INTRADAY MARKET DYNAMICS

BY

YASHAR HEYDARI BARARDEHI

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Graduate College of the University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Doctoral Committee:

Professor Dan Bernhardt, Chair
Associate Professor Ryan J. Davies
Professor Timothy C. Johnson
Professor Roger Koenker
Abstract

The revolutionary technological and regulatory changes in financial markets over the first few years of the new millennium have radically altered trading routines and strategies. Algorithms have taken over trade executions in an environment where interactions between virtual traders happen faster than blink of an eye. New trading strategies of long-term investors, e.g. institutional investors, have moved from one of submitting a few large orders to one of finely splitting orders over time and across trading venues in order to minimize their market impacts. Instead if human decisions, instructions that algorithms follow in order to locate liquidity, arbitrage opportunities, pattern detection, etc. determine the size and timing of transactions. As a result, individual transactions are far from reflecting economic decisions, and classical models of market microstructure may not be used to describe phenomena at transaction level.

I develop a novel aggregation approach that accounts for features modern markets; for a given stock, I identify successive sequences of transactions where cumulative dollar volume of each sequence is a fraction of previous month’s market-capitalization plus a fixed dollar-amount. Time durations of these trade sequences measure trading activity, and the corresponding price changes reflect market impacts of a fixed dollar volume traded at variable intensities. With this approach I (a) control for the temporal dependence across individual transactions induced by dynamic order-splitting, (b) finely isolate different market conditions, e.g. volume spikes from low trading activity,
(c) tell apart trading activity from trading volume, (d) reduce the effect of odd-lots bias that exists at transaction level, and (e) provide a measure of trading activity that helps us study intraday dynamics of trading activity and prices.

I first show that, for most stocks, price impacts of fixed dollar-positions significantly fall in trading activity. But price impacts and trading activity, on average, are endogenously determined: trading activity rises when liquidity (depth near good prices) is unusually high which presents itself as small price impacts. I then show that one can predict this variation using a simple instrument. Moreover, the relationships between price impacts (trading costs) and instrumented trading activity are very similar across differently-sized stocks post 2006, suggesting greater cross-stock homogeneity post RegNMS. In sharp contrast, greater heterogeneity obtains if one examines the levels of price impacts (trading costs): smaller (less liquid) stocks became less liquid post 2007, but the opposite holds for larger (more liquid) stocks. Using a CAMP that includes four Fama-French factors and key stock characteristics, I show that this divergence in liquidity is translated to greater liquidity premia post financial crisis. I findings indicate that the massive changes in the design of markets did not led to uniform improvements in stock liquidity and that the asymmetric evolution of liquidity across different stocks affected investment decisions.

I then begin to investigate the intraday dynamics of trading activity and price movements by contrasting two separate cases of changes in trading activity: I capture a relative increase in trading activity by a pair of successive trade sequences whose first sequences has a longer time durations—the opposite pattern reflects a decline in trading activity. I show that, surprisingly, increases in trading activity are associated with return momentum, but declines in trading activity are associated with price reversals. Return momentums are stronger when starting/concluding activity levels are higher and signed trades are less balanced. In sharp contrast, price reversals are
stronger when starting/concluding activity levels are lower and signed trades are more balanced. I conclude that these patterns are liquidity driven, e.g. price reversals of falling activity reflects rewards to liquidity provision after a phase of high activity. I then document more interesting time of day patterns: while increases in trading activity are least likely in earlier trading hours, return momentum of rising activity is strongest at these times; similarly, while activity decrease are least likely near close price reversals of falling activity are strongest in later trading hours. These findings highlight the highly variable nature of trading over the course of trading day. Earlier hours witness execution of overnight trading decisions that raise trading activity and persistent price impacts. Later trading hours, however, feature lower competition to provide liquidity since traders target flat closing positions; thus greater rewards to liquidity provision in expected.

I conclude my work by trying to model the dynamic structure of trading activity in the form I measure it. I employ the ACD models of Engle and Russell (1998) that were designed to model the time durations between individual transactions (inter-transaction durations). In todays markets, however, individual transactions are hard to reconcile with economic behavior. Thus, estimates of ACD models or any other dynamic structure that utilized inter-transaction durations have limited economic interpretations. An important contributions of my work is to introduce an alternative input to ACD models that fit features of modern financial markets and can provide a basis for economic interpretations. Moreover, my approach indirectly addresses other computations and statistical challenges one would face dealing with inter-transaction durations. Performing stock-year specific estimates of ACD models, I identify several interesting routes for future research in the fields of empirical market microstructure and financial econometrics.
To my beloved family...
Many individuals have contributed to my academic life whose most recent output is this document. I would like to extend my gratitude to every single one of them knowing that I will unintentionally exclude some of these individuals to whom I apologize.

First of all, I am grateful to my advisor, Dan Bernhardt, who has been inspiring me over the past five years. His commitment to train advisees well is one of a kind, and it is a key factor to one of his many significant contributions to the profession. I would not be able to develop my research, had he not provided generous help and guidance.

I also would like to thank members of my dissertation committee, Ryan J. Davies, Timothy C. Johnson, and Roger Koenker for their valuable advice. Each of these scholars are top-quality experts in their fields of study, and following their comments and suggestions played a significant role in shaping my work.

I would like to thank Jiaying Gu, Adam Calrk-Joseph, Praveen Kumar, Stefan Negal, George Pennacchi, Ronni Sadka, Mao Ye, and the seminar participants at the 2014 FMA annual meeting and the departments of Finance and Economics at the University of Illinois at Urbana-Champaign for helpful discussions and comments.

I must thank people who first taught me how to communicate thoughts: my parents. They
were not only my first teachers; they literally devoted their lives to provide for my wellbeing and education. I owe them my life. I highly appreciate every bit of their efforts, hoping that I can somehow pay back at least a tiny portion of my debt.

Finally, I recognize huge contributions of a person who devoted the best days of her life to stand by me while I was in the PhD program. My wonderful wife, Mina Kamouie, spent seven years of her life very patiently next to me. I had no chance of completing this degree in her absence; her words and attitude were my last resorts in times of crisis when nothing else would cheer me up. Undoubtedly, she gets much credit for any academic achievement I have made or will make in my life. I simply cannot thank her enough!
## Contents

List of Tables ................................................................. x

List of Figures ............................................................... xv

Chapter 1  Trading costs and priced illiquidity in high frequency trading markets 1

I  Introduction ............................................................ 1
   A  Related Literature ................................................. 8

II  Measuring Activity .................................................. 12
   A  Data ................................................................. 16
   B  Overview of activity measure ................................. 18

III Analysis ................................................................. 20

IV  Priced liquidity and changing illiquidity premia ................. 40

V  Conclusion .............................................................. 50

Chapter 2  Intraday Market Dynamics ................................. 52

I  Introduction ............................................................ 52

II  Reconstructing comparable trade sequences ..................... 60

III Data .................................................................... 64

IV Results ..................................................................... 65
   A  Changes in stock-specific market activity and the correlation structure of returns 68
   B  Information or liquidity? .......................................... 75
   C  Time-of-day and economic decisions .......................... 97
   D  RegNMS and the correlation structure of price dynamics 103
Chapter 3 ACD Models in Modern Markets

I Introduction ................................................................. 115
II Durations ................................................................. 119
III Data ................................................................. 122
IV Model ................................................................. 123
V Results ................................................................. 125
VI Future work: to believe or to not believe in WACD? ........................................ 137
   A To believe in WACD .............................................. 137
   B To not believe in WACD ........................................ 139

Appendix A ................................................................. 142

Appendix B ................................................................. 152

References .............................................................. 159
List of Tables

1.1 Year-by-year summary statistics of the final sample. Closing share prices, book-to-market ratios, and log of market capitalizations are reported as the medians of their monthly values. $-spreads are the median time-weighted dollar bid-ask spread at the NBBO. The last column reports the median number of trade sequences (each with an aggregate value of at least $V_{j,t}$) in a year across stocks. 18

1.2 Trading costs and instrumented trading activity. For each stock-year observation, the absolute return realized over a trade sequence is regressed on the corresponding instrumented trading activity percentile. The empirical distributions of these point estimates within each size decile are used to draw statistical inference. The sampling distribution is assumed to follow a student $t$ distribution (panel A). Observations are pooled across stocks within a size decile. Absolute return is regressed on the corresponding instrumented trading activity percentile; partial regression outputs are reported (panel B). The point estimates give the amount by which trading costs change as instrumented trading activity rises from 1st percentile to the 99th. Results are reported for 2001–2012, and subperiods 2001–2006 and 2007–2012 37

1.3 Augmented CAPM including trade-time liquidity measure. Estimation results for (1.6) using the trimmed sample, where the error term allows for stock fixed effects. Standard errors are clustered at the stock level. Month dummies capture any common variation in excess returns. $\beta(r_{t}^{mkt} - r_{t}^{f})$ is the product of a firm’s market beta and excess monthly return based on monthly holding period returns from CRSP and 1-month treasury bills. $BM_{t-1}$ is the previous period’s Book-to-Market ratio. $Ln(M_{t-1})$ is the natural logarithm of the firm’s market capitalization at the end of previous month. $BBD_{t-1}$ is the previous month’s trade-time illiquidity measure based on $\theta = 0.04$. Symbols *, **, and *** denote significance at 10%, 5%, and 1% levels, respectively. Robust standard errors are reported in parentheses. Premium is the product of the $BBD$ coefficient and the median value of $BBD$. 43
Augmented CAPM using standard illiquidity measures. The table reports panel estimates of (1.6) using AML, AMH, PSP, and DSP. Month dummies capture common variation caused by market condition changes, i.e., by common variation in excess returns over time. Standard errors are clustered at the stock level. Stock fixed effects capture any fixed heteroskedasticity. Observations in the top 1% tail of the relevant illiquidity measure are excluded. Symbols *, **, and *** denote significance at 10%, 5%, and 1% levels, respectively.

Augmented CAPM estimates for orthogonal decompositions. The table reports panel estimation results for (1.6) using the full sample. Standard errors are clustered at the stock level. Stock fixed effects capture any fixed heteroskedasticity. Month dummies capture any common variation caused by market condition changes. We separately present results that contrast BBD against each alternative measure AML, AMH, and PSP. Durations are generated based on $\theta = 0.04$. Observations in the upper 1% tail of either BBD or $F \in \{AML, AMH, PSP\}$ in a month are excluded. Robust standard errors are reported in parentheses. Symbols *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Subsample estimation of augmented CAPM with different liquidity measures. Panel estimation results of (1.6) for three subsamples: April 2001–Dec 2007, Jan 2008–Dec 2009, and Jan 2010–Dec 2012. Standard errors are clustered at the stock level. Stock fixed effects are introduced to capture any fixed heteroskedasticity. Month dummies are included to capture any common variation caused by market condition changes. The first column for a given sub-sample present results for our BBD measure based on $\theta = 0.04$ computed over all trade sequences. The other three columns of a panel present estimation results for the standard liquidity measures, AML, AMH, and PSP. Observations in the upper 1% tail of the relevant liquidity measure are excluded. Robust standard errors are reported in parentheses. The symbols *, **, and *** denote significance at 10%, 5%, and 1% levels, respectively.

Sample mean and confidence intervals of correlation parameters, isolating increases from reductions in stock-specific market activity. Model 2.5 is estimated each year, stock by stock, on subsamples representing increases in activity (yielding $\hat{\rho}_{1y}$) and decreases in activity (yielding $\hat{\rho}_{1y}$). Sample means and 99% confidence intervals of correlation parameters ($\hat{\rho}$ in panel A and $\hat{\rho}$ in panel B) are reported by size group in periods 2001–2006 and 2007–2012. The cumulative trade value over each trade sequence is $80,000 plus 0.025% of market-capitalization.

Sample mean and confidence intervals of correlation parameters, isolating increases from reductions in stock-specific market activity by year. Model 2.5 is estimated each year, stock by stock, on subsamples representing increases in activity (yielding $\hat{\rho}_{1y}$) and decreases in activity (yielding $\hat{\rho}_{1y}$). Sample means and 99% confidence intervals of correlation parameters ($\hat{\rho}$ and $\hat{\rho}$) are reported by year. The cumulative trade value over each trade sequence is $80,000 plus 0.025% of market-capitalization.
2.3 Sample means and confidence intervals of correlation parameter estimates, isolating increases from reductions in activity by starting stock-specific market activity level. Model 2.5 is estimated year-by-year, stock-by-stock, on subsamples representing increases in activity (yielding $\hat{\rho}_{jy}$) and decreases in activity (yielding $\hat{\rho}_{jy}$). Sample means and 99% confidence intervals of the correlation parameter are reported by starting activity group. Panel A reports the sample mean confidence intervals correlation parameters given no restrictions on activity changes. Panel B reports sample means and confidence intervals for activity changes that are less than twenty activity percentiles.

2.4 Sample means and confidence intervals of correlation parameter estimates, isolating increases from reductions in activity by concluding trading activity level. Model 2.5 is estimated year-by-year, stock-by-stock, for each activity quintile, on subsamples representing increases in activity (yielding $\hat{\rho}_{jy}$) and decreases in activity (yielding $\hat{\rho}_{jy}$). Sample means and 99% confidence intervals of the correlation parameter are reported by concluding activity group. Panel A reports the sample mean confidence intervals correlation parameters given no restrictions on activity changes. Panel B reports sample means and confidence intervals for activity changes that are less than twenty activity percentiles.

2.5 Conditional likelihood of return persistence as a function of stock-specific market activity level and the size of activity change. For each stock-year, pairs of successive trade sequences are grouped into quintiles of starting/concluding activity. For each activity quintile, observations are classified into groups of changes in activity percentile: percentile changes are categorized into $(-20, -10)$, $(-10, 0)$, $(0,10)$, and $(10,20)$ to reflect large decreases, small decreases, small increases, and large increases, respectively. Relative frequencies of successive returns of same sign are computed at the stock level. Reported conditional likelihoods reflect averages of such relative frequencies across stocks and time by category. (*) symbolizes the categories whose results are subject to potential selection biases.

2.6 Sample means and confidence intervals of correlation parameter estimates, isolating increases from reductions in activity at highest and lowest levels of signed trade imbalance by stock size decile. Model 2.5 is estimated year by year, stock by stock, for the highest and lowest starting/concluding imbalance quintiles, on subsamples representing increases in activity (yielding $\hat{\rho}_{jy}$) and decreases in activity (yielding $\hat{\rho}_{jy}$). Increases and decreases in activity that exceed twenty percentiles are excluded. Sample means and 99% confidence intervals of the correlation parameter are reported by stock size decile. Panel A (Panel B) reports the sample means and confidence intervals of correlation parameters at different starting (concluding) signed trade imbalance levels.

