted to be present, and being short when a block seller is predicted to be present. The lower line is the long-only strategy for the same periods.

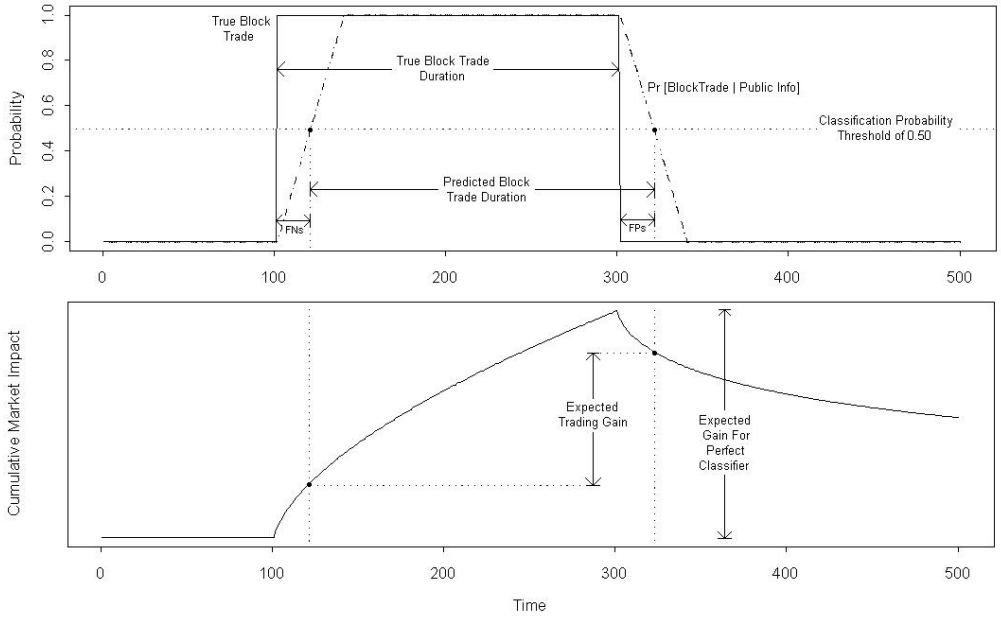


Figure 2.7: Potential trading gains from block trade identification scheme with lags. The top panel shows when a block purchase occurs and the identification scheme's estimated probability. The classifier predicts a block purchase when the probability is greater than 0.50. The scheme has an identification lag which causes 'false negative' classifications (FNs) at the beginning of the block trade, and 'false positive' classifications (FPs) after the block trade. The expected gains from trading using the classifier to signal when to take a long position are shown in the bottom panel. The trading scheme captures the cumulative market impact of the block trade, less the market impact during the FNs, and less the decay during the FPs.

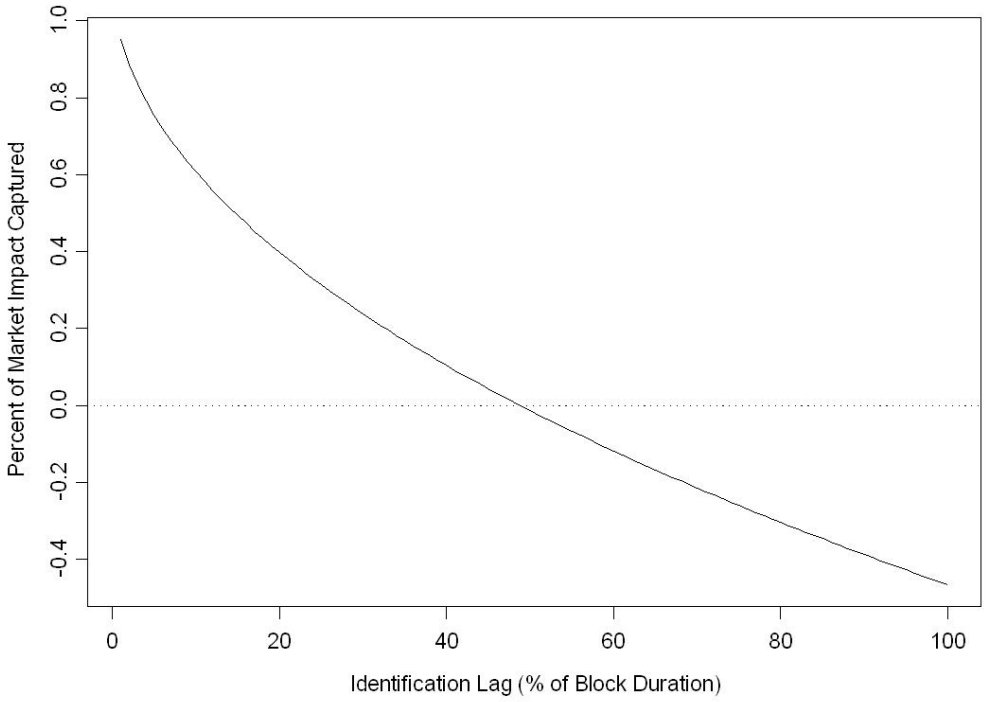


Figure 2.8: Decay of trading performance as identification lag increases. The horizontal axis represents the delay (as a percentage of the block trade duration) after a block trade begins or ends that a classifier requires to identify the changed state. The vertical axis represents the portion of the total market impact that the trader can capture. As the identification lag increases, the expected trading gain decreases and eventually becomes negative. Identification lags at the beginning and end of a block trade are assumed to be equal.

Chapter 3

Mini Flash Crashes and High Frequency Trading

3.1 Introduction

On May 6th, 2010, an historic financial event occurred which immediately was named the Flash Crash. This crash was unique in comparison to other historical market crashes in how quickly prices fell and then recovered. In a matter of minutes, U.S. stocks, stock indices, futures, options, and exchange-traded funds experienced a rapid drop in value of more than 5%, and a few minutes later, markets rebounded to near pre-crash levels. For those few minutes, approximately $1 trillion in market value disappeared. This was purely a liquidity shock resulting from the new high-speed electronic market structure that has evolved over the past few years. Researchers, policy makers, and market participants are concerned with the question of how likely flash crashes are to be in the future.