2.7 Mean relative frequency of increases/decreases in trading activity at different time-of-day windows by size decile. For each stock year, pairs of successive trade sequences are categorized by the start time of the second sequence into three time windows: early, 9:30:00AM–11:30:00AM; mid-day, 11:30:00AM–2:00:00PM; and late, 2:00:00PM–4:00:00PM. The relative frequencies of increases and decreases in activity are computed across time windows for each stock-year. Relative frequencies are averaged across stocks and years by size decile.
2.8 Sample mean and 99% confidence intervals of of $\rho_{j,y}$ estimates that distinguish increases from decreases in activity at different time-of-day windows 

by stock size decile Model (2.5) is estimated year-by-year, stock-by-stock by time-of-day window: early, 9:30:00AM–11:30:00AM; mid-day, 11:30:00AM–2:00:00PM; and late, 2:00:00PM–4:00:00PM. Model (2.5) is estimated year-by-year, stock-by-stock, for each time window, on subsamples representing decreases in activity (yielding $\hat{\rho}_{jy}$, Panel A) and increases in activity (yielding $\tilde{\rho}_{jy}$, Panel B). Activity changes are less then twenty percentiles. Stocks are sorted by their market-capitalizations at the beginning of each year into size deciles. Sample means and confidence intervals are computed by stock size decile.

2.9 Sample mean and 99% confidence intervals of of $\rho_{j,y}$ estimates that distinguish increases from decreases in activity at different time-of-day windows over time Model (2.5) is estimated year-by-year, stock-by-stock by time-of-day window: early, 9:30:00AM–11:30:00AM; mid-day, 11:30:00AM–2:00:00PM; and late, 2:00:00PM–4:00:00PM. Model (2.5) is estimated year-by-year, stock-by-stock, for each time window, on subsamples representing decreases in activity (yielding $\hat{\rho}_{jy}$, Panel A) and increases in activity (yielding $\tilde{\rho}_{jy}$, Panel B). Sample means and confidence intervals are computed year-by-year.

3.1 Overview of transaction sequence duration: The median time duration of each stock's transaction sequences is computed annually. Each year, stocks are grouped into size deciles based on market-capitalizations at the end of the previous year; size decile 1 contains stocks with smallest market-caps and size decile 10 contains those with largest. The second column of the table reports the cross-stock averages of stock-specific median durations, the third column reports the inter-quartile range of stock-specific median durations, and the last column reports the inter-quartile range of stock-specific median durations.

3.2 Number of insignificant estimates: WACD(2,2)s are fit on a stock-year basis. Reported are the number of cases where a particular parameter in an ACD fit is insignificant at 5% level. Durations measure the time it takes for a sequence of transactions worth $80,000 plus 0.025% of market capitalization. Durations are adjusted for diurnal patterns and overnight regime changes according to 3.4.

3.3 Predictability of duration and standardized duration series: $y_{jt} = \pi_0 + \sum_{i=2}^{20} \pi_i I_{jt} + u_{jt}$ is estimated using the stock-year specific series if $\tilde{x}_{jt}$ and $\epsilon_{jt}$. Mean point estimates and mean t-statistics of $\{\pi_2, \ldots, \pi_{20}\}$ for the two series $\tilde{x}_{jt}$ and $\epsilon_{jt}$ are reported. The last row of the table presents the mean regression F-statistics of the two cases.
A.1 Signed trade imbalances versus trading activity and instrumented trading activity. For each stock-year observation, the absolute return realized over a trade sequence is regressed on the corresponding activity percentile (Panel A) or instrumented activity percentile (Panel B). The empirical distributions of these point estimates within each size decile are employed to draw statistical inference. The sampling distribution is assumed to follow a student $t$ distribution (“Stock level estimates”). Observations are pooled across stocks within a size decile. Absolute return is regressed on the corresponding activity percentile; partial regression outputs are reported (“Size group level estimates”). The point estimates reflect the amount by which trade imbalances change as activity (instrumented activity) rises from lowest to highest percentile.

A.2 Illiquidity coefficients in the two sub-periods, after orthogonal decomposition. The table reports panel estimation partial results of (1.6) for two subsamples, April 2001–Dec 2006 and Jan 2007–Dec 2012. Standard errors are clustered at the stock level. Stock fixed effects capture any fixed heteroskedasticity. Month dummies capture common variation caused by market condition changes. The coefficients of $z_{Ft-1}$ and $\tilde{z}_{Ft-1}$ are reported, the version of $BBD$ based on all trade sequences is contrasted against the three standard liquidity measures. Observations in the upper 1% tail of the relevant liquidity measures are excluded. Robust standard errors are reported in parentheses. The symbols *, **, and *** denote significance at 10%, 5%, and 1% levels, respectively.

B.1 Sample means and confidence intervals of correlation parameter estimates, isolating increases from reductions in trading activity in most and least active markets by stock size decile. Model (2.5) is estimated year by year, stock by stock, for the highest and lowest starting/concluding activity quintiles, on subsamples representing increases in activity (yielding $\hat{\rho}_{ij}$) and decreases in activity (yielding $\tilde{\rho}_{ij}$). Increases and decreases in activity that exceed twenty percentiles are excluded. Sample means and 99% confidence intervals of the correlation parameters are reported by stock size decile. Panel A (Panel B) reports the sample means and confidence intervals of the correlation parameters at different starting (concluding) activity levels.
## List of Figures

1.1 **An illustration of how we measure the durations of trade sequences with an aggregate value of at least $V_{j,t}$.** ........................................... 13

1.2 **Trading activity by year and by firm size.** Average median duration (in minutes) of a sequence of trades with a cumulative aggregate value of at least 0.04% of a firm’s market capitalization. Median durations are calculated on a stock-by-stock basis. The average is then computed for stocks in a size decile each year. Size decile 1 contains the smallest firms; decile 10 contains the largest. ........................................... 19

1.3 **Distribution of durations of trade sequences for representative stocks.** Cumulative distribution functions of durations (in minutes) for International Business Machines (IBM), Archer Daniels Midland (ADM), and Hanger (HGR) from January 1, 2001 to December 31, 2012. Durations are based on the elapsed time for a sequence of trades with a cumulative aggregate value of at least 0.04% of the firm’s market capitalization. .................................................. 20

1.4 **Average signed dollar-volume and average number of trades by trading activity level.** Average signed trade imbalance and average number of trades for trade sequences with cumulative value of 0.04% of market capitalization in 2001–2012 versus activity level for the three deciles of smallest and largest stocks. Reported average signed trade imbalance and average number of trades are, respectively, the average of firm-specific annual mean estimated signed trade imbalance and number of trades (realized over trade sequences), computed at a trading activity level. Firm-specific means at a given trading activity level are averaged over all firms in a size decile, over the entire sample period. ........................................... 21

1.5 **Average absolute return and return volatility by duration group.** Market impacts of $\theta = 0.04\%$ of market capitalization in 2001–2012 versus activity level for the three deciles of smallest and largest stocks. The reported average price impact is the average of firm-specific monthly means of absolute returns (realized over trade sequences), computed per duration group. Reported average return volatility is the average of firm-specific monthly standard deviations of returns (realized over trade sequences), computed per duration group. The averages of means and standard deviations are taken over all firms in a given size decile, over the entire sample period by activity level. ........................................... 24
### 1.6 Estimated interpercentile ranges for realized returns by duration group.

Plot of the interpercentile ranges of the distribution of realized returns for different activity levels. The percentile statistics of returns (realized over trade sequences) are first computed stock-by-stock, annually, for each duration group, and then averaged across stocks/years. Finally, the differences between selected average percentile statistics ($\bar{r}_{.99} - \bar{r}_{.01}$, $\bar{r}_{.95} - \bar{r}_{.05}$, and $\bar{r}_{.75} - \bar{r}_{.25}$) are computed by activity group. Durations are based on $\theta = 0.04$. 

### 1.7 Empirical distribution of linear correlations and log-polynomial $F$-statistics by size decile.

In the top graphs, for each stock-year observation, we find the correlation between current duration and lagged number of trades. The kernel density of these correlation coefficients is estimated by size groups. In the bottom graphs, for each stock-year observation, the current log duration is regressed on a second-order polynomial of the lagged number of trades and the corresponding $F$-statistic is obtained. We then estimate, by size groups, the kernel density for the empirical distributions of the $F$-statistics. The vertical dashed lines indicate $F$-statistics of 10. Values exceeding 500 are omitted for visual clarity.

### 1.8 Average absolute return by predicted activity group.

Average absolute return of 0.04% of market-cap in 2001–2012 versus predicted trading activity for the deciles of smaller and larger stocks. The reported average absolute return is the average of firm-specific annual mean absolute return (realized over trade sequences), computed per predicted activity group. The average of means is taken across all firms in a given size decile, over the entire sample period by predicted activity level. Average absolute return versus activity is also displayed to highlight differences.

Stock-year observations with first stage regression $F$-statistics below 10 are dropped.

### 1.9 Temporal changes in average price impacts for small and large firms.

Stock-specific mean price impacts of 0.04% market-capitalization are computed quarterly. Then averages of these means are calculated quarterly across stocks in a market-cap quintile. Average price impacts and the corresponding 99% confidence intervals are presented for small (bottom 20% market-cap) and large (top 20% market-cap) stocks (top row)—the horizontal lines give medians of average price impacts are computed over the entire sample. The time series of price impacts normed by these sample medians are compared in the bottom row both in absolute and relative terms.

### 1.10 Temporal changes in levels and the ratio of the 20th and the 80th percentiles of $BBD$.

$BBD$ is constructed based on all trade sequences. The 20th and the 80th percentiles of $BBD$ are determined on a monthly basis. The ratio reflects the 80th percentile divided by the 20th percentile every month.

### 1.11 Monthly median values of different illiquidity measures over time.

The $BBD$ measure (based on $\theta = 0.04$) is calculated using observations from all trade sequences. Also shown is the high frequency Amihud measure ($AMH$), the low frequency Amihud measure ($AML$) and percentage bid-ask spreads ($PSP$).

### 2.1 Calculations of trade sequences and returns for successive trade sequences.

xvi
2.2 Time duration of trade sequences by year and by firm size. The figure reports the average median duration (in minutes) of a sequence of trades with a cumulative aggregate value of at least $80,000 plus 0.04% of a firm’s market capitalization. The median duration is first calculated on a stock-by-stock basis, and then the average is taken across stocks within a size decile in a year. Size decile 1 contains the smallest firms; decile 10 contains the largest.

2.3 Average return by activity group. Average return trade sequences with cumulative value of $80,000 plus 0.025% of market capitalization in 2001–2012 versus activity level by stock size categories. Reported average return is the average of firm-specific annual mean returns (realized over trade sequences), computed per activity group. The average of means is taken across all firms in a given size decile, over the entire sample period by activity group.

2.4 Empirical distributions of correlation parameter estimates and the corresponding t-statistics, isolating increases from decreases in stock-specific market activity. Model 2.5 is estimated year-by-year, stock-by-stock, on subsamples representing increases in activity (yielding $\hat{\rho}_{jy}$) and decreases in activity (yielding $\hat{\rho}_{jy}$). The top row reports the empirical distributions of $\hat{\rho}$ and $\hat{\rho}$, separately for stocks falling in the bottom 30%, middle 40%, and top 30% of beginning-of-the-year market-capitalizations, for the periods 2001–2006 and 2007–2012. The bottom row exhibits the corresponding t-statistics with vertical lines standing for the 95% critical values for relevant one-sided significance tests. The cumulative trade value over each trade sequence is $80,000 plus 0.025% of market-capitalization.

2.5 Empirical distributions of correlation parameter estimates for signed past returns, isolating increases from decreases in stock-specific market activity. I estimate model 2.5 each year, stock by stock, on subsamples representing increases in activity (yielding $\hat{\rho}_{jy}$) and decreases in activity (yielding $\hat{\rho}_{jy}$), conditioning on the sign of return realized over the first trade sequence, $r_{jy}(k - 1)$. The cumulative trade value over each trade sequence is $80,000 plus 0.025% of market-capitalization.

2.6 Empirical distributions of correlation parameter estimates that isolate increases from reductions in activity by starting stock-specific market activity level. Model 2.5 is estimated each year stock by stock, for each starting activity quintile, on subsamples representing increases in activity (yielding $\hat{\rho}_{jy}$) and decreases in activity (yielding $\hat{\rho}_{jy}$). The top row displays the empirical kernel densities of estimated correlations with no restrictions on activity changes. The bottom row reports similar densities whose estimates are based on activity changes that are less than twenty activity percentiles.

2.7 Empirical distributions of correlation parameter estimates that isolate increases from reductions in activity by concluding stock-specific market activity level. Model 2.5 is estimated year-by-year, stock-by-stock, for each concluding activity quintile, on subsamples representing increases in activity (yielding $\hat{\rho}_{jy}$) and decreases in activity (yielding $\hat{\rho}_{jy}$). The top row displays the empirical kernel densities of estimated correlation estimates with no restrictions on activity changes. The bottom row reports similar densities whose estimates are based on activity changes that are less than twenty activity percentiles.
2.8 Empirical distributions correlation parameter estimates that isolate decreases in activity by starting/concluding signed trade imbalance level for stock size groups. Model (2.5) is estimated year-by-year, stock-by-stock, controlling for concluding/start ing imbalance quintiles for the subsample of trade sequences associated with a decrease in activity (yielding $\hat{\rho}_{jy}$). Decreases in activity that exceed twenty percentiles are excluded. The top (bottom) row reports the empirical density estimates at the highest and lowest starting (concluding) signed trade imbalances.

2.9 Empirical distributions of correlation parameter estimates that isolate increases in activity by starting/concluding signed trade imbalance level for stock size groups. Model (2.5) is estimated year-by-year, stock-by-stock, controlling for concluding/start ing imbalance quintiles for the subsample of trade sequences associated with a increase in activity (yielding $\hat{\rho}_{jy}$). Increases in activity that exceed twenty percentiles are excluded. The top (bottom) row reports the empirical density estimates at highest and lowest starting (concluding) signed trade imbalance.

2.10 Sensitivity of current signed trade imbalance to past changes in activity by concluding stock-specific market activity level. The sample is decomposed by year, size decile, and concluding activity group—observations are pooled across different stocks. Within each category, signed trade imbalance is regressed on the past change in activity percentile, inversely weighting observations on each stock by its number of trade sequences that year. The point estimates and the associated t-statistics are reported by activity groups and sub-periods (top row). Similar estimates are obtained, isolating samples of increases and decreases in trading activity (bottom row).

2.11 Sensitivity of current price impact to past changes in activity by concluding stock-specific market activity level. The sample is decomposed by year, size decile, and concluding activity group—observations are pooled across different stocks. Within each category, price impact is regressed on the past change in activity percentile, inversely weighting observations on each stock by its number of trade sequences that year. The point estimates and the associated t-statistics are reported by activity groups and sub-periods (top row). Similar estimates are obtained, isolating samples of increases and decreases in trading activity (bottom row).

2.12 Empirical distributions of $\hat{\rho}_{jy}$ estimates that distinguish increases from decreases in activity at different time-of-day windows by stock size group. Model (2.5) is estimated year-by-year, stock-by-stock by time-of-day window: early, 9:30:00AM–11:30:00AM; mid-day, 11:30:00AM–2:00:00PM; and late, 2:00:00PM–4:00:00PM. Model 2.5 is estimated year-by-year, stock-by-stock, for each activity quintile, on subsamples representing increases in activity (yielding $\hat{\rho}_{jy}$, top row) and decreases in activity (yielding $\hat{\rho}_{jy}$, bottom row). Activity changes are less then twenty percentiles. Stocks are sorted by their market-capitalizations at the beginning of each year into bottom 30%, middle 40%, and top 30% size groups.
2.13 The temporal evolution of the commonality in the return correlations associated with increases and decreases in stock-specific market activity. Every quarter, I estimate the correlation structure of returns realized over two successive trade sequences for the sample of increases in activity and for the sample of reductions in activity, stock-by-stock. I then estimate the cross-stock correlations of these two correlation parameters quarterly.

2.14 The temporal evolution of mean change in proportion of buyer-initiated trading when stock-specific market activity falls by signed past returns. Samples of successive trade sequence pairs that represent reductions in activity are decomposed by signing the return realized over the first trade sequence of each pair (past return). To avoid selection bias, the reduction in activity must be smaller than twenty activity percentiles. For each stock, the mean change in the proportion of buyer-initiated trades, associated with a decline in activity after signed returns, is computed annually. Averages of those means are taken across stocks of a size group (bottom 30%, middle 40%, and top 30% of market-capitalizations at the beginning of each year) year-by-year.

2.15 The relation between concluding stock-specific market activity level and the change in the proportion of buyer-initiated trades as trading activity declines. Samples of successive trade sequence pairs that represent reductions in activity are decomposed by signing the return realized over the first trade sequence of each pair (past return). To avoid selection bias, the reduction in activity must be smaller than twenty activity percentiles. For each stock, the mean change in the proportion of buyer-initiated trades, associated with a decline in activity after signed returns, is computed by activity decile, annually. Averages of those means are taken across stocks of a size group (bottom 30%, middle 40%, and top 30% of market-capitalizations at the beginning of each year) and over time within the periods 2001–2006 and 2007–2012 by activity decile. Solid and dashed curves signify positive (negative) past returns, respectively.

2.16 The relation between concluding stock-specific market activity level and the change in the proportion of buyer-initiated trades as trading activity declines by time of day in 2007–2012. Samples of successive trade sequence pairs that represent reductions in activity are decomposed by signing the return realized over the first trade sequence of each pair (past return). To avoid selection bias, the reduction in activity must be less than twenty activity percentiles. For each stock, the mean change in the proportion of buyer-initiated trades, associated with a decline in activity after signed returns, is computed by activity decile and time of day, annually. Averages of those means are taken across stocks in a size group and over time within the period 2007–2012 by activity decile.

3.1 Calculations of successive transaction sequences of a given stock.

3.2 Empirical distributions of parameter estimates: WACD(2,2)s are fitted on a stock-year basis. The histograms reflect the empirical distributions of stock-year-specific estimates of parameters \(\{\alpha_1, \alpha_2, \beta_1, \beta_2, \omega, \gamma\}\). Durations measure the time it takes for a sequence of transactions worth $80,000 plus 0.025% of market capitalization. Durations are adjusted for diurnal patterns and overnight regime changes according to 3.4.
3.3 **Pair-wise correlations of parameters:** WACD(2,2)s are fitted on a stock-year basis. The scatter plots reflect the pair-wise associations between stock-year-specific estimates of parameters \( \{ \alpha_j^1, \alpha_j^2, \beta_j^1, \beta_j^2, \omega_j, \gamma_j \} \). Durations measure the time it takes for a sequence of transactions worth $80,000 plus 0.025% of market capitalization. Durations are adjusted for diurnal patterns and overnight regime changes according to 3.4.

3.4 **Shape parameter estimates and log-number of durations against stock size:** Stock-year specific shape parameter estimates \( \gamma_j \)'s and log-number of durations are plotted against stock size percentile. Size percentiles reflect annual normalized rank statistics that sort stocks from smallest to largest by market-capitalization each year—market-capitalizations are measured at the end of the previous year. Durations measure the time it takes for a sequence of transactions worth $80,000 plus 0.025% of market capitalization.

3.5 **Empirical distributions of AR(2) coefficients and their corresponding t-statistics:**
\[ y_{jt} = \delta_j^0 + \delta_j^1 y_{jt-1} + \delta_j^2 y_{jt-2} + u_{jt} \]

is estimated using the series of \( \hat{x}_{jt} \) and \( \epsilon_{jt} \) for each stock-year set of observations. The empirical distributions of the corresponding parameter estimates and t-statistics for the two series types are contrasted. Vertical short-dashed lines in the bottom row represent two-sided critical values at 0.1% significance level. Durations measure the time it takes for a sequence of transactions worth $80,000 plus 0.025% of market capitalization.

3.6 **Point estimates and t-statistics of mean hazard errors:** WACD(2,2)s are fitted on a stock-year basis. Using stock-year-specific estimates, hazard errors (ratios of empirical-to-predicted hazard) are computed for each \( \epsilon_{jt} \). Observations on each stock-year are sorted into twenty equally-sized groups of standardized duration \( (\epsilon_{jt}) \), and mean hazard errors are computed by standardized duration group per stock-year set of observations. The top row presents the averages of mean hazard errors at different levels of standardized duration; the bottom row reports the corresponding t-statistics for a null of average error equal to unity (the two horizontal lines stand for critical values at 0.1% significance level). Averages and t-statistics are calculated across stocks and years, controlling for stock size decile. Size deciles are formed at the beginning of each year by sorting stocks into ten equally-sized groups of market-capitalizations at the end of the previous year.

A.1 **Average signed trade imbalances versus predicted trading activity.** Average signed trade imbalances of trade sequences versus predicted trading activity for the deciles of smaller and larger stocks. The reported average signed-dollar volume is the average of firm-specific annual mean dollar-weighted proportion of buy- (sell-) oriented trades (realized over trade sequences), computed per predicted activity group. The average of means is taken across all firms in a given size decile, over the entire sample period by predicted activity group. Average signed trade imbalances versus trading activity is displayed to highlight differences. Stock-year observations with first stage regression \( F \)-statistics below 10 are dropped.
A.2 Average daily return volatility by activity group. Average return volatility of 0.04% of market capitalization in 2001–2012 versus activity level for the three deciles of smallest and largest stocks. The reported average daily return volatility is the average of firm-specific annually computed daily standard deviations of returns $S^d_i(y,j)$, computed per duration group. The average of daily standard deviations is taken across all firms in a given size decile, over the entire sample period, by activity level.

A.3 Empirical cumulative density functions of duration by size decile. Durations are computed based on $\theta = 0.04$ and their CDFs are presented for two subsets of large and small stocks, over the entire period (2001–2012), and the early (2001–2006) and later (2007–2012) sub-periods. To calculate the CDF of durations for stock $j$ in year $y$, every month, we first sort durations into 20 buckets, each containing 5% of observations. Then we pool buckets across different months. The median of these durations corresponds to the quantile statistic falling at the midpoint of that duration bucket. For example, the median in the pooled bucket of shortest 5% estimates quantile statistic 2.5%; that for next bucket estimates quantile statistic 7.5%; etc.

A.4 Average absolute return, return volatility, and average number of trades versus signed trade imbalance. Average absolute return, average return volatility, and mean number of trades for trade sequences versus signed trade imbalance for deciles of smaller and larger stocks. We first sort trade sequences into ten equally-sized groups of percent signed trades, every month. The reported averages are the sample averages of firm-specific annually-computed mean absolute return, return standard deviation, and mean realized depth (all assessed over sequences), computed per signed trade imbalance decile. The averages of means are taken across all firms in a given size decile, over the entire sample period by signed trade imbalance level.

B.1 Empirical distributions correlation parameter estimates that isolate decreases in activity by starting/concluding stock-specific market activity level for stock size groups. Every year, model (2.5) is estimated stock by stock, controlling for concluding/concluding trading activity quintiles. Estimates are obtained using the sample of trade sequences that are associated with a decrease in activity (yielding $\hat{\rho}_{jy}$). Decreases in activity that exceed twenty percentiles are excluded. The top (bottom) row reports the empirical density estimates at highest and lowest starting (concluding) activity levels.

B.2 Empirical distributions correlation parameter estimates that isolate increases in activity by starting/concluding stock-specific market activity level for stock size groups. Every year, model (2.5) is estimated stock by stock, controlling for concluding/concluding trading activity quintiles. Estimates are obtained using the sample of trade sequences that are associated with an increase in activity (yielding $\tilde{\rho}_{jy}$). Increases in activity that exceed twenty percentiles are excluded. The top (bottom) row reports the empirical density estimates at highest and lowest starting (concluding) activity levels.
B.3 Sensitivity of current signed trade imbalance to past changes in activity by concluding stock-specific market activity level. The sample is decomposed by year, size decile, and starting activity group—observations are pooled across different stocks. Within each category, signed trade imbalance is regressed on the past change in activity percentile, inversely weighting observations on each stock by its number of trade sequences that year. The point estimates and the associated t-statistics are reported by activity groups and sub-periods (first row). Similar estimates are obtained, isolating samples of increases and decreases in trading activity (second row).

B.4 Sensitivity of current price impact to past changes in activity by concluding stock-specific market activity level. The sample is decomposed by year, size decile, and starting activity group—observations are pooled across different stocks. Within each category, signed trade imbalance is regressed on the past change in activity percentile, inversely weighting observations on each stock by its number of trade sequences that year. The point estimates and the associated t-statistics are reported by activity groups and sub-periods (first row). Similar estimates are obtained, isolating samples of increases and decreases in trading activity (second row).
Chapter 1

Trading costs and priced illiquidity in high frequency trading markets

I Introduction

Prior to 1997, stock prices were quoted in eighths, and this tick size drove trading costs and limit order book depth. Today, the typical tick size is a penny, and depth near inside quotes has collapsed. The primary providers of market-making services are now algorithmic and high frequency traders (HFT henceforth) who submit many thousands of small, fleeting orders each trading day. As a result, average trade sizes, inside bid-ask spreads and depth are tiny.

To establish and unwind positions, institutional investors now employ specialized brokers who use sophisticated algorithms to split intended trading positions into many small orders that are dynamically submitted over time and across trading venues. This makes institutional “child orders” temporally dependent, a phenomena that HFT can profitably exploit, further raising institutional trading costs. Individual transactions and the corresponding price changes largely reflect execution processes, not institutional trading decisions. The price change associated with each small individual trade trivially contributes to trading costs, implying that traditional measures of trading costs and liquidity no longer describe concerns of institutional investors. Rather, institutional investors care about the cumulative price impacts of these trades that together comprise their larger intended trade positions.
The key contributions of our paper are to measure these trading costs, how they vary with the level of market activity and firm characteristics, and to show how these trading costs explain the cross-section of expected returns in asset pricing models. We identify the primitive economic forces underlying the variations in trading costs, trade sizes and signed trade imbalances at different trading activity levels, teasing out implications for theorists seeking to understand today’s markets. We also identify how the divergence of trading costs in the cross-section of stocks translates to greater liquidity premia post RegNMS.

We first develop a stock-specific measure of trading activity that serves as a tool to isolate different trading activity conditions. We then use it to measure how trading costs vary with the level of trading activity, firm size and time. Our findings indicate that, on average, endogenous consumption of liquidity, not information arrival, is the primary driver of trading activity and the associated price impacts. This inference highlights the importance of identifying when to begin to establish or unwind a position. We employ a simple instrument to control for this endogeneity, and show how an institutional investor can predict future stock-specific trading activity levels, and thereby optimally time trades.

We then establish that, over time, the shapes of trading cost-trading activity level relationships have become more similar across stocks—in recent years, as trading activity rises, trading costs decline in similar ways for most stocks—suggesting that markets of different stocks became more homogeneous in an important way. In other ways, however, the RegNMS and advent of algorithmic trading seem to have increased cross-stock heterogeneity: we find that although the trading costs of larger and more liquid stocks have fallen in recent years, those of smaller and less liquid stocks have not. Finally, we establish that this divergence in trading costs translates into post-financial crisis illiquidity premia that exceed pre-crisis levels.
O’Hara (2015) highlights the ways in which changes in the market make individual trades minimally relevant for existing microstructure models. Any measure of stock-specific trading activity must aggregate sufficiently to control for the temporal dependence of trades, but not by so much that one aggregates over very different market conditions. Aggregation is also necessary to deal with issues with available data sets\(^1\). The question becomes: how to aggregate? There are two popular approaches.

Amihud (2002) aggregates volume over a trading day to construct his price impact measure. One could modify his approach by aggregating over shorter (e.g., twenty minute) time intervals within a trading day, where greater volumes indicate greater activity. For our purposes, this approach does not work. Such aggregation gives rise to massive variation in trading volumes across time intervals. As a result, intervals with very low volume deliver noisy measures of market metrics such as returns, and represent far lower volumes than would be relevant for institutional investors. In addition, intervals with volume spikes aggregate over different market conditions. Furthermore, aggregating in this way means that relationships between price impacts and trading volumes conflate both activity and volume: it is unclear whether a positive relationship between price impacts and volume in a fixed time interval is due to higher activity, or to activity being linked to volume, which in turn is positively correlated with price volatility (Andersen and Bondarenko (2014)).

Easley et al. (2012) instead divide trading volume on each day into a fixed number of equally-large volume bundles (e.g., 50). Within a trading day, bundles that span longer time periods indicate lower activity. For our purposes, this approach also does not work: it leads to huge variation

\(^1\) O’Hara et al (2012) highlight that monthly TAQ data misses odd-lots that, on average, account for only 7% of trading volume, but 35% of price discovery. Braccini (2014) reports that 25% of trades coincide given millisecond time scales, which makes analysis of time series properties of trades challenging. Murvayev and Picard (2014) highlight issues of mechanical trade clusters. Holden and Jacobsen (2013) documents shortcomings of monthly TAQ relative to daily TAQ.
from one day to the next in volume bundle sizes, a variation that is compounded by the exploding trading volumes in recent years. Issues arise when trying to aggregate across days or over time for estimation purposes, because a low level of activity on one day is not comparable with a low level on another day.

To circumvent these issues, we develop a novel aggregation method that groups consecutive transactions together into a trade sequence with a fixed cumulative target dollar value (e.g., 0.04% of a firm’s market capitalization). Once a stock’s target dollar value is reached, we start over, grouping the next set of consecutive trades until their cumulative dollar value reaches its target value, continuing in this way iteratively over a year. We measure trading activity by the time it takes for a trade sequence to reach its target dollar value—shorter durations indicate more active markets. Controlling for the dollar value of a sequence preserves two key features: (1) it delivers comparable fundamentals measured over different trade sequences of a stock; (2) it lets us isolate the effects of activity from those of trading volume. By setting target dollar values neither too large nor too small, we address concerns about (a) aggregating over multiple activity levels due to excessively large targets, (b) mis-attributing variation in activity due to excessively small targets, and (c) the impact of missing odd-lots; and it delivers trade volumes of interest to institutional investors, objects that are economically and empirically tractable.