This paper contributes to our understanding of the likelihood of future flash crashes by examining the frequency of 'flash-type' events before and after the big Flash Crash of May 6th. These 'mini flash crashes' are of a smaller scale than the big crash, but are still substantial enough to be considered serious liquidity shocks. Since they have similar characteristics to the big crash, any change in frequency of these mini flash crashes may tell us about the increasing or decreasing likelihood of another big one.

The financial media has been reporting occurrences of miniature flash crashes since May 6th, and within the data used in this paper there have been more flash-type events with

shorter inter-flash durations. Despite this suggestive evidence however, survival analysis of these events does not establish the significance of these differences from before and after the big Flash Crash.

This paper proceeds by first detailing the essential features of the big Flash Crash of May 6th, and then discusses the new high-speed market structure in which the crash occurred. Next, the data that is used in this study is described and mini flash crashes are characterized. Tests are performed to determine the relative frequency of these liquidity shocks and the results are discussed, along with avenues for future research.

3.2 The Flash Crash of May 6th

To get a sense of the pace of events on May 6th, consider the following rough time line of price moves for the June 2010 E-Mini S&P 500 futures contract. Between 13 : 32 : 00 and 13 : 45 : 38 (CT), the futures price fell from 1127.45 to 1056.00 - a 6.3% fall in under 15 minutes. At certain points in the descent, prices fell at a rate of more than 1% in a second, triggering the CME Globex Stop Logic Functionality which paused trading for 5 seconds. Between 13 : 45 : 38 and 14 : 06 : 00, the futures price rapidly rose to 1123.75 - a 6.4% rise in just over 20 minutes. Prior to this, the market was experiencing "unusually high volatility and thinning liquidity" [67], and the price had already fallen more than 2%. To illustrate the magnitude and velocity of the price changes during the Flash Crash, Figure 1 shows how the SPDR S&P 500 ETF behaved during the crash. This extreme intraday volatility also saw the Dow Jones Industrial Average suffer its largest one-day point decline of 998.5 in it's history. Additionally, many individual equities suffered extreme drops, with some large firms seeing their shares briefly trading for pennies.

[FIGURE 1 ABOUT HERE]

Immediately, theories emerged to explain the cause of the crash. These theories included 'fat finger' theories whereby a trade was mistakenly submitted with price and quantity that immediately consumed all of the available liquidity. Other theories pointed to predatory practices by high-frequency trading firms. In truth, many diverse factors converged to create this market dislocation. The joint investigation by the Securities and Exchange Committee (SEC) and the Commodity Futures Trading Commission (CFTC) led to a report [67] that investigated the causes and circumstances that led to the crash. They identified a large sell

order of 75,000 E-Mini S&P 500 futures contracts (with a nominal value of approximately $4.1 billion) placed at 13 : 32 : 00 from a mutual fund complex as the precipitating cause of the crisis.

This large block trade was intended as a hedge against the funds' equity position. It was performed through a trading algorithm that was set to execute at a rate of 9% of the market's volume from the previous minute. Given the large volumes that were seen during the flash crash, with peak minute-volume reaching approximately 80,000 contracts compared to a normal rage of approximately 5,000 to 20,000, this trade would have had significant market impact. It should be noted that the CME Group disputes this trade as the primary cause [23], noting that the trade was executed over a 20 minute period and that half of the block's volume executed as the market rallied following the crash. [56] points out that this trader had a history of executing very large trades in a short amount of time, but that it's last trade of that size took more than 5 hours to execute.

While large trades have long been known to have significant market impact [12, 58], the fund complex's large trade is not the full story since it does not explain how the Flash Crash immediately spread to so many markets. In the investigations by regulators, as well as in the small but growing academic literature [56, 27], the behavior of other traders is believed to have led to turning the market impact of a large trade into a market wide crash. In particular, a new class of high-frequency traders, who have become the de facto market makers in recent years, have come under close scrutiny. The next section describes this class of trader, and explores their role in the Flash Crash.

3.3 The High-Frequency Trading Environment

Due to regulatory restrictions, prior to 1999, stock exchanges in the U.S. had a monopoly in the provision of liquidity for most stocks. In that year, Regulation ATS opened up the exchanges to competition from alternative trading venues. At that time, the low cost of technology enabled the establishment of markets to compete directly with the more traditional exchanges. One of the ways that trading venues competed for market share was by offering faster access to market data and lower latency for order submissions. The traders who exploited this fast trading technology the most are called high-frequency traders.

Currently, a high-frequency trader can place thousands of orders per second, and when these orders are executed, securities may be held for just a fraction of a second. Once at

the exchange these trades can be processed in less than 500 microseconds. The general aim of these traders is to make tiny profits per trade, but make an enormous number of trades. Consequently, the average daily volume on the NYSE has increased by more than 160% since 2005, with high-frequency traders participating in more than 70% of equity transactions. Of course, no human is fast enough to make such trades, so these traders use trading algorithms executed on the fastest computers available, usually located as close to the exchange's computers as possible in order to shave microseconds off of the time it takes to receive information and transmit orders.

Since fundamental information about a company is not likely to change in a matter of milliseconds, the majority of orders submitted by high frequency traders are in response to information generated by the market itself. One way that high-frequency traders exploit this information is by acting as the de facto market makers in many markets. This activity is particularly profitable because of the liquidity rebates that almost all trading venues pay to liquidity providers. Since trading venues earn profits from transaction fees, there is keen competition among them to attract high-frequency traders and their large number of trades. One way that exchanges do this is by offering small payments to its largest and most active traders for 'adding liquidity' by placing nonmarketable orders. These payments are made per share transacted and are currently around the $0.0025 level. On the other hand, similar fees are charged to those traders who 'take liquidity' by placing marketable orders which remove orders from the order book. Exchanges cover the commission costs and exchange fees of these high-volume traders, thus rebate traders can earn half a cent per share that they buy and sell.