We first show that both trade size and signed trade imbalance rise with trading activity: on average, as the duration of a trade sequence for a stock shrinks, (1) fewer trades comprise the sequence and (2) trade becomes less balanced.\footnote{We use the Lee and Ready (1991) algorithm to identify the share of value-weighted trades in a trade sequence that is buyer- or seller-initiated (whichever is highest).} Going from least to most active markets, signed trade imbalances rise by roughly 20%, and for small stocks, mean trade sizes double.
It is not surprising that more active markets feature more aggressive trading. Classical models of speculation (Kyle (1985), Easley and O’Hara (1992), Glosten (1994)) predict precisely this: speculators with substantial private information will establish larger positions, submitting larger marketable orders that cause trading activity, signed trade imbalance, and trade sizes all to rise. But, a very different explanation presents itself. Liquidity provision may be state-varying—more depth may be available near inside quotes at some times than others. When the market is unusually liquid on one side, traders respond aggressively, submitting larger marketable orders that are filled at these “good” prices. As a result, the corresponding trade sequences have shorter durations, with fewer trades and less balanced signed trade. In contrast, with little depth, traders submit less aggressive orders (i.e., marketable orders that consume less liquidity), and submit more passive orders (i.e., limit orders that supply liquidity) to reduce price impacts. The result is lower trading activity, more balanced signed trade, and more trades comprising a trade sequence.

To distinguish between these two explanations, we look at price impacts. Models of speculation predict that more aggressive trading should lead to larger price impacts; but, if endogenous consumption of unusually high depth at “good” prices is the primary driver, price impacts should fall as activity rises. We measure price impacts in two ways: (1) the average absolute return over a trade sequence at a trading activity level; and (2) the average annual standard deviation of those returns. For small and mid-sized stocks, both price impact measures fall sharply as trading activity rises—declines on the order of 40%. Thus, even though volatility is positively associated with volume and activity, the opposite relationship holds once one controls for volume. In contrast, for larger stocks, price impacts peak at intermediate activity levels, and, indeed, the standard deviation of returns is smallest when markets are least active.

\[^{3}\text{See Bloomfield et al. (2005), Goettler et al. (2005), Hollifield et al. (2004), Hollifield et al. (2006) or Easley et al. (2012).}\]
The uniformly negative relationship between price impacts and activity for small and mid-sized stocks indicates that varying endogenous consumption and provision of liquidity is the key driver of variation in trading activity and price impacts, not information. In contrast, fat tails of the return distribution in active markets of large stocks suggests that, in unusual times, information arrival drives some higher activity, but, in normal times, activity mostly reflects variation in liquidity.

Advice to investors who seek to minimize trading costs to “trade when markets are liquid at good prices” has modest value. Thus, we proceed to isolate stable and predictable activity–trading cost relationships: we examine how price impacts vary with predicted trading activity.

To control for information arrival and the endogeneity of order composition, we use the number of trades filling the previous trade sequence to instrument for the time duration of the current trade sequence. The insight underlying this instrument is that liquidity persists over time—present and past states of the order book are positively correlated. The endogenous composition of current orders, however, hinges only on the contemporaneous state of the order book, which strongly co-varies with current activity and price impacts. Effects of past order book states on current price impacts only transmit via the current state of the order book and the dollar volume in a trade sequence is large enough to minimize the temporal dependence associated with endogenous order composition. Intuitively, the endogenous responses quickly offset imbalances in liquidity—when liquidity is high, it is consumed, and when it is low, it is provided. Appropriate aggregation is crucial for this result, since at ultra-high frequency, it takes some time for “unusual” liquidity to be absorbed, as Hirschey (2013) shows.

We provide extensive evidence that our instrument choice is a good one, and that signed trade imbalances only rise weakly with predicted trade activity. In sharp contrast, very stable and uniform
trading costs–predicted trading activity relationships obtain across all firm sizes. Instrumenting flattens the price impact/activity relationship for small stocks by roughly 75%; but for larger stocks, the price impact/activity relationship not only becomes monotone, it becomes steeper. The uniform reduction in average price impacts (of a trade sequence) of five to eight basis points as predicted market conditions rise from least active to most active levels are meaningful from the perspective of an institutional investor.

We next investigate how trading costs evolved between 2001 and 2012 as U.S. equity markets were transforming. We first show that in more recent years the relationship between predicted trading activity and trading costs became more similar across stocks—trading costs fall by comparable amounts as predicted trading activity rises. But, markets did not become uniformly more liquid: levels of price impacts of more liquid stocks fell post crisis, but those of less liquid stocks rose. As such, the narrower bid-ask spreads of recent years are a misleading indicator of changes in trading costs. These findings lead us to explore their asset pricing implications.

To create a measure of individual stock liquidity that is appropriate for today’s markets we scale our price impact measure by the dollar amount traded. Our illiquidity measure, denoted $BBD$, is a rolling average of the per-dollar price impact of trading a fixed proportion of a stock’s market capitalization that controls for trading activity level. $BBD$ is a trade-time analogue of Amihud’s (2002) temporal measure of liquidity that isolates the effects of trading activity from trading volume and controls for the variation in trading costs with activity.

Regardless of which trading activity levels are used to calculate $BBD$, it is priced in an augmented capital asset pricing model (CAPM), and indeed similar illiquidity premia obtain. We verify that $BBD$ better captures illiquidity considerations of investors than do standard illiquidity
measures—average percentage bid-ask spreads, Amihud’s per dollar price impact, and a high frequency version of Amihud’s measure constructed at hourly frequencies—especially, post RegNMS. To do this, given the high correlation in the measures, we employ the residuals from the regression of one measure on a second in the pricing regressions. The \textit{BBD} residual is always priced, whereas the other residuals are not.

We then show that illiquid stocks become more illiquid post RegNMS, and that liquid stocks became more liquid. Splitting the sample into pre-crisis, crisis, and post-crisis periods reveals that despite far higher trading volumes associated with high-frequency trading, illiquidity premia post-crisis exceed pre-crisis levels. That is, increased cross-stock dispersion in trading costs and illiquidity translates to greater dispersion in expected excess returns\footnote{In contrast, Ben-Rephael et al. (2015) find lower liquidity premia in 2000-2011 than in earlier sub-periods between 1964–1999. We compare their approach with ours in detail in Section IV, reconciling the findings.}

Our paper is organized as follows. We next discuss related research. Section II develops our trading activity measure. Section III shows how various fundamentals vary with activity. Section IV evaluates the illiquidity measures in an asset pricing model. Section V concludes. Chapter A explores calendar time relationships between volatility and stock-specific trading activity, and shows our results are robust to different error structure specifications.

A Related Literature

Trade time and market dynamics. Mapping financial market dynamics onto trade-time space is not a new idea. Dufour and Engle (2000) examine the duration between successive trades and show that price adjustments are faster when trading intensities are high (i.e., durations are short). They argue that high trade intensities reflect informed trading and conclude that these periods
represent periods of market illiquidity. Engle and Russell (1998) and Engle (2000) develop related models of autoregressive conditional duration (ACD) that provide semi-parametric estimates of trade intensities based on trade arrival rates. ACD models are largely designed to estimate the expected cost and time to execute a single order, submitted within a short time frame, not the cost of a series of orders submitted over a longer window.

In modern markets, the information content of one trade is small; and consecutive trades may represent a single trading decision, where a larger order is crossed against multiple smaller orders in the book. To account for these features, Gouriéroux et al. (1999) measure activity by the time it takes the market to execute an exogenous level of volume or dollar volume. They focus on the evolution of trading activity over the trading day on Paris Bourse for a single stock over 18 trading days, exploring the average time duration to unwind a fixed position at different times. They find activity peaks at times corresponding to the opening of the London Stock Exchange and NYSE. Easley et al. (2012) are similarly motivated to group consecutive trades across time according to the time required to trade an exogenous level of order flow. Within each group, they use the sign of price changes over one-minute periods to infer buyer-initiated versus seller-initiated trades, creating a measure of volume “toxicity”, VPIN, the high frequency analogue to their probability of informed trading (PIN) measure.

Feldhütter (2012) develops the notion of an imputed roundtrip trade, based on the empirical feature that trades in corporate bonds often occur infrequently, but when they do trade, there are often several trades in a short period of time. He argues that these trades are likely linked and

---

5Our findings indicate a tension between concerns about using and consuming liquidity, and identification of probabilities of informed trade off of a zero-profit limit order pricing condition. If liquidity provision is competitive, traders should always consume liquidity—it is because markets are sometimes more liquid than others that a trader sometimes consumes and sometimes supplies liquidity (Andersen and Bondarenko (2014)). Based on execution algorithms of ITG, O’Hara (2015) provides evidence of order-splitting strategies that employ variable weights of aggressive and passive orders at different levels of institutional trade intensity.
should be considered together when measuring trading costs.

**Relation between trading activity and liquidity.** There has been a long debate about the relation between activity and liquidity. Demsetz (1968) finds that more active stocks tended to be more liquid, and argues that more frequent trading is associated with low dealer costs. Lippman and McCall (1986) argue that active markets have a lower opportunity cost of searching and are thus more liquid. Jones (2002) and Fujimoto (2004) find no relation between liquidity measures and changes, or shocks, in turnover. Johnson (2008) develops a model in which volume is positively related to liquidity risk, rather than liquidity, and provides supporting evidence from U.S. government bond markets.

**Liquidity Measures.** Amihud and Mendelson (1986) show that when the standard tick size was one-eighth, bid-ask spreads were priced and were a good measure of liquidity. Other quote-based liquidity measures include effective bid-ask spreads, Roll’s (1984) measure, and Hasbrouck’s Gibbs estimate.⁶ Lesmond et al. (1999) argue that periods of zero returns may represent an absence of informed trading, periods where informed trading costs are high.

Variants of Kyle’s λ (Kyle (1985)) have also been used as liquidity measures (see, e.g., Glosten and Harris (1988), Brennan and Subrahmanyam (1996), or Pástor and Stambaugh (2003)). Bernhardt and Hughson (2002) estimate price impacts using a structural model whose equilibrium, unlike Kyle’s, explicitly incorporates orders that are not pooled.

Chordia et al. (2011), Angel et al. (2011), Kim and Murphy (2013), and others have questioned the accuracy and feasibility of many existing liquidity measures in a high frequency trading world with tiny spreads and little depth at the inside quote. Holden and Jacobsen (2014) document

---

⁶Effective bid-ask spreads measure the spread between the trade price and quote midpoint. Roll (1984) focuses on how the tick size induces negative autocorrelation in returns. Hasbrouck (2009) proposes a Bayesian platform to generate a Gibbs estimate of Roll’s effective trading cost indicator.
realized dollar spreads and share depth as low as 0.88¢ and 564 shares, respectively, for a sample of randomly selected stocks in 2008. Today, liquidity concerns do not revolve around avoiding minimum price ticks, so spreads do not capture liquidity from that perspective. Order-splitting makes consecutive trades dependent (Hasbrouck (2009), Easley et al. (2012)) and interpretations of estimates of $\lambda$ heroic—especially since TAQ data omit odd-lots, which now account for 35% of price discovery. Moreover, high trading volumes result in few instances of zero returns (Mazza (2013)). Further, the increased use of hidden orders makes it difficult to measure depth directly. Deuskar and Johnson (2012) propose a measure of liquidity that uses information from the cumulative depth in the book, not just the inside quote. Unfortunately, using order book information is not feasible for most applications, and their measure is not designed to capture dynamic order splitting.

In light of these issues, liquidity measures based on volume, rather than orders have been developed. These aggregative measures are minimally affected by dynamic order-splitting and omission of odd-lots. The most widely used is Amihud’s (2002) measure, which is based on the ratio of absolute daily returns divided by daily dollar volume. This measure has been useful at explaining the cross-section of returns, and, indeed, in time series (Amihud (2002) and Jones (2002)). Goyenko et al. (2009) show that Amihud’s measure remains useful in the post-decimalization era. With increased high frequency trading, however, such low frequency measures may lose valuable information by aggregating across heterogeneous conditions.

Acharya and Pedersen (2005) incorporate a portfolio-level aggregate of Amihud’s measure in a market model and argue that premia can be explained by covariations of stock liquidity with market-wide return and this liquidity proxy. Akbas et al. (2011) extend Acharya and Pederson (2005) by adding a measure of idiosyncratic liquidity volatility. These papers are related to research on the commonality of market-wide liquidity (Chordia et al. (2000), Sadka (2006)). Our focus is
on illiquidity as a stock characteristic, as in Ben-Rephael et al. (2014).

II Measuring Activity

To start, we define a trade sequence and explain how we use the time duration of a trade sequence to measure stock-specific trading activity. Each year, we number trades in stock $j$ sequentially, using index $n_j$. For trade $n_j$, we use $\tau_j(n_j)$, $Q_j(n_j)$ and $P_j(n_j)$ to denote respectively, (i) its time measured in seconds from the beginning of the year, (ii) its size (in shares), and (iii) its price. We construct our activity measure by calculating the time it takes to execute a sequence of consecutive trades that have an aggregate value of at least $V_{j,t}$ for stock $j$ in month $t$. Thus, a shorter time duration indicates a more active market. For the reasons detailed below, we set $V_{j,t}$ to be proportional to stock $j$’s market capitalization at the end of the previous month, $M_{j,t-1}$. The first trade sequence begins with the first trade of the year, and each subsequent trade sequence begins with the first trade following the previous sequence. Figure 3.1 illustrates a typical pattern.

Formally, we iteratively solve for the last trade of the $k^{th}$ trade sequence, $k = \{1, 2, 3, \ldots \}$, as:

$$n^k_j = \arg\min_{n^*} \left\{ \sum_{n=n_j^{k-1}+1}^{n^*} P_j^C(n) \times Q_j(n) \left| \sum_{n=n_j^{k-1}+1}^{n^*} P_j^C(n) \times Q_j(n) \geq V_{j,t} \right. \right\},$$

(1.1) 

where $n^0_j = 0$ and the value of aggregate trades is measured using the previous day’s closing price, $P_j^C(n_j)$.

Then we obtain the time duration of the $k^{th}$ trade sequence,

$$\text{dur}_{j}(k) = \tau_j(n^k_j) - \tau_j(n^{k-1}_j + 1),$$

(1.2) 

7The last quoted bid-ask midpoint is used when the closing price is not available.
Figure 1.1: An illustration of how we measure the durations of trade sequences with an aggregate value of at least $V_{j,t}$.

the corresponding return,

$$r_j(k) = \frac{P_j(n_j^k)}{P_j(n_j^{k-1} + 1)} - 1, \quad (1.3)$$

and the dollar value of the sequence,

$$DVOL_j(k) = \sum_{n=n_{j,1}^k} P_j^C(n) \times Q_j(n). \quad (1.4)$$
The per-dollar price impact for stock $j$ of the $k^{th}$ trade sequence is given by

$$DIMP_j(k) = \frac{|r_j(k)|}{DVOL_j(k)}.$$  

To compute per-dollar price impacts, we divide by $DVOL_j(k)$ rather than $V_{j,t}$ because the size of the last trade, $n^k_j$, may slightly exceed the level needed to deliver a total value of $V_{j,t}$.

We construct trade sequences that span two trading days, but exclude them from the analysis. We do this for several reasons: (i) calculating the overnight trade sequences and then excluding them ensures a random starting point for the first trade sequence of a given day, precluding any systematic bias; (ii) closing prices used to calculate dollar volumes vary from one day to the next; (iii) it circumvents issues associated with overnight price adjustments or information arrival; and (iv) it avoids combining trading activity levels from near close with those just after open which generally differ.

We sort the remaining trade sequences for stock $j$ in month $t$ into ten equally-sized groups according to their durations: groups with shorter durations capture higher trading activity. This is important: by calculating relative trading activity for a stock on a monthly basis by using monthly sorts, we control for the long-term upward trend of trading volumes, and monthly changes in $V_{j,t}$. Each year, we also sort firms into market capitalization deciles.

We can now compare the market impact of unwinding a given position at different trading activity levels. We consider two different market impact measures:

- **Price impact**: Each year, for each stock, we calculate the mean of the absolute return of

---

8We use the previous day’s closing price to calculate trade values, which makes time durations of trade sequences independent of contemporaneous price movements and allows clean comparisons of the relationship between activity and trading costs.
each trade sequence in a given duration group. We then average these annual mean absolute returns across years for all stocks in a given size decile.

- **Return volatility.** Each year, for each stock, we calculate the standard deviation of the return of trade sequences in each duration group. We then average these annual return standard deviations across years for stocks in a given size decile. Computing annual standard deviations before averaging minimizes temporal aggregation issues.

Later, we use a scaled 3-month rolling average of the price impact measure, $DIMP_j(k)$, for different subsets of trade sequences, as a priced liquidity measure, $BB\text{ }D_{j,t}$.

**Choice of $V_{j,t}$.** We focus on the market impact and the time to execute a sequence of trades that for a firm $j$ have an aggregate value $V_{j,t}$ equal to $\theta\%$ of its market capitalization at the end of the previous month, $M_{j,t-1}$. Our base formulation using $\theta = 0.04$ generates a median duration (across stocks and years) of about 20 minutes. A robustness analysis verifies that our qualitative findings are not driven by the choice of $\theta$: similar findings obtain using $V_{j,t} = 0.03M_{j,t-1}$, $V_{j,t} = 0.05M_{j,t-1}$ or $V_{j,t} = 0.025M_{j,t-1} + \$80,000$. Several considerations enter our base specification: (i) the positions are large enough to be relevant for institutional traders, and to control for dynamic order splitting and division of orders against the book; (ii) the positions are not so large that a single trade sequence spans different activity levels; (iii) the proportional specification captures the feature that institutional traders tend to establish larger positions in larger firms, and it facilitates cross-stock comparisons; (iv) the proportional specification flattens out the distribution of durations across market capitalizations (see Figure 1.2). As a result, trade sequence durations of small market capitalization stocks are not too much longer, delivering enough observations for our analysis: we

---

*By averaging stock-specific moments, we make cross-stock statistics independent of the number of trade sequences for a stock—each stock in a stock size decile contributes equally to the statistic of interest.*
drop stock-year observations for stocks without enough trade sequences in the year to avoid small sample issues, and setting $\theta = 0.04$ ensures that we do not lose many stock-year observations, even in early years. In contrast, a fixed dollar-amount based on a medium-sized stock’s market capitalization would deliver too few observations for small-cap firms, and far too many for large-cap firms.

By fixing a target dollar-volume in a month and letting execution times vary, the durations of our trade sequences control for variations in business-clock speed, providing metrics of activity that are independent of daily volume variations. This approach also facilitates temporal comparisons and temporal aggregation across roughly homogeneous entities.

A Data

Our sample period runs from January 1, 2001 to December 31, 2012. Each year, we construct a sample of U.S.-based NYSE-listed ordinary common shares (CRSP share codes 10 and 11). We drop a stock-year observation if it has less than 800 trade sequences in the year, or if the stock had a closing price\textsuperscript{11} below $1.00 on any trading day in the year. Monthly market capitalizations of firms come from the CRSP monthly stock file. We sort stocks into size-based deciles using their market capitalizations on the first trading day of a year. We use the previous day’s closing price from the CRSP daily stock file to compute dollar volumes so that duration measurements are not biased by any associated price impact\textsuperscript{12}.

We calculate CAPM betas, $\beta_{j,t}^{mkt}$, $\beta_{j,t}^{smb}$, $\beta_{j,t}^{hml}$, and $\beta_{j,t}^{umd}$, for stock $j$ in month $t$ by estimating

\textsuperscript{10}Moreover, by controlling for market capitalization, one can convert our characterizations into analyses of the costs of trading fixed dollar amounts.

\textsuperscript{11}The average of the closing bid and ask prices is used if a last trade price is not available.

\textsuperscript{12}Since we exclude trade sequences spanning more than one trading day, we do not need to adjust $r_j(k)$ for stock splits or dividend distributions.
a four-factor Fama-French model. We regress excess weekly stock returns (versus the weekly secondary market 6-month Treasury bill rate) against the excess weekly Fama-French factors, over the previous two-year period \((t-24 \text{ to } t-1)\). If a stock is not listed over the previous 24 months, we employ the longest time period available, requiring at least 26 weekly observations, where we drop the initial observation to avoid a listing effect. Firm \(j\)'s book-to-market ratio, \(BM_{j,t-1}\), is given by its book value per share, obtained from Compustat, divided by its share price, at the end of month \(t-1\).

We obtain trade prices, quantities and time stamps from the consolidated trade history in the NYSE TAQ database. We consider all stock trades on U.S.-based trading venues during regular market hours, i.e., between 9:30AM and 4:00PM (EST), together with trades corresponding to delayed reporting of market-on-close orders that are executed after 4:00PM. We exclude trade sequences with absolute returns that exceed 10%. We calculate percentage bid-ask spreads using the TAQ National Best Bid and Offer (NBBO) series on WRDS.

Table 1.1 presents annual summary sample statistics. Median stock prices largely mimic overall market performance. The median annual number of trade sequences increases threefold over the sample period, reaching a peak of 6,305 in 2009 due to the lower market capitalizations after the 2008 financial crisis. While the median market capitalizations are almost equal in 2001 and 2009 (the values of \(V_{j,t}\) are similar in those years), the median number of trade sequences is much higher in 2009, highlighting the large increase in trade volumes. Median dollar-spreads fall from 7.2¢ in 2001 to 2.4¢ in 2012.

\(^{13}\)We use daily factors from WRDS to construct weekly observations.

\(^{14}\)These trades are flagged “MOC” in TAQ data.
Table 1.1: **Year-by-year summary statistics of the final sample.** Closing share prices, book-to-market ratios, and log of market capitalizations are reported as the medians of their monthly values. $\$\text{-spreads}$ are the median time-weighted dollar bid-ask spread at the NBBO. The last column reports the median number of trade sequences (each with an aggregate value of at least $V_{j,t}$) in a year across stocks.

<table>
<thead>
<tr>
<th>Year</th>
<th># of Stocks</th>
<th>Share Price</th>
<th>Book-to-Market</th>
<th>ln(Market Cap)</th>
<th>$$\text{-spreads}$</th>
<th># of Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>1,091</td>
<td>26.50</td>
<td>0.47</td>
<td>21.14</td>
<td>0.072</td>
<td>2,288</td>
</tr>
<tr>
<td>2002</td>
<td>1,181</td>
<td>24.78</td>
<td>0.49</td>
<td>21.11</td>
<td>0.052</td>
<td>2,563</td>
</tr>
<tr>
<td>2003</td>
<td>1,170</td>
<td>23.62</td>
<td>0.53</td>
<td>20.98</td>
<td>0.035</td>
<td>2,752</td>
</tr>
<tr>
<td>2004</td>
<td>1,277</td>
<td>28.86</td>
<td>0.45</td>
<td>21.28</td>
<td>0.031</td>
<td>2,968</td>
</tr>
<tr>
<td>2005</td>
<td>1,268</td>
<td>30.95</td>
<td>0.43</td>
<td>21.41</td>
<td>0.030</td>
<td>3,317</td>
</tr>
<tr>
<td>2006</td>
<td>1,277</td>
<td>32.75</td>
<td>0.41</td>
<td>21.53</td>
<td>0.030</td>
<td>3,855</td>
</tr>
<tr>
<td>2007</td>
<td>1,255</td>
<td>33.70</td>
<td>0.41</td>
<td>21.63</td>
<td>0.029</td>
<td>4,907</td>
</tr>
<tr>
<td>2008</td>
<td>1,277</td>
<td>24.28</td>
<td>0.53</td>
<td>21.31</td>
<td>0.030</td>
<td>6,283</td>
</tr>
<tr>
<td>2009</td>
<td>1,198</td>
<td>19.62</td>
<td>0.67</td>
<td>21.11</td>
<td>0.023</td>
<td>6,305</td>
</tr>
<tr>
<td>2010</td>
<td>1,192</td>
<td>24.95</td>
<td>0.55</td>
<td>21.37</td>
<td>0.020</td>
<td>5,200</td>
</tr>
<tr>
<td>2011</td>
<td>1,183</td>
<td>28.66</td>
<td>0.50</td>
<td>21.54</td>
<td>0.022</td>
<td>5,050</td>
</tr>
<tr>
<td>2012</td>
<td>1,169</td>
<td>29.09</td>
<td>0.52</td>
<td>21.52</td>
<td>0.024</td>
<td>4,296</td>
</tr>
</tbody>
</table>

**B Overview of activity measure**

Figure [1.2] shows how durations of trade sequences vary over time and firm size. In a given year, duration and firm size have a U-shaped relationship: while larger stocks are more actively traded, we set $V_{j,t}$ proportional to firm size, which flattens the duration-size relationship. The very large positions for which we calculate durations for the largest stocks drive their longer durations; and the modest trading activity for small stocks drives their longer durations. Overall, the number of trade sequences does not vary wildly with market capitalization.

Durations tend to fall over time, reflecting the rise in trading activity. The average median duration for the second size decile of stocks falls from 33 minutes in 2001 to 26 minutes in 2012; while that for the largest stocks goes from 42 to 26 minutes. Durations are shortest in 2009 due to the lower market valuations in that year. This highlights the importance of controlling for fixed month effects by calculating a stock’s relative activity on a monthly basis.
Figure 1.2: **Trading activity by year and by firm size.** Average median duration (in minutes) of a sequence of trades with a cumulative aggregate value of at least 0.04% of a firm’s market capitalization. Median durations are calculated on a stock-by-stock basis. The average is then computed for stocks in a size decile each year. Size decile 1 contains the smallest firms; decile 10 contains the largest.

We next show how the distribution of durations is related to firm size. Figure 1.3 plots the cumulative distribution functions (CDFs) of durations over the entire sample period for representative small, medium and large market capitalization stocks: Hanger (HGR), Archer Daniels Midland (ADM), and International Business Machines (IBM). Each stock has durations that are as short as a few seconds and as long as one hundred minutes. The large firm (IBM) has far more evenly-distributed durations (its PDF has thinner tails). In contrast, ADM and HGR have many more short durations (about 80% are less than 40 minutes). HGR has fatter probability density function tails than ADM, indicating greater contributions from very short and very long durations. Figure A.3 in the chapter A shows that, quite generally, distributions for large stocks are much less skewed than those for small stocks.
Figure 1.3: Distribution of durations of trade sequences for representative stocks. Cumulative distribution functions of durations (in minutes) for International Business Machines (IBM), Archer Daniels Midland (ADM), and Hanger (HGR) from January 1, 2001 to December 31, 2012. Durations are based on the elapsed time for a sequence of trades with a cumulative aggregate value of at least 0.04% of the firm’s market capitalization.

III Analysis

Activity, signed trade imbalances, and trade size. We begin our primary analysis by documenting well-known, but poorly understood, relationships between trading activity, trade sizes and signed trade imbalances. Because each trade sequence has the same dollar volume, trade sequences with fewer trades have larger average trades. To identify buyer- and seller-initiated trades, we use the Lee and Ready (1991) algorithm.\textsuperscript{15} We define signed trade imbalance as the share of value-weighted trades out of the total dollar volume in a trade sequence that is buyer- or seller-initiated.

\textsuperscript{15}A trade is marked buyer-initiated (seller-initiated) if the execution price is above (below) the mid-point quoted price during the same second. If execution and mid-quote prices are equal, the trade is marked buyer-initiated (seller-initiated) if the execution price increased (decreased) with respect to the most recent trade.
Figure 1.4: **Average signed dollar-volume and average number of trades by trading activity level.** Average signed trade imbalance and average number of trades for trade sequences with cumulative value of 0.04% of market capitalization in 2001–2012 versus activity level for the three deciles of smallest and largest stocks. Reported average signed trade imbalance and average number of trades are, respectively, the average of firm-specific annual mean estimated signed trade imbalance and number of trades (realized over trade sequences), computed at a trading activity level. Firm-specific means at a given trading activity level are averaged over all firms in a size decile, over the entire sample period.

Figure 1.4 shows that (i) average signed trade imbalances rise with activity—more active markets feature more aggressive trading on one side of the market; (ii) average signed trade imbalances...
rise especially sharply in highly active (short duration) markets; (iii) trade sizes sharply rise with trading activity; (iv) average signed trade imbalances decline with firm size; and (v) changes in these measures become sharp in highly active markets. The changes in these measures are large: going from least to most active markets, signed trade imbalances rise by roughly 20%, and for smaller stocks, mean trade sizes double. These patterns are far too strong to be attributed to Lee-Ready classification errors; indeed, any such measurement error would likely weaken the observed relationships. In unreported results, we confirm that similar patterns hold when we condition on the time of day at which trade sequences were initiated or completed (allowing us to control for differences in liquidity trade at open and close).

By themselves, these facts seem innocuous. Classical models of speculation (e.g., Kyle (1985), Easley and O’Hara (1992), Glosten (1994)) would suggest that signed trade imbalances and trade sizes should rise with activity: speculators with substantial private fundamental information will establish substantial positions based on that information, demanding liquidity by submitting larger market orders, which will cause activity, signed trade imbalance, and trade sizes all to rise. Admati and Pfleiderer (1988) show that liquidity traders have incentives to cluster at open and close to draw more informed traders who compete away more of their information rents, leading to higher trade imbalances and smaller price impacts. However, once one controls for timing, predictions similar to those of classical models obtain.