While the market making activities of high-frequency traders is not a contentious issue, other forms of high-frequency proprietary trading has generated a lot of concern. Opponents of high-frequency trading claim that these traders are so quick and influential that they can manipulate prices and volatilities, thereby creating (rather than discovering) profitable trading opportunities. However, such claims are difficult to verify since the details of their trading strategies are closely guarded secrets. The power of these trading secrets, as well as the desire to keep them secret, is exemplified by the case of Sergey Aleynikov - a former Goldman Sachs employee - who was arrested for leaving the company with parts of it's high frequency trading algorithm. At the bail hearing, the federal prosecutor wanted Mr. Aleynikov held without bail since the code he possessed could "unfairly manipulate" stock

prices. There have been cases of charges brought against high-frequency traders for manipulative practices such as 'quote stuffing.' For example, in 2007, FINRA brought charges against Trillium Trading LLC for using non-bona fide orders to improve executions on other orders on more than 46,000 occasions over a three-month period.

While their volume of trade is very high, the actual number of high frequency trading firms is very low. Only approximately 2% of the 20,000 trading firms in the U.S. employ high-frequency trading [51]. In the detailed study of the E-Mini S&P 500 futures market during the Flash Crash by [56], audit-trail data revealed that only 15 of the almost 12,000 traders were high-frequency traders, yet they were involved in approximately 30% of the transactions. With such a large portion of trading volume concentrated in a small number of trading firms, there is concern that they represent a significant source of systemic risk.

Despite this risk and the manipulative behavior of some traders, the bulk of the empirical evidence indicates that high-frequency trading, in general, improves market quality [48, 47, 19]. [43, 44] investigate the order dynamics in this new high-speed environment, and while the order submissions and cancelations of high-frequency traders can be erratic and occur in high-volume bursts, these episodes do not significantly increase the number of transactions or volatility. Their typical behavior, which usually improves market quality, can sometimes lead to their withdrawal of their liquidity provision, and can even lead to their becoming liquidity demanders. This was the case during the flash crash as buy-side liquidity dramatically fell just as it was needed.

[27] explains the Flash Crash as the consequence of three features of the new market structure: (i) the concentration of liquidity provision into a small number of high-frequency trading firms, (ii) the reduced participation of retail investors, leading to more informed trading and therefore more persistence in net liquidity demand, and (iii) the high sensitivity of high-frequency traders to intraday losses. The high-frequency traders depend on high capital turnover and aim to keep inventories very close to zero. Having provided significant liquidity for a large part of the 75,000 contract order placed by the mutual fund complex, their inventories were substantially large, and in their attempt to lower these inventories, they reduced their liquidity to the buy-side of the market. Eventually, as prices fell, their aversion to intraday losses led them to become buy-side liquidity demanders just as prices were falling.

These dynamics are established features of the market, and consequently, the consensus is that it is just a matter of time before another flash crash happens.

3.4 The Rise of the Mini Flash Crash?

Since the Flash Crash, the financial media has been on the look out for signs of further flash-crash-like activity. With headlines like "The apparent rise of the mini flash crash"[1], "The Flash Crash, in Miniature"[2], and "Plunge in 10 ETFs triggers 'flash crash' memories"[3], one might think that the new 'flash-type' liquidity shocks are more common since May 6th.

These mini flash crashes are happening on a regular basis in individual markets. For example, Figure 2 shows two mini flash crashes in Apple and Citigroup stocks that were reported in the financial media.

[FIGURE 2 ABOUT HERE]

Yet, the media can selectively focus it's attention (see for example [29]). Since the phrase 'flash crash' was not in use before May 6th, the media coverage is not a good indicator of the frequency of such events. In fact the higher awareness of the risks of flash crashes may be making market participants, trading venues, and regulators take steps to reduce their frequency. On the other hand, algorithms being trained on historical data may lead to even quicker withdrawal of liquidity, resulting in more flash crashes. This study is an attempt to formally study the question of the relative frequency of flash crashes before and after May 6th, 2010.

3.5 The Data and Characterization of Flashes

The data used in this study is U.S. equity transaction data for seven months before the Flash Crash (Oct 5th, 2009, to May 5th, 2010) and seven months afterwards (May 7th, 2010, to Dec 7th, 2010). The stocks chosen for the study are the 30 stocks that constitute the Dow Jones Industrial Average and are listed in Appendix A at the end of this paper. These are large, closely watched companies, and their shares are heavily traded. While considered very liquid, all of the Dow 30 stocks fell substantially during the Flash Crash.

[1]By Izabella Kaminska on September 28, 2010 on ft.com and available at `http://ftalphaville.ft.com/blog/2010/09/28/354876/the-apparent-rise-of-the-mini-flash-crash/`

[2]By Graham Bowley on November 8, 2010, on nytimes.com and available at `http://www.nytimes.com/2010/11/09/business/09flash.html`

[3]By Reuters on Apr 01, 2011, and available on `http://www.moneycontrol.com/news/market-news/plunge10-etfs-triggers-flash-crash-memories_533355.html`

Thus, these stocks are suspectable to flash-like liquidity shocks, but they are considered to be relatively stable stocks.

In order to determine whether mini-flashes are more or less frequent since the Flash Crash, a formal definition is required that captures the essence of a flash crash and allows us to search for them in the data. The essential character of the flash crash was it's size, speed, and the extent of its reversion. The reversion is important in identifying flash crashes since it captures the liquidity-shock aspect of the flash and distinguishes them from fast permanent adjustment of price in response to fundamental information. To capture these features, three parameters are needed: (i) a time limit within which a flash episode must be completed, (ii) a minimum size for the fall, and (iii) a minimum amount of recovery after the fall within the time window.