But, those fundamentals can covary for very different reasons. If liquidity provision is time varying—if markets are more liquid (“deeper” near good inside quotes) at some times than others—then traders will respond endogenously to the extant liquidity in their order composition. If liquidity is unusually high on one side, (informed) institutional traders will adapt their order submission strategy to include more large aggressive (market) orders to take advantage of this temporary
liquidity, resulting in (i) higher activity, (ii) fewer trades comprising a trade sequence, and (iii) higher signed trade imbalances. In contrast, if depth is modest, institutional traders will use a more balanced order submission strategy that reflects the desire to consume the limited liquidity available via market orders, and the desire to reduce price impacts (and possibly receive liquidity rebates in a maker-taker fee model) by submitting limit orders. Under this situation, (i) durations are longer, (ii) more trades comprise a trade sequence, and (iii) signed trade is more balanced.\footnote{Boulatov and George (2015) show that when hidden orders are not allowed, informed traders prefer to consume liquidity.}

Activity and price impacts. Thus, both information arrival and endogenous liquidity provision choices can lead to higher signed trade imbalances and larger trades in more active markets. The question becomes: which is it? We observe that if high signed trade imbalances reflect information arrival, then price impacts should rise with activity; but if high signed trade imbalances reflect endogenous consumption of unusually high depth at good prices, then price impacts should fall with activity.

Figure 1.5 shows how the two market impact measures vary with trading activity. For small and mid-sized stocks, both mean absolute returns and the standard deviation of returns\footnote{In the Appendix, we calculate the implied calendar time volatility were trading activity to stay constant over a fixed time interval (e.g., a trading day), showing that implied calendar time volatility \textit{rises} with trading activity. The virtue of measuring activity over time in this way is that it does not mix different levels of trading activity that might arise over a non-trivial interval of time.} fall sharply as activity rises.\footnote{The unreported patterns for mid-sized stocks are similar to those of smaller stocks.} This activity-impact relationship is \textit{quite} pronounced for smaller firms. For the decile of smallest firms, the price impact of the same dollar volume \textit{rises} by almost 50\% as markets go from highest to lowest trading activity levels. In contrast, for larger stocks, market impacts do \textit{not} fall monotonically as trading activity rises. Instead, they peak at relatively high activity levels. For the decile of largest firms, market impacts are smallest when trading activity is \textit{lowest}.}
Figure 1.5: **Average absolute return and return volatility by duration group.** Market impacts of $\theta = 0.04\%$ of market capitalization in 2001–2012 versus activity level for the three deciles of smallest and largest stocks. The reported average price impact is the average of firm-specific monthly means of absolute returns (realized over trade sequences), computed per duration group. Reported average return volatility is the average of firm-specific monthly standard deviations of returns (realized over trade sequences), computed per duration group. The averages of means and standard deviations are taken over all firms in a given size decile, over the entire sample period by activity level.

Volatility falls off sharply when trading activity is low, dropping over 20% from its peak.\(^{20}\)

To shed more light, Figure 1.6 depicts the relation between inter-percentile ranges for returns...
(e.g., $r_{.99} - r_{.01}$) and trading activity for small and large firms. For small firms, all inter-percentile ranges narrow uniformly as trading activity rises—every measure of price impact falls as market activity rises. In contrast, for large stocks, (i) the interpercentile range $r_{.99} - r_{.01}$ is highest when trading activity is highest—extreme returns are more likely in active markets, and less likely in inactive ones; but (ii) the interpercentile ranges of intermediate percentile statistics (e.g., $r_{.75} - r_{.25}$) rise slightly as trading activity rises from low to intermediate levels, before falling as trading activity rises further. This combination underlies the single-peak patterns of average mean returns and return volatility in trading activity for large stocks.

![Figure 1.6: Estimated interpercentile ranges for realized returns by duration group.](image)

Figure 1.6: Estimated interpercentile ranges for realized returns by duration group. Plot of the interpercentile ranges of the distribution of realized returns for different activity levels. The percentile statistics of returns (realized over trade sequences) are first computed stock-by-stock, annually, for each duration group, and then averaged across stocks/years. Finally, the differences between selected average percentile statistics ($r_{.99} - r_{.01}$, $r_{.95} - r_{.05}$, and $r_{.75} - r_{.25}$) are computed by activity group. Durations are based on $\theta = 0.04$.

---

Footnote: We first order returns for a stock at a given trading activity level over a year, and compute the conditional return percentile statistics. We then average the returns corresponding to a selected percentile $\pi$ across stock-years at the trading activity level to estimate the return $r_{\pi}$ at percentile $\pi$. 

---

We refer to Figure A.4 in the chapter A presents the relationships between signed trade imbalances and (i) mean absolute returns, (ii) return volatility, and (iii) mean number of trades in a trade sequence,
showing that save for the largest stocks, price impacts and mean number of trades fall as signed trade imbalances rise.\footnote{In unreported results, we verify that these results hold at each level of activity.}

Cumulatively, this evidence indicates that for small and mid-sized firms, time-varying liquidity provision and consumption drive the activity/trade size/signed trade imbalance relationships, and not information arrival. For these stocks, price impacts are uniformly smaller in active markets, which is inconsistent with information arrival driving market impacts. Instead, this evidence together with the evidence that the signed trade imbalances and trade sizes rise with activity indicate that active markets are driven by traders choosing to quickly consume unusually high depth near good inner quotes. In contrast, for large stocks, time-varying liquidity provision seems to drive the market impact-activity relationship in “normal” times (captured by less extreme percentile differences); while in “unusual” times, high information arrival results in high trading activity and large price impacts.

In sum, for small and mid-sized stocks, more aggressive trading on one side of the market is associated with higher activity, but smaller price impacts. These patterns are inconsistent with the predictions of classical models of speculation—Kyle-style competitive dealership markets and Glosten-style competitive limit order markets. In the former, informed traders optimally choose market order sizes, and dealers set an asset’s price equal to its expected value given the information in net order flow. In the latter, each informed trader chooses the size of his limit order when competitive liquidity providers price limit orders to break even conditional on the information contained in the fact that their limit orders were filled. In both of these settings, a larger order or a greater net order flow, i.e., greater signed trade imbalance, has a large price impact—the opposite of what we show.
A key premise of these models is that market liquidity is invariant: traders do not have to choose between consuming and supplying liquidity when establishing or unwinding positions. In contrast, our findings highlight the impact of time-varying, imperfectly competitive liquidity provision, and the consequences for traders’ choices of how much liquidity to consume and provide (see, for example, Goettler et al. (2005), Hollifield et al. (2004), and Hollifield et al. (2006)). For most stocks, the impacts of the endogenous choice of order type in response to extant liquidity dominate the predicted relationships between signed trade imbalance, trading activity and price impact of standard models of speculation.

We emphasize the importance in our analysis of fixing dollar trading volume and then using the time it takes for that dollar volume to be executed to measure trading activity. One can then derive the relationships between trading activity and market outcomes that are relevant to an institutional investor who cares about the costs of establishing a given dollar position. If, instead, one fixed the measurement horizon—time—and then used dollar trading volume to measure market activity, one would conflate the effects of trading volume with trading activity, making inferences difficult to draw.

To illustrate the issues, in unreported results, we measure trading activity using dollar trading volume executed in fixed twenty-minute time intervals. For all stocks, mean absolute returns rise by nearly a factor of three as trading volume rises from the lowest decile of dollar volume to the highest—price impacts rise sharply with volume. It also misleads to put these price impacts on a per-dollar basis—per-dollar absolute returns cannot be used to discern the costs of establishing a target position. Concretely, in a time interval that has $100,000 of dollar volume, how does one extrapolate to obtain the cost of establishing a million dollar position? Such an exercise requires strong assumptions about how trade and price impacts covary. Indeed, with twenty minute time
intervals, as volume rises from the lowest decile to the highest, per-dollar price impacts fall by factors of *eight to thirty* depending on firm size—declines that clearly do not capture the true costs of trading at different market activity levels.

**Controlling for information arrival and endogenous liquidity consumption.** An institutional investor seeking to minimize trading costs, while establishing or unwinding a large position, would like to time trading according to activity. Such an investor, however, presumably cannot exploit information arrival to which it is not privy, and the investor may not be able to predict when liquidity provision will be “unusually high”.\(^{23}\) This observation leads us to try to separate the impacts of information arrival and the endogenous consumption and provision of liquidity, to isolate activity/trading cost relationships that are stable and predictable. That is, we identify how trading costs vary with *predicted* trading activity.

We implement a non-parametric analogue of 2SLS estimation, where the first stage is parametric and the second stage is non-parametric. We control for both endogeneity and information arrival using the number of trades filling the previous trade sequence to instrument for the duration of the current trade sequence. Recall that sequences with fewer trades correspond to higher trading activity and smaller price impacts. Thus, these sequences reflect instances of greater depth near good prices, when execution of larger marketable orders have smaller price impacts. The observations underlying our instrument are that liquidity persists over time—present and past states of the order book are positively correlated—but the endogenous composition of orders hinges only on the *current* state of the order book. That is, the effects of past order book states on current price impacts are *only* transmitted via the current state of the order book. Importantly, the nontrivial dollar volume in a trade sequence minimizes the temporal dependence associated with endogenous

\(^{23}\text{Given the extant liquidity, the complicated algorithms of brokers intelligently route orders of various types into various trading venues to extract as much liquidity as possible.}\)
order composition. Intuitively, the endogenous responses quickly offset imbalances in liquidity—
unusually high liquidity on one side of the market is consumed, and limit orders fill in the book
when liquidity is unusually low—returning the limit order book to a more balanced state.

We next present evidence that cumulatively validate our instrument. For each stock, we first
regress the natural logarithm of duration on a quadratic polynomial of the number of trades in
the previous trade sequence, *annually* (fitting regressions annually accounts for temporal changes
in order sizes). The fitted values of these regressions serve as *instrumented* levels of activity. We
then sort the monthly samples by instrumented/predicted activity and investigate relationships just
as we did with raw activity.

To show the instrument’s merits, we first estimate the linear correlation between the duration
of the current trade sequence and the number of trades filling the previous trade sequence for each
stock-year. The empirical distribution of this statistic reveals the significance of the correlation
parameter for a typical stock. The top row of Figure 1.7 shows this empirical distribution for
different-sized firms. The correlation coefficient is positive and large for a typical stock in each size
category. Its average rises from 0.2 for small stocks to 0.25 for large stocks—past liquidity predicts
current trading activity.

The bottom row of Figure 1.7 presents the empirical distributions of the *F*-statistics for the
significance of the first stage log-polynomial regressions at the stock-year level. Over 96% of the
*F*-statistics exceed 10, indicating that the past number of trades is a valid statistical instrument for
activity (Staiger and Stock (1997) suggest first stage regression *F*-statistics of at least 10 to assure
maximum estimation bias of less than 10%).

---

24 Linear or higher order polynomial specifications lead to similar results; but the quadratic specification has a
stronger first stage fit.
Figure 1.7: Empirical distribution of linear correlations and log-polynomial $F$-statistics by size decile. In the top graphs, for each stock-year observation, we find the correlation between current duration and lagged number of trades. The kernel density of these correlation coefficients is estimated by size groups. In the bottom graphs, for each stock-year observation, the current log duration is regressed on a second-order polynomial of the lagged number of trades and the corresponding $F$-statistic is obtained. We then estimate, by size groups, the kernel density for the empirical distributions of the $F$-statistics. The vertical dashed lines indicate $F$-statistics of 10. Values exceeding 500 are omitted for visual clarity.
One might worry that this persistence could be mirrored by equal persistence in the endogenous provision and consumption of liquidity. To establish that this is not so, we first calculate the analogous first stage regressions with signed trade imbalance—which measures endogenous consumption and provision of liquidity—rather than trading activity, as the endogenous variable. In these unreported results, only about 67% of the associated $F$-statistics exceed 10, indicating a weaker relationship between past number of trades and current signed trade imbalance than that with current activity. More concretely, Figure A.1 and Table A.1 in the Appendix contrast results of parametric and non-parametric estimations of the relationships between (a) instrumented activity and signed trade imbalances with those for (b) trading activity and signed trade imbalances. They show that while signed trade imbalance still rises with instrumented trading activity, the relationship is far weaker (by factors of at least two) than that between signed trade imbalance and raw trading activity.

In sharp contrast, the relationship between trading costs and activity dramatically changes when we instrument using predicted trading activity, delivering strong and consistent impacts of predicted activity on trading costs. Figure 1.8 details how trading costs vary with predictable components of trading activity. For all stock size groups, trading costs consistently fall as predicted trading activity rises. Trading costs fall on average by 5–8 basis points depending on stock size as predicted trading activity rises from its lowest level to its highest.

- For small stocks, instrumenting flattens the price impact–trading activity relationship by roughly 75%, highlighting the pronounced market-chasing of liquidity providing orders in less active markets of smaller stocks that lead to large cumulative price impacts;
- For larger stocks, instrumentation dramatically unwinds the non-monotone relationship be-
Figure 1.8: **Average absolute return by predicted activity group.** Average absolute return of 0.04% of market-cap in 2001–2012 versus predicted trading activity for the deciles of smaller and larger stocks. The reported average absolute return is the average of firm-specific annual mean absolute return (realized over trade sequences), computed per predicted activity group. The average of means is taken across all firms in a given size decile, over the entire sample period by predicted activity level. Average absolute return versus activity is also displayed to highlight differences. Stock-year observations with first stage regression $F$-statistics below 10 are dropped.
tween activity and price impacts. Instead, just like with smaller stocks, the average price impact falls as predicted activity rises. Indeed, the relationship not only becomes monotone, it becomes steeper: the difference between the peak and trough of the trading cost – trading activity relationship rises.

The reductions in price impacts as predicted trading activity rises are meaningful from the perspective of an institutional investor, especially since we use a simple instrument to predict trading activity, and investors can employ more sophisticated prediction methods to achieve greater savings on trading costs.

This uniform relationship indicates that the instrumentation removes most of the effects of information arrival and the endogenous contemporaneous consumption and provision of liquidity, which vary sharply with stock size. Indeed, by contrasting activity/price impact relationships with and without instrumentation, one can identify the key roles that the contemporaneous choice of order type and information arrival play to shape the relationships between market impacts and trading activity.

**Parametric Analysis.** We now buttress these non-parametric findings with two 2SLS regression estimation approaches, one that pools observations within a stock decile, and one that estimates the model at the stock-year level, and uses the distribution of the roughly 14,000 coefficient estimates to make inferences.

To use the information from each stock to estimate the relationships between trading costs and predicted trading activity, we first standardize measurements across stocks by defining stock-specific duration rank statistics. *Each month*, we sort a stock’s trade sequences by predicted activity from *longest* to *shortest* durations. For example, if a stock has 200 trade sequences in a month, the
longest duration becomes percentile 0.5, and the shortest duration becomes percentile 100. This permits clean cross-stock comparisons of the impact of predicted trading activity on trading costs regardless of the number of trade sequences for a firm.

**Pooled approach.** We pool observations across stocks in a size decile over time, inversely weighting observations for a stock-year by the number of trade sequences it has in that year. This weighting strategy offsets the effect on estimates of variation in the number of trade sequences across stock-years (e.g., rising over time, balancing contributions of more and less actively-traded stocks)—see Solon et al. (2013). For each decile, we then regress absolute returns in basis points on instrumented activity percentile.