For this study, a window of 30 minutes was used since this was the time frame within which the Flash Crash of May 6th was completed. The choice of the minimum price change to characterize a mini flash crash requires consideration of several factors. First, the size should not be too large so that there are insufficient observations to perform statistical tests. Second, since different stocks experience different levels of intraday volatility, the price change threshold needs to be adjusted for each stock's volatility. Since crashes are often described as 'N-Sigma' moves, this was the strategy chosen for this study. Daily price data, adjusted for dividends and stock splits, were used to calculate the standard deviation of daily returns of each stock, $\sigma_{i,day}$, over the full 14-month sample period. These were then scaled by $\sqrt{\frac{1}{13}}$ to get an estimate of the 30-minute volatility, $\widetilde{\sigma}_{i,30\,\mathrm{min}}$. This scaling is not technically correct as noted by [25], but is widely used in practice and accepted by regulators. For the purposes of generating a price change threshold that can be used for cross-stock comparisons, the estimation errors should not make a significant difference. Table I lists the daily standard deviations and the 5-sigma thresholds for the stocks in the sample.

[Table I ABOUT HERE]

With this 30-minute sigma value, the price change threshold for each stock was set at $5 \times \widetilde{\sigma}_{i,30\,\mathrm{min}}$. That is, a stock must fall at least this much to be considered a mini flash crash. 5-Sigma price changes are substantial moves. For example, under the normally distributed innovations of a Brownian motion which is common in financial modeling, a 5-sigma price move would be expected only once every $1,744,278$ intervals (once every 530 years under

the assumption of 250 6.5-hour trading days per year). Clearly, the normal distribution does not describe intraday financial returns very well.

In addition to a 5-sigma price change, a flash crash requires a quick reversion since it is essentially a short-lived liquidity shock, and once liquidity is restored, prices should return to levels more reflective of fundamental value. A minimum reversion move of 2.5-sigma from the crash's minimum is used for this study. The 5-sigma fall and 2.5-sigma reversion must be completed within the 30-minute window to be counted as a mini flash crash.

The 5-sigma threshold value is based on an unconditional estimate of stock return volatility. Since return volatility is known to change over time, a time-varying threshold might be considered a more consistent threshold. For this study, a time-varying threshold was not used for two reasons. First, if flash crashes were more frequent since May 6th, then this would likely increase the time-varying threshold thereby making the flashes more difficult to identify. Second, a flash may temporarily increase the threshold, and if flashes cluster in time, subsequent flash crashes may not be detected. So while the constant threshold is based on a crude measure of volatility, its simplicity allows for easy comparison without the inconsistent identification of equal sized mini-crashes in different times in the sample.

While flash crashes are more likely to generate public concern because of the long-bias of their portfolios, inventory and intraday loss pressures are equally likely to be felt by high-frequency market makers from both the long and short sides of the market. Therefore, we may observe flash spikes (price increases followed by downward reversion) in addition to flash crashes (price dips). Tables II lists the 5-sigma flash dips and spikes experienced during the seven months before and after the flash crash of May 6th by each of the 30 DJIA components.

[Table II ABOUT HERE]

3.6 Survival Analysis

An appropriate framework to examine the likelihood of flash spikes and dips before and after May 6th is survival analysis. Survival analysis has been developed to deal with 'time-to-event' data. These techniques are used extensively in medical research (e.g. time-to-death under treatment and control conditions) and electronic component testing (e.g. time-to-failure of a component). In finance research, survival analysis has been used to model company defaults as well as transaction durations. A key aspect of this data that survival

analysis deals with is censoring - where a sample period may end before an event occurred or may start after the interval began. For example, in a medical context, a measurement of time-to-death will be 'right censored' if the patient is still alive at the end of the experiment. Prior to modern survival analysis, censored data resulted in data loss and biased estimates.

In the context of this study, the events studied are the mini flash crashes and spikes described in the previous section. The time-to-event observations are the lengths of time between two adjacent flash events (either a dip or a spike) for a single stock. This results in 445 interval measurements that are split into two groups: the first group are those flash events that occurred in the seven months before May 6th, and the second group are those that occurred in the seven months after. In the sample before May 6th, the first intervals for each stock are left-censored. There are no right-censored observations in this group since all stocks experienced a flash crash on May 6th. Similarly, there are no left-censored observations in the later group, however, the last interval for each stock in the later group is right-censored.

If one flash-event makes another much more (or much less) likely in the near future, then this kind of temporal dependency can easily be handled by the nonlinear nature of survival analysis. However, if flash-events in one stock make flash-events in other stocks more likely, then this would have to be accounted for for any hypothesis testing. This is a substantial possibility since contagion of this kind occurred during the big flash crash of May 6th. For the smaller flash crashes and spikes studied here, this does not appear to be a problem. A visual inspection of the data found very few occurrences of multiple flashes at the same time across different stocks.

There were 206 intervals between flash events (or sample end points) in the seven months before May 6th, and 239 in the seven months afterwards. While these counts are supportive of the claim of more flash events since the big Flash Crash, further analysis is required. To get a sense of the distribution the time-to-flash data for the two groups, Figure 3 shows the empirical survival function for each of the groups.

[Figure 3 ABOUT HERE]

The solid line, representing the period before May 6th, tends to lie above the dotted line representing the later group. This indicates longer periods between flashes in the earlier group, and this is confirmed in Table III.

[Table III ABOUT HERE]

Despite this evidence that supports the media's claim that flashes are becoming more frequent, the Mantel-Haenszel test (a.k.a. the log-rank test) fails to reject the null hypothesis that the survival functions are significantly different[4].

3.7 Conclusion

Despite the occurrence of more mini flash events since the May 6th Flash Crash, survival analysis of these events in the Dow 30 stocks does not establish the significance of these differences. Wider time samples over more stocks may improve confidence in these results, or may demonstrate that the small differences in frequency are in fact significant. Additionally, only a single characterization of flash events (a 5-sigma move followed by a 2.5-sigma reversion within 30 minutes) was considered. For a robustness test, other characterizations could be examined. Future research could also examine the order and trade activity around these mini flash events to determine whether toxic order flow is at play in these mini events as it was on May 6th, as established by [27].

[4]The null hypothesis that the survival functions are different is not rejected for nearby parameterizations of flash events including changing the window size to 20 minutes or changing the threshold to 4-sigma (with a 2-sigma reversion).

Appendix: Dow Jones Industrial Average Components

Ticker	Company	Ticker	Company
AA	Alcoa	JPM	JP Morgan Chase & Co.
AXP	American Express	KFT	Kraft Foods
BA	Boeing Company	KO	Coca-Cola
BAC	Bank of America	MCD	McDonald's
CAT	Caterpillar	MMM	3M Company
CSCO	Cisco Systems	MRK	Merck & Co
CVX	Chevron	MSFT	Microsoft
DD	E.I. du Pont de Nemours	PFE	Pfizer
DIS	Walt Disney	PG	Procter & Gamble
GE	General Electric	T	AT&T Inc.
HD	Home Depot, Inc.	TRV	The Travelers Co.
HPQ	Hewlett-Packard	UTX	United Technologies
IBM	IBM	VZ	Verizon Communications
INTC	Intel Corporation	WMT	Wal-Mart
JNJ	Johnson & Johnson	XOM	Exxon Mobil

Company	$\widetilde{\sigma}_{day}$	$5 \times \widetilde{\sigma}_{30\,\mathrm{min}}$	Company	$\widetilde{\sigma}_{day}$	$5 \times \widetilde{\sigma}_{30\,\mathrm{min}}$
AA	0.025	0.035	JPM	0.019	0.027
AXP	0.021	0.029	KFT	0.011	0.015
BA	0.019	0.026	KO	0.010	0.013
BAC	0.023	0.033	MCD	0.010	0.013
CAT	0.020	0.028	MMM	0.013	0.018
CSCO	0.020	0.028	MRK	0.014	0.019
CVX	0.013	0.018	MSFT	0.014	0.019
DD	0.017	0.023	PFE	0.014	0.019
DIS	0.015	0.021	PG	0.009	0.012
GE	0.017	0.024	T	0.010	0.014
HD	0.015	0.021	TRV	0.013	0.018
HPQ	0.016	0.022	UTX	0.013	0.018
IBM	0.012	0.016	VZ	0.010	0.015
INTC	0.016	0.022	WMT	0.009	0.012
JNJ	0.008	0.011	XOM	0.012	0.016

Table 3.1: Daily and 30-minute standard deviation estimates for the 30 components of the DJIA index. The daily estimate is from returns calculated on daily end-of-day prices, adjusted for any splits and dividends, from Oct 5th, 2009, to Dec 7th, 2010. The 30-minute estimates are derived from the daily estimates by scaling by $\sqrt{\frac{1}{13}}$. The 30-minute estimates are then multiplied by 5 to get the 5-Sigma threshold for the stock.

	Before May 6th			After May 6th		
Company	*Dips*	*Spikes*	*Total*	*Dips*	*Spikes*	*Total*
AA	5	0	5	0	2	2
AXP	3	1	4	1	2	3
BA	1	1	2	4	3	7
BAC	9	8	17	0	3	3
CAT	2	2	4	1	1	2
CSCO	0	1	1	1	4	5
CVX	2	3	5	5	6	11
DD	1	3	4	2	2	4
DIS	2	0	2	4	3	7
GE	0	3	3	2	5	7
HD	0	0	0	3	2	5
HPQ	0	1	1	4	5	9
IBM	0	3	3	2	0	2
INTC	7	7	14	3	5	8
JNJ	2	7	9	4	7	11
JPM	3	2	5	6	9	15
KFT	1	6	7	3	1	4
KO	3	4	7	3	4	7
MCD	6	0	6	3	6	9
MMM	2	0	2	4	4	8
MRK	5	0	5	4	4	8
MSFT	3	2	5	11	12	23
PFE	5	1	6	2	4	6
PG	3	3	6	1	2	3
T	13	13	26	2	5	7
TRV	2	0	2	1	3	4
UTX	0	2	2	0	3	3
VZ	4	6	10	5	5	10
WMT	5	3	8	0	7	7
XOM	5	1	6	4	3	7
Total	94	83	177	85	122	207

Table 3.2: Number of 5-Sigma 30-minute dips and spikes during the seven months before and after May 5th, 2010.

	Before May 6th	After May 6th
Number of Intervals	206	239
Mean Interval Length	346	295
(Std. Dev.)	34.3	28.0
Median	142	122
LogRank	1.07	
(p-value)	(0.301)	

Table 3.3: Kaplan-Meier survival estimates and the Mantel-Haenszel test (log-rank).