**Disaggregated approach.** This innovative estimation approach regresses at the stock-year level the absolute return (in basis points) realized over a trade sequence on the corresponding instrumented trading activity percentile. To make inferences, we use the empirical distributions of the slope coefficients from these regressions. We treat these point estimates as a random sample. Relying on the central limit theorem, we use the sample mean and standard deviation of slope point estimates to construct confidence intervals for the average slope for a given size decile. Constructing confidence intervals in this way is conservative because it uses estimates from a pool of different stocks (over 110 stocks per size decile) over a 12 year period in an unbalanced panel. The large cross-sectional heterogeneity and the temporal changes in sample composition and market structure clearly introduce greater variance in the empirical distribution of estimates than the limited stock-specific time series correlation structure reduces them, leading to conservative estimates of standard errors.

\[25\]

\[25\]In contrast, the Fama-MacBeth (1973) method treats estimates from different periods as a random sample of i.i.d. entities with one entity per period. In practice, a firm’s characteristics strongly persist over time, sharply reducing the variability of estimates relative to the i.i.d. premise.
This approach has other key virtues:

- It allows for stock-year fixed effects and random coefficient estimation, as Swamy (1970) recommends, preserving an arbitrary stock-year-specific linear structure between price impacts and predicted trading activity. Mean durations of trade sequences vary extensively in the cross-section and over time. By estimating at the stock-year level, our approach finely controls for this variation.

- Statistical inference is based on the empirical distributions of stock-year-specific point estimates. As a result, we do not have to make any assumptions regarding the correlation structure of error terms over time or across stocks. This latter feature circumvents the difficulties associated with identifying a parametric structure to deal with correlated errors in the cross-section of stocks. Using the empirical distribution of estimates renders it unnecessary to specify an error term correlation structure that would deliver robust standard errors.

The point estimate of the slope in the pooled approach and the mean stock-specific slope in the disaggregated approach measure by how much (in basis points) trading costs change as instrumented activity rises from the 1st percentile to the 99th. Table 1.2 shows that both estimation methods deliver qualitatively identical sensitivities of trading costs to activity at a given size decile. The estimates indicate that trading costs fall as predicted trading activity rises—the trading costs of establishing or unwinding a position equal to 0.04% of a stock’s market capitalization fall by 4–8 basis points going from the lowest predicted activity level to the highest. These estimates also reveal that the predicted cost saving from trading in active markets is greater for smaller stocks—the sensitivity of trading costs to trading activity roughly doubles as one goes from the largest decile of stocks to the smallest. These findings are consistent with the non-parametric findings illustrated
in Figure 1.8

Table 1.2 also breaks down the sample period into 2001–2006 and 2007–2012 periods. This pre-
and post-RegNMS decomposition reveals that the size heterogeneity associated with the sensitivity
of price impacts to changes in predicted trading activity is driven by the pre-RegNMS period. In
2007–2012, for all stock deciles save the smallest, trading costs fall by 5 basis points as predicted
trading activity rises from the first percentile to the 99th. This uniformity in the sensitivity of
price impacts to predicted trading activity suggests that changes in the market design, the replace-
ment of market makers by HFTs and automated liquidity providers, and improvements in trading
technologies have made liquidity provision considerations more homogeneous across stocks in an
important way.

Table 1.2 only provides information about how trading costs change as predicted trading activity
rises; it reveals nothing about how the mean levels of those costs have evolved. We next investigate
how the revolutionary changes in market structure have affected levels of trading costs. In particular,
have the costs of trading different stocks converged or diverged?

To find out, we explore the temporal variations of the price impacts associated with execution
of 0.04% of market-capitalization, controlling for stock size. We compute the stock-specific mean
price impacts (absolute returns), quarterly, and then average those means by stock size quintile
every quarter. We find that in the post-RegNMS period, average price impacts tend to fall for large
stocks, but to rise for small ones.

The top row of Figure 1.9 presents the temporal variations of average price impacts and the
 corresponding 99% confidence intervals for small and large firms, where the horizontal lines indicate
the sample median of price impacts over the whole sample period. For both small and large stocks,
Table 1.2: Trading costs and instrumented trading activity. For each stock-year observation, the absolute return realized over a trade sequence is regressed on the corresponding instrumented trading activity percentile. The empirical distributions of these point estimates within each size decile are used to draw statistical inference. The sampling distribution is assumed to follow a student t distribution (panel A). Observations are pooled across stocks within a size decile. Absolute return is regressed on the corresponding instrumented trading activity percentile; partial regression outputs are reported (panel B). The point estimates give the amount by which trading costs change as instrumented trading activity rises from 1st percentile to the 99th. Results are reported for 2001-2012, and subperiods 2001–2006 and 2007–2012.

### Panel A: Stock level estimates

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean slope</td>
<td>Std. Err.</td>
<td>95% CI</td>
</tr>
<tr>
<td>1</td>
<td>-7.91</td>
<td>0.37</td>
<td>(-8.64, -7.18)</td>
</tr>
<tr>
<td>2</td>
<td>-8.23</td>
<td>0.30</td>
<td>(-8.82, -7.65)</td>
</tr>
<tr>
<td>3</td>
<td>-7.29</td>
<td>0.22</td>
<td>(-7.73, -6.86)</td>
</tr>
<tr>
<td>4</td>
<td>-6.99</td>
<td>0.19</td>
<td>(-7.37, -6.62)</td>
</tr>
<tr>
<td>5</td>
<td>-6.29</td>
<td>0.17</td>
<td>(-6.62, -5.95)</td>
</tr>
<tr>
<td>6</td>
<td>-5.60</td>
<td>0.15</td>
<td>(-5.90, -5.30)</td>
</tr>
<tr>
<td>7</td>
<td>-5.37</td>
<td>0.14</td>
<td>(-5.65, -5.09)</td>
</tr>
<tr>
<td>8</td>
<td>-5.19</td>
<td>0.14</td>
<td>(-5.47, -4.91)</td>
</tr>
<tr>
<td>9</td>
<td>-4.95</td>
<td>0.14</td>
<td>(-5.22, -4.68)</td>
</tr>
<tr>
<td>10</td>
<td>-3.98</td>
<td>0.15</td>
<td>(-4.28, -3.68)</td>
</tr>
</tbody>
</table>

### Panel B: Size group level estimates

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Std. Err.</td>
<td>95% CI</td>
</tr>
<tr>
<td>1</td>
<td>-7.60</td>
<td>0.14</td>
<td>(-7.87, -7.33)</td>
</tr>
<tr>
<td>2</td>
<td>-7.92</td>
<td>0.09</td>
<td>(-8.10, -7.74)</td>
</tr>
<tr>
<td>3</td>
<td>-7.13</td>
<td>0.07</td>
<td>(-7.26, -6.99)</td>
</tr>
<tr>
<td>4</td>
<td>-6.94</td>
<td>0.06</td>
<td>(-7.05, -6.82)</td>
</tr>
<tr>
<td>5</td>
<td>-6.23</td>
<td>0.05</td>
<td>(-6.33, -6.14)</td>
</tr>
<tr>
<td>6</td>
<td>-5.58</td>
<td>0.05</td>
<td>(-5.67, -5.48)</td>
</tr>
<tr>
<td>7</td>
<td>-5.41</td>
<td>0.04</td>
<td>(-5.50, -5.32)</td>
</tr>
<tr>
<td>8</td>
<td>-5.21</td>
<td>0.04</td>
<td>(-5.29, -5.12)</td>
</tr>
<tr>
<td>9</td>
<td>-4.96</td>
<td>0.04</td>
<td>(-5.05, -4.88)</td>
</tr>
<tr>
<td>10</td>
<td>-4.15</td>
<td>0.05</td>
<td>(-4.26, -4.04)</td>
</tr>
</tbody>
</table>
average price impacts drop below the median in 2004–2007. Surprisingly, different trends present themselves post crisis (2010–2012): on average, price impacts of large stocks drop below the median, but those for small stocks do not. More generally, in 2010-2012, mean trading costs are lower for firms above the 40th percentile of firm size, but not for firms below the 40th percentile. To illuminate this divergence, we normalize the price impacts for small and large stocks by dividing
by their respective sample median price impacts. The bottom panel of Figure 1.9 plots how these median-adjusted price impacts and their ratio vary over time. These plots make clear that even as the impacts of trading activity levels have become more homogeneous in recent years, the mean levels of trading costs have diverged—on average, beginning in 2007, relative to the costs for larger stocks, the costs of trading smaller stocks have risen by almost forty percent.

More striking evidence obtains when we identify stock illiquidity more precisely. To do this, we average per-dollar absolute returns of fixed-dollar positions for a stock on a monthly basis (see the detailed description in Section IV of the construction of our BBD illiquidity measure), to investigate the evolution of trading costs. With this measure, on average, the per-dollar price impact of executing .04% of market-cap of a more illiquid stock is greater. We order all stocks in a month by this illiquidity measure from the most liquid to the least liquid. Figure 1.10 plots the temporal evolutions of the 20th and the 80th percentiles of this illiquidity measure and their ratio. The figure shows that the costs of trading less liquid (80th percentile) stocks have almost doubled between 2005 and 2012 relative to those for more liquid (20th percentile) stocks, with the bulk of the increase occurring post full implementation of RegNMS in 2007. This temporal evolution indicates that illiquid stocks became increasingly illiquid both in levels and in relative terms post crisis: the RegNMS reform and the advent of algorithmic trading seem to have reduced the trading costs of stocks that were already more liquid, at the expense of less liquid stocks.

A plausible explanation is that while specialists were effectively crowded out of markets for all stocks in recent years, larger stocks have attracted disproportionately more liquidity provision from automated liquidity providers, reducing trading costs. Indeed, Yao and Ye (2015) show that, on average, HFTs have twice the market share of depth and volume in large cap stocks (top third of market cap) than small cap stocks (bottom third). Moreover, today’s fragmented markets make it
more difficult to locate liquidity for less liquid stocks without drawing attention of HFTs, who may front-run and steal liquidity.

IV Priced liquidity and changing illiquidity premia

The asymmetric evolution of trading costs for liquid and illiquid stocks raises the question: what has happened to the premia that investors demand to hold less liquid stocks? To investigate how liquidity premia have changed, we now estimate illiquidity premia in an asset pricing model and examine their evolution over time.

We begin our asset pricing analysis by showing that our measure of illiquidity successfully explains cross-sectional variations in expected returns. A priori, it is not obvious which activity levels matter “most” to traders—price impacts are higher in less active markets, but institutional traders may be able to time trades strategically. As such, we construct measures that control for the level of trading activity monthly on a stock-by-stock basis. The trade-time illiquidity measure
based on trading activity level $\zeta = \text{low (bottom 30%)}, \text{medium (middle 40%)}, \text{high (top 30%)}$ or all, for stock $j$ in month $t$, $BBD_{j,t}^\zeta$, is given by the average per-dollar absolute return (scaled by $10^6$) of stock $j$’s trade sequences at trading activity level $\zeta$ during the 3-month period ending on the last day of month $t$. Thus, our base measure weights equally the price impacts across all trade sequences for a stock, while the alternative measures isolate particular trading activity levels.

We estimate an augmented four-factor CAPM that also includes book-to-market, size, previous month’s return, and the $BBD$ liquidity measure as stock characteristics:

$$r_{j,t} - r_{t}^f = \gamma_0 + \gamma_1 \beta_{j,t}^{mkt} (r_{t}^{mkt} - r_{t}^f) + \gamma_2 \beta_{j,t}^{hml} HML_t + \gamma_3 \beta_{j,t}^{smb} SMB_t + \gamma_4 \beta_{j,t}^{umd} UMD_t + \gamma_5 BM_{j,t-1} + \gamma_6 \ln(M_{j,t-1}) + \gamma_7 r_{j,t-1} + \gamma_8 BBD_{j,t-1}^\zeta$$

$$+ \sum_{m=4}^{120} \gamma_m^D Dum_m + \gamma_{10} \Delta(YIR)_{j,t} + \epsilon_{j,t}, \quad (1.6)$$

where $\Delta(YIR)_{j,t}$ denotes the change (from month $t-1$) in the difference between the industry and market return $r_{t}^{mkt}$, where stock $j$ is excluded when computing the return of the industry to which stock $j$ belongs. $Dum_m = 1$ if $j$ is observed in month $m$ and is 0 otherwise. These monthly dummies capture any common variation in excess returns resulting, for example, from changes in the trading environment.

We use GLS estimation with stock fixed effects in the error term. This strategy helps identify excess return differences caused by unobserved temporally-constant firm attributes. We allow for firm clusters and estimate robust standard errors to avoid downward biases in standard error estimates that might result from firm-specific serial correlations in the error term (see Petersen’s (2009) discussion of the benefits of this estimation strategy). We control for time-specific error

---

26 We also considered variants based on predicted trading activity level or that focused on the 40% of trade sequences with medium levels of signed trade imbalance. Qualitatively similar estimates obtain.

27 The sample period begins in April 2001, due to our three-month rolling window construction.
term autocorrelations using $\Delta(YIR)$: $\Delta(YIR)$ captures common industry shocks that could lead to contemporaneously correlated errors for stocks in the same industry. In the Chapter A, we establish robustness of our findings to alternative ways of dealing with the error term structure (e.g., imposing an AR(1) error structure). As in Amihud (2002), each month, we drop the top 1% of observations of the illiquidity measure(s) to avoid spurious results driven by the high variability of extreme observations.\textsuperscript{28}

Table 1.3 presents estimates of (1.6) using each $BBD$ measure. Regardless of the $BBD$ measure selected, its coefficient is positive and significant after controlling for standard sources of systematic risk and stock characteristics. Moreover, all coefficients on other factors and characteristics have the expected signs, are statistically significant, and are virtually unaffected by the choice of $BBD$.

Recall that $BBD$ reflects the average price impact per million dollars, and that average $BBD$ varies with the level of trading activity on which it is based. To obtain an illiquidity premium, we substitute in the median level of $BBD$. This substitution yields the estimated monthly return premium that investors demand for facing the cost of possibly unwinding a million dollar position of a typical stock relative to a hypothetical “perfectly liquid” stock. The estimated monthly premium for the $BBD$ measures range from 6.0 to 7.6bp, and the average is 6.6bp.\textsuperscript{29} Given these qualitatively similar results, our subsequent presentation of results focuses on the version of $BBD$ that is based on all trade sequences.

\textsuperscript{28}Amihud (2002) also trims the bottom 1%; this additional trimming has no qualitative effects on estimates reflecting the positively-skewed illiquidity measure distribution. Similar estimates obtain with alternative trimming thresholds (e.g., top 2% or top 5%).