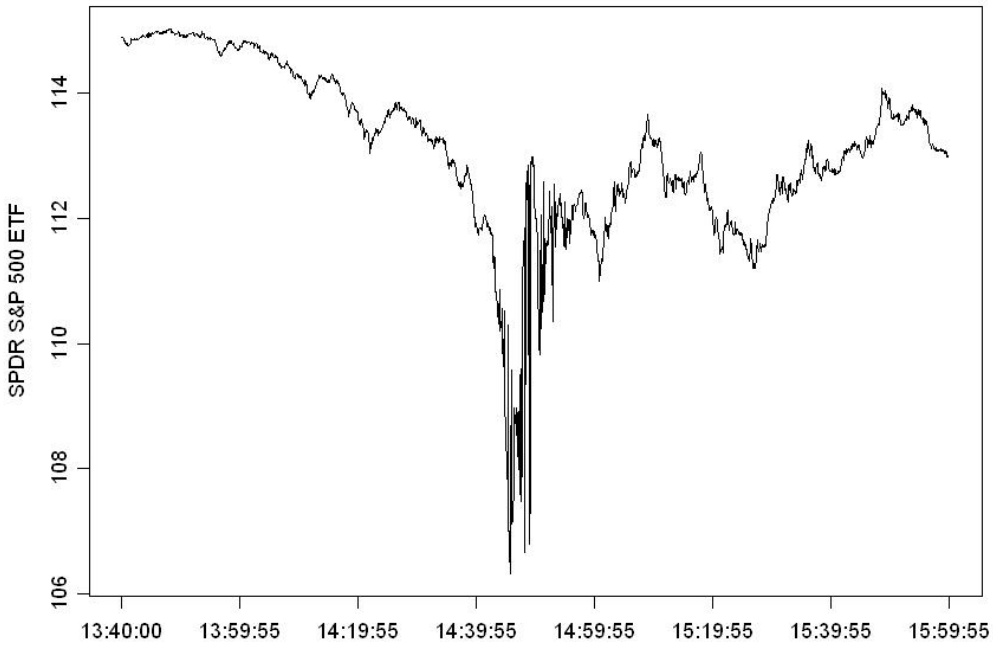


Figure 3.1: SPDR S&P 500 ETF on May 6th, 2010, during the Flash Crash. The SPDR SPY ETF generally corresponds to the price and yield performance of the S&P 500 Index. Prices are the closing price at five-second intervals. Source: NYSE TAQ.

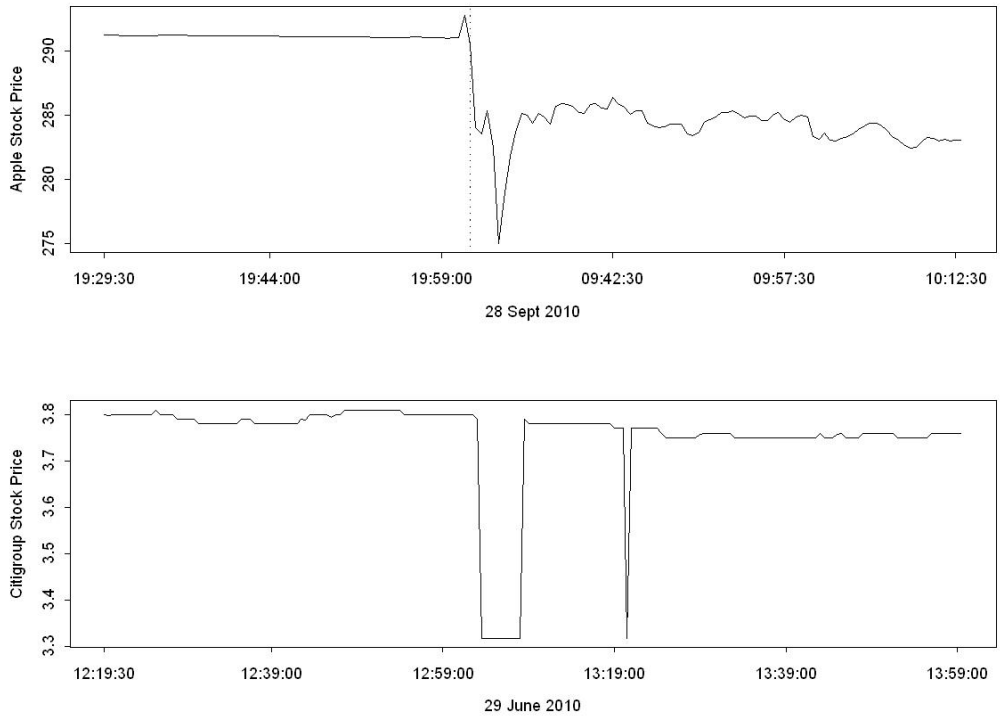


Figure 3.2: Two examples of mini flash crashes in Apple and Citigroup stock. The dotted vertical line on the Apple chart indicated the day's market opening. Prices are from the minimum price seen on 30-second intervals. Source: NYSE TAQ.

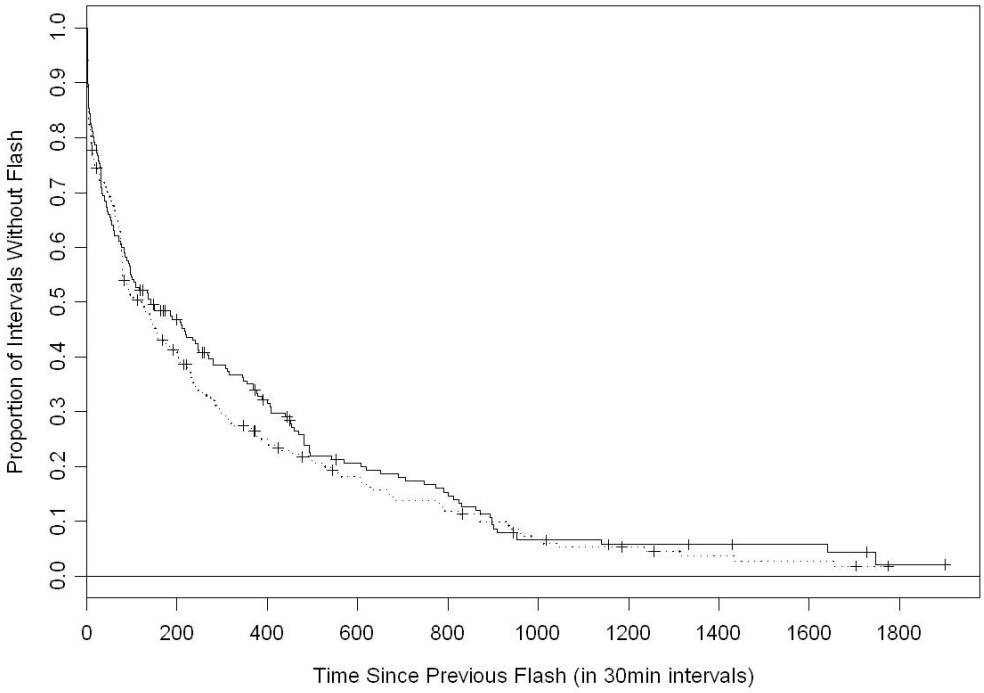


Figure 3.3: Empirical survival function for time to flash events (dips or spikes) before May 6th (solid line) and after (dotted line). The + signs represent censored data.

Bibliography

[1] R. Almgren. Optimal execution with nonlinear impact functions and trading-enhanced risk. *Applied Mathematical Finance*, 10:1–18, 2003.

[2] R. Almgren. Optimal trading in a dynamic market. *Technical Report*, 2009.

[3] R. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Risk*, 3:5–39, 2000.

[4] R. Almgren and J. Lorenz. Algorithmic trading iii: precision, control, execution. *Institutional Investor Journal*, 2007.

[5] R. Almgren, C. Thum, E. Hauptmann, and H. Li. Direct estimation of equity market impact. 2005.

[6] T. Andersen and T. Bollerslev. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4):885–905, May 1998.

[7] T. Andersen, T. Bollerslev, P. Christoffersen, and F. Diebold. Volatility and correlation forecasting. In G. Elliott, C. W. J. Granger, and A. Timmermann, editors, *Handbook of Economic Forecasting*. North Holland, 2006.

[8] J. Angel. Limit versus market orders. *School of Business Administration, Georgetown University*, 1994.

[9] F. Bandi and J. Russell. Microstructure noise, realized variance, and optimal sampling. *Review of Economic Studies*, 75(2):339–369, 2008.

[10] O. Barndorff-Nielsen, S. Kinnebrock, and N. Shephard. Measuring downside risk - realised semivariance. *CREATES Research Paper No. 2008-42*, 2008.

[11] D. Bertsimas and A. Lo. Optimal control of execution costs. *Journal of Financial Markets*, 1:150, 1998.

[12] F. Black. Towards a fully automated exchange, part i. *Financial Analysts Journal*, 27:29–34, 1971.

[13] B. Bouchard, N. Dang, and C.A. Lehalle. Optimal control of trading algorithms: a general impulse control approach. *Technical Report*, 2009.

[14] J.P. Bouchaud, J.D. Farmer, and F. Lillo. How markets slowly digest changes in supply and demand. *SSRN*, 2004.

[15] B. Brandt and K. Kavajecz. Price discovery in the U.S. treasury market: The impact of orderflow and liquidity on the yield curve. University of Pennsylvania Working Paper, 2002.

[16] S. Briese. Commitments of traders as a sentiment indicator. *Technical Analysis of Stocks and Commodities*, 8(5):199–204, May 1990.

[17] C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2:121–168, 1998.

[18] CFTC. Backgrounder. *Commodity Futures Trading Commission*, 2007.

[19] A. Chaboud, B. Chiquoine, E. Hjalmarsson, and C. Vega. Rise of the machines: Algorithmic trading in the fireign exchange market. *Board of Governers of the Federal Reserve System International Finance Discussion Papers*, 980, 2009.

[20] L Chan and J. Lakonishok. The behavior of stock prices around institutional trades. *Journal of Finance*, 50(4):1147–1174, 1995.

[21] T. Chordia, R. Roll, and A. Subrahmanyam. Market liquidity and trading activity. *Journal of Finance*, 56(2):501–530, 2001.

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

[23] CME Group. Statement on the joint cftc sec report regarding the events of may 6. October 2010.

[24] M. M. Dacorogna, R. Gençay, U. A. Müller, R. B. Olsen, and O. V. Pictet. *An Introduction to High Frequency Finance*. Academic Press, San Diego, 2001.

[25] F. Diebold, A. Hickman, A. Inoue, and T. Schuermann. Converting 1-day volatility to h-day volatility: Scaling by $\sqrt{h}$ is worse than you think. *Working Paper*, 1996.

[26] J. Durbin and S. Koopman. *Time Series Analysis by State Space Methods*. Oxford University Press, Oxford, 2001.

[27] D. Easley, M. de Prado, and M. O'Hara. The microstructure of the 'flash crash': Flow toxicity, liquidity crashes and the probability of informed trading. *Working Paper*, 2010.

[28] F.R. Edwards and M.S. Canter. The collapse of metallgesellschaft: Unhedgeable risks, poor hedging strategy or just bad luck? *Journal of Futures Markets*, 15, 1995.

[29] S. Fagan and R. Gencay. Health care decision making and media coverage: The case of in.uenza vaccination. *Working Paper*, 2010.

[30] E.F. Fama and K.R. French. Commodity future prices: Some evidence on forecast power, premiums, and the theory of storage. *Journal of Business*, 60(1):55–73, Jan 1987.

[31] X. Gabaix, P. Gopikrishnan, V. Plerou, and H.E. Stanley. Institutional investors and stock market volatility. *Quarterly Journal of Economics*, 121:461–504, 2006.

[32] J. Gatheral. No-dynamic-arbitrage and market impact. *Working Paper*, 2009.

[33] R. Gibson and E. Schwartz. Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance*, 45(3):959–976, Jul 1990.

[34] S. Goldfeld and R. Quandt. A markov model for switching regressions. *Journal of Econometrics*, 1:3–16, 1973.

[35] R. Goyenko, C. Holden, and C. Trzcinka. Do liquidity measures measure liquidity? *Journal of Financial Economics*, (forthcoming), 2008.

[36] J. Hamilton. A new approach to the economic analysis of nonstationary time series subject to changes in regime. *Econometrica*, 57:357–384, 1989.

[37] J. Hamilton. Anaylsis of time series subject to changes in regime. *Journal of Econometrics*, 45:39–70, 1990.

[38] J. Hamilton. *Time Series Analysis*. Princeton University Press, Princeton, 1994.

[39] L.P. Hansen, J. Heaton, and A. Yaron. Finite-sample properties of some alternative gmm estimators. *Journal of Business and Economic Statistics*, 14:262–280, 1996.

[40] L. Harris. Optimal dynamic order submission strategies in some stylized trading problems. *Financial Markets, Institutions and Instruments*, 7:1–76, 1998.

[41] L. Harris and J. Hasbrouck. Market vs. limit orders: The superdot evidence on order submission strategy. *Journal of Financial and Quantitative Analysis*, 31:213–231, 1996.

[42] J. Hasbrouck. check reference from how markets slowly digest changes in supply and demand. 1991.

[43] J. Hasbrouck and G. Saar. Technology and liquidity provision: The blurring of traditional definitions. *Journal of Financial Markets*, 12:143172, 2009.

[44] J. Hasbrouck and G. Saar. Low latency trading. *Johnson School Research Paper Series*, 35, 2010.

[45] J. Hasbrouck and G. Sofianos. The trades of market makers: An empirical analysis of NYSE specialists. *Journal of Finance*, 48(5):1565–1593, 1993.

[46] Hausman, Lo, and MacKinlay. An ordered probit analysis of transaction stock prices. *Journal of Financial Economics*, 31(3):319–379, June 1992.

[47] T. Hendershott, C. Jones, and A. Menkveld. Does algorithmic trading improve liquidity. *Journal of Finance*, Forthcoming, 2010.

[48] T. Hendershott and R. Riordan. Algorithimic trading and information. *Working Paper*, 2009.

[49] C. Hopman. Do supply and demand drive stock prices? *Quantitative Finance*, 2006.

[50] R. Huang, J. Cai, and X. Wang. Inventory risk-sharing and public information-based trading in the treasury note interdealer broker market. University of Notre Dame Working Paper, 2001.

[51] R. Iati. High frequency trading technology. *TABB Group*, 2009.

[52] T. Joachims. Text categorization with support vector machines: Learning with many relevant features. *Proceedings of the Tenth European Conference on Machine Learning*, pages 137–142, 1998.

[53] D. Keim and A Madhavan. The upstairs market for large-block transactions: analysis and measurement of price effect. *Review of Financial Studies*, 9(1):1–36, 1996.

[54] C. Kim. Dynamic linear models with markov switching. *Journal of Econometrics*, 60:1–22, 1994.

[55] C. Kim and C. Nelson. *State Space Models with Regime Switching*. MIT Press, Cambridge, 1999.

[56] A. Kirilenko, A. Kyle, M. Samadi, and T. Tuzun. The flash crash: The impact of high frequency trading on an electronic market. January 2011.

[57] S. Koopman, N. Shephard, and J. Doornik. Statistical algorithms for state space models using ssfpack 2.2. *Econometrics Journal*, 2:113–166, 1999.

[58] A. Kraus and H. Stoll. Price impacts of block trading on the new york stock exchange. *Journal of Finance*, 27:569–588, 1972.

[59] C. Lee and M. Ready. Inferring trade direction from intraday data. *Journal of Finance*, 46:733–746, 1991.

[60] A. Lo, A. MacKinlay, and J. Zhang. Econometric models of limit order execution. *Journal of Financial Economics*, 65:31–71, 2002.

[61] A.S. Mello and J.E. Parsons. Maturity structure of a hedge matters: Lessons from the metallgesellschaft debacle. *Journal of Applied Corporate Finance*, 7, 1995.

[62] T.M. Mitchell. *Machine Learning*. McGraw-Hill, New York, 1997.

[63] A. Obizhaeva and J. Wang. Optimal trading strategy and supply/demand dynamics. *Discussion Paper MIT Sloan School of Management*, 2005.

[64] R. Quandt. The estimation of parameters of linear regression systems obeying two separate regimes. *Journal of the American Statistical Association*, 55:873–880, 1958.

[65] R. Roll. A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance*, 39:1127–1139, 1984.

[66] D. Sanders, K. Boris, and M. Manfredo. Hedgers, funds, and small speculators in the energy futures markets: An analysis of the cftc's commitments of traders reports. *Finance Research Letters*, 26(3):425–445, 2004.

[67] SEC and CFTC. Findings regarding the market events of may 6, 2010: Report of the staffs of the cftc and sec to the joint advisory committee on emerging regulatory issues. September 2010.

[68] M. Spiegel. Patterns in cross market liquidity. *Finance Research Letters*, 5:2–10, 2008.

[69] P. Tetlock, M. Saar-Tsechansky, and S. Macskassy. More than words: Quantifying language to measure firms' fundamentals. *Journal of Finance*, LXIII(3):1437–1467, June 2008.

[70] N. Torre. *BARRA Market Impact Model Handbook*. BARRA Inc., Berkeley, 1997.

[71] F. Vaglica, F. Lillo, E. Moro, and R. Mantegna. Scaling laws of strategic behavior and size heterogeneity in agent dynamics. *Physical Review E.*, 77, 2008.

[72] D. Vayanos and J. Wang. Liquidity and asset prices: A unified framework. *Technical Report 15215, NBER*, 2009.

[73] C. Wang. Investor sentiment, market timing, and futures returns. *Applied Financial Economics*, 13(12):891–898, 2003.

[74] E. Zivot, J. Wang, and S. Koopman. *State Space Models in Economics and Finance Using SsfPack in S+FinMetrics*. A. Harvey, S.J. Koopman, and N. Shephard (eds.), *Unobserved Components Models*. Camebridge University Press, Cambridge, 2004.