\textsuperscript{29}In unreported results, we find that the log-likelihood is maximized by illiquidity measures that place more weight on active markets. This finding suggests that traders may have the ability to time some trading to occur in more active markets. The log-likelihood, however, is quite flat, indicating that the measures contain roughly the same amount of information.
Table 1.3: Augmented CAPM including trade-time liquidity measure. Estimation results for (1.6) using the trimmed sample, where the error term allows for stock fixed effects. Standard errors are clustered at the stock level. Month dummies capture any common variation in excess returns. $\beta(r_{it}^{mk} - r_{it}^{f})$ is the product of a firm’s market beta and excess monthly return based on monthly holding period returns from CRSP and 1-month treasury bills. $BM_{t-1}$ is the previous period’s Book-to-Market ratio. $Ln(M_{t-1})$ is the natural logarithm of the firm’s market capitalization at the end of previous month. $BBD_{t-1}$ is the previous month’s trade-time illiquidity measure based on $\theta = 0.04$. Symbols *, **, and *** denote significance at 10%, 5%, and 1% levels, respectively. Robust standard errors are reported in parentheses. Premium is the product of the $BBD$ coefficient and the median value of $BBD$.

<table>
<thead>
<tr>
<th>Co-variate</th>
<th>All activity levels</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta^{mk}(r_{it}^{mk} - r_{it}^{f})$</td>
<td>0.437***</td>
<td>0.437***</td>
<td>0.436***</td>
<td>0.437***</td>
</tr>
<tr>
<td>(0.026)</td>
<td>(0.026)</td>
<td>(0.026)</td>
<td>(0.026)</td>
<td></td>
</tr>
<tr>
<td>$\beta^{HML-HML}$</td>
<td>0.235***</td>
<td>0.237***</td>
<td>0.235***</td>
<td>0.237***</td>
</tr>
<tr>
<td>(0.026)</td>
<td>(0.026)</td>
<td>(0.026)</td>
<td>(0.026)</td>
<td></td>
</tr>
<tr>
<td>$\beta^{SMB-SMB}$</td>
<td>0.324***</td>
<td>0.322***</td>
<td>0.323***</td>
<td>0.323***</td>
</tr>
<tr>
<td>(0.024)</td>
<td>(0.024)</td>
<td>(0.024)</td>
<td>(0.024)</td>
<td></td>
</tr>
<tr>
<td>$\beta^{UMPUMD}$</td>
<td>0.396***</td>
<td>0.396***</td>
<td>0.397***</td>
<td>0.400***</td>
</tr>
<tr>
<td>(0.026)</td>
<td>(0.026)</td>
<td>(0.026)</td>
<td>(0.026)</td>
<td></td>
</tr>
<tr>
<td>$BM_{t-1}$</td>
<td>0.005***</td>
<td>0.005***</td>
<td>0.005***</td>
<td>0.004***</td>
</tr>
<tr>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>$Ln(M_{t-1})$</td>
<td>-0.022***</td>
<td>-0.022***</td>
<td>-0.022***</td>
<td>-0.023***</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>$r_{t-1}$</td>
<td>0.025***</td>
<td>0.025***</td>
<td>0.025***</td>
<td>0.027***</td>
</tr>
<tr>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td>$BBD_{t-1}$</td>
<td>0.182***</td>
<td>0.145***</td>
<td>0.196***</td>
<td>0.192***</td>
</tr>
<tr>
<td>(0.034)</td>
<td>(0.028)</td>
<td>(0.033)</td>
<td>(0.039)</td>
<td></td>
</tr>
</tbody>
</table>

Log-likelihood 135863 135831 136070 136831
Within-$R^2$ 0.31 0.31 0.31 0.31
Observarions 158903 158905 159511 159515

Premium (bps) 6.9 6.2 7.6 6.0

Contrasts with Ben-Rephael et al. (2015)’s finding of no significant liquidity premia post decimalization. Differences in the construction of our illiquidity measures likely explain why we uncover illiquidity premia, but they do not. Our monthly illiquidity measure is calculated using average illiquidity premia, but they do not. We show shortly that similar findings obtain for standard illiquidity measures constructed at monthly frequencies, including Amihud’s (2002) measure, which serves as illiquidity measure in Ben-Rephael et al. (2015). This conjecture does not undermine their central conclusion that illiquidity premia declined from levels in much earlier periods, and, in particular, are lower post-decimalization than pre-decimalization.
per-dollar price impacts over the previous three months. As a result, changes in illiquidity are incorporated on a month-by-month basis as they occur. In contrast, because Ben-Rephael et al. (2015) must use daily observations for their historical analysis, they measure key independent variables, including stock illiquidity, at annual frequencies. Thus, they use illiquidity measures from the preceding year to explain the monthly cross-section of returns: the same illiquidity measurement for a stock applies to all 12 months of a given year. This limits the variation of these independent variables, and it amplifies measurement error that biases their coefficients toward zero. In addition, by assuming no adjustment in investors’ assessments of stock-specific illiquidity over a year, they may mix fixed year effects with those of cross-stock illiquidity.

**Comparisons with other illiquidity measures.** We next compare the asset pricing performance of *BBD* with those of commonly-used illiquidity measures—Amihud (2002), dollar spreads, and percentage spreads. The standard Amihud illiquidity measure aggregates over a trading day and hence may not exploit informative intraday variations in trading activity. This property leads us also to consider a high-frequency version of Amihud based on hourly observations of price changes and dollar volumes, to ensure that performance differences are not due to excessive temporal aggregation in the daily Amihud measure.

To calculate the standard (low-frequency) Amihud measure (*AML*), we first compute the daily per-dollar price impact for each stock *j* on each day *q*, 

\[ DIMP_j(q) = \frac{|r_j(q)|}{DVOL_j(q)} \]

where \( r_j(q) \) is the return on stock *j* on day *q*, and \( DVOL_j(q) \) is the daily dollar trading volume. The low-frequency Amihud measure for stock *j* in month *t* is given by the average of \( DIMP_j(q) \) over trading days *q* in the 3-month period ending on the last day of month *t*, scaled up by \( 10^6 \).

To calculate the high-frequency analogue, *AMH*\(_{j,t}\), we first calculate the hourly per-dollar price
Figure 1.11: Monthly median values of different illiquidity measures over time. The BBD measure (based on $\theta = 0.04$) is calculated using observations from all trade sequences. Also shown is the high frequency Amihud measure ($AMH$), the low frequency Amihud measure ($AML$) and percentage bid-ask spreads ($PSP$).

Impact for a stock as the absolute hourly realized return divided by the corresponding executed hourly dollar trading volume. Daily averages of these hourly price impacts produce daily per-dollar price impacts. The high-frequency Amihud measure for stock $j$ in month $t$ is given by the average of the daily per-dollar price impacts over trading days $q$ in the 3-month period ending on the last day of month $t$, scaled by $10^6$.

Average dollar bid-ask spreads ($DSP$) and percentage bid-ask spreads ($PSP$) are calculated using a time-weighted average of spreads based on the NBBO at a one-second frequency, on a daily basis. We calculate a rolling 3-month average of spreads to be consistent with the other measures. Hence, for stock $j$, the spread used in month $t$ is the simple average of $DSP$ or $PSP$ from months
$t-1$, $t-2$, and $t-3$, denoted $DSP_{j,t-1}$ and $PSP_{j,t-1}$.

Figure 1.11 plots the monthly median value of each illiquidity measure. The measures are highly correlated and display similar temporal variation. The simple average sample correlations between $BBD$ and $PSP$, $AML$, and $AMH$ are 0.77, 0.80, and 0.66, respectively. Notably, $BBD$ varies more than the other measures.

Table 1.4 summarizes estimates of (1.6) using these alternative illiquidity measures in lieu of $BBD$ (using the same error term structure and trimming criteria). Dollar bid-ask spreads is the only alternative measure that is not significant (even at $\alpha = 0.1$). Consequently, our subsequent analysis compares $BBD$ with the priced measures: $AML$, $AMH$, and $PSP$.

The illiquidity measures are highly correlated. A pricing regression that includes $BBD$ and another illiquidity measure will have multicollinearity issues. To address this issue, we decompose each measure $F \in \{AML, AMH, PSP\}$ into two linearly-orthogonal components with respect to $BBD$. We first obtain the residuals $z_{tj}^F$ and $\tilde{z}_{tj}^F$ from the regressions,

$$BBD_{j,t} = \hat{\alpha}_1 + \hat{\alpha}_2 F_{j,t} + z_{tj}^F,$$

(1.7)

$$F_{j,t} = \hat{\alpha}'_1 + \hat{\alpha}'_2 BBD_{j,t} + \tilde{z}_{tj}^F.$$

(1.8)

We then estimate (1.6) with $z_{tj}^F$ or $\tilde{z}_{tj}^F$ as the illiquidity measure. This approach isolates the information content in one measure that is not in the other measure since, by construction,

$\text{Cov}(z_{tj}^F, F_{jt}) = \text{Cov}(\tilde{z}_{tj}^F, BBD_{jt}) = 0$. Table 1.5 reveals that $BBD$ captures more information

\footnote{In unreported results, we find that if both $BBD$ and another illiquidity measure are included in the same regression, the coefficient on $BBD$ is positive and statistically significant, while that on the other illiquidity measure is negative and insignificant.}
about stock illiquidity that do the standard measures: After filtering out the commonality between $BBD$ and each alternative measure, what remains (i.e., the $BBD$ residual, $z^{p}_{t-1}$) is positively and significantly correlated with expected returns; but the converse is not true. In Chapter A, we provide evidence that $BBD$’s better performance is more pronounced in the latter half of the sample, indicating that our measure better captures liquidity considerations of investors in today’s markets.
Table 1.5: **Augmented CAPM estimates for orthogonal decompositions.** The table reports panel estimation results for (1.6) using the full sample. Standard errors are clustered at the stock level. Stock fixed effects capture any fixed heteroskedasticity. Month dummies capture any common variation caused by market condition changes. We separately present results that contrast BBD against each alternative measure AML, AMH, and PSP. Durations are generated based on $\theta = 0.04$. Observations in the upper 1% tail of either BBD or $F \in \{AMI, AMH, PSP\}$ in a month are excluded. Robust standard errors are reported in parentheses. Symbols *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

<table>
<thead>
<tr>
<th>Co-variates</th>
<th>BBD against the alternative illiquidity measure $F$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F = AML$</td>
</tr>
<tr>
<td>$\beta^mkt (r_{kt}^0 k_t - r_t^0)$</td>
<td>0.437*** (0.026)</td>
</tr>
<tr>
<td>$\beta^{hml} HML_t$</td>
<td>0.236*** (0.026)</td>
</tr>
<tr>
<td>$\beta^{smb} SMB_t$</td>
<td>0.329*** (0.025)</td>
</tr>
<tr>
<td>$\beta^{umd} UMD_t$</td>
<td>0.398*** (0.025)</td>
</tr>
<tr>
<td>$BM_{t-1}$</td>
<td>0.005*** (0.026)</td>
</tr>
<tr>
<td>$Ln(M_{t-1})$</td>
<td>-0.024*** (0.001)</td>
</tr>
<tr>
<td>$r_{t-1}$</td>
<td>0.026*** (0.005)</td>
</tr>
<tr>
<td>$z_{t-1}^F$</td>
<td>0.287*** (0.061)</td>
</tr>
<tr>
<td>$\tilde{z}_{t-1}$</td>
<td>-0.273*** (0.089)</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>135765</td>
</tr>
<tr>
<td>Within-$R^2$</td>
<td>0.31</td>
</tr>
<tr>
<td>Observations</td>
<td>158478</td>
</tr>
</tbody>
</table>

We next investigate the asset pricing implications of the asymmetric evolution of trading costs across different stocks in recent years—trading costs of liquid and illiquid stocks have diverged. That is, we seek to answer the question: how have the asymmetric changes in stock liquidity been reflected in the cross-section of asset prices? To answer this question, we divide the sample into three periods, April 2001–December 2007 (pre crisis), January 2008–December 2009 (crisis) and January 2010–December 2012 (post crisis), estimate model 1.6 in each period, and then compare
Table 1.6: Subsample estimation of augmented CAPM with different liquidity measures. Panel estimation results for three subsamples: April 2001–Dec 2007, Jan 2008–Dec 2009, and Jan 2010–Dec 2012. Standard errors are clustered at the stock level. Stock fixed effects are introduced to capture any fixed heteroskedasticity. Month dummies are included to capture any common variation caused by market condition changes. The first column for a given sub-sample present results for our BBD measure based on $\theta = 0.04$ computed over all trade sequences. The other three columns of a panel present estimation results for the standard liquidity measures, $AML$, $AMH$, and $PSP$. Observations in the upper 1% tail of the relevant liquidity measure are excluded. Robust standard errors are reported in parentheses. The symbols *, **, and *** denote significance at 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta^{alt}(r_{it}^m-r_{it})$</td>
<td>0.308*** (0.027) 0.309*** (0.027) 0.313*** (0.027) 0.314*** (0.027)</td>
<td>0.520*** (0.045) 0.523*** (0.045) 0.521*** (0.045) 0.529*** (0.044)</td>
<td>0.488*** (0.048) 0.487*** (0.047) 0.485*** (0.048) 0.487*** (0.048)</td>
</tr>
<tr>
<td>$\beta_{bolo}^{HML_t}$</td>
<td>0.163*** (0.024) 0.164*** (0.024) 0.162*** (0.024) 0.164*** (0.024)</td>
<td>0.287*** (0.057) 0.299*** (0.056) 0.297*** (0.056) 0.303*** (0.058)</td>
<td>0.334*** (0.067) 0.344*** (0.067) 0.344*** (0.067) 0.339*** (0.068)</td>
</tr>
<tr>
<td>$\beta_{bolo}^{SMB_t}$</td>
<td>0.250*** (0.023) 0.255*** (0.024) 0.254*** (0.023) 0.288*** (0.023)</td>
<td>0.405*** (0.057) 0.398*** (0.056) 0.395*** (0.056) 0.395*** (0.058)</td>
<td>0.500*** (0.067) 0.498*** (0.067) 0.499*** (0.068) 0.505*** (0.068)</td>
</tr>
<tr>
<td>$\beta_{bolo}^{HMD_t}$</td>
<td>0.396*** (0.029) 0.373*** (0.028) 0.372*** (0.028) 0.370*** (0.028)</td>
<td>0.467*** (0.032) 0.468*** (0.032) 0.468*** (0.031) 0.471*** (0.031)</td>
<td>0.503*** (0.035) 0.507*** (0.035) 0.507*** (0.036) 0.507*** (0.036)</td>
</tr>
<tr>
<td>$BM_{t-1}$</td>
<td>0.007* (0.004) 0.008* (0.004) 0.007* (0.004) 0.006* (0.004)</td>
<td>0.002 (0.003) 0.002 (0.003) 0.001 (0.003) 0.002 (0.003)</td>
<td>0.006 (0.005) 0.005 (0.005) 0.006 (0.005) 0.003 (0.005)</td>
</tr>
<tr>
<td>$L_i(M_{t-1})$</td>
<td>-0.028*** (0.002) -0.030*** (0.002) -0.031*** (0.002) -0.025*** (0.002)</td>
<td>-0.153*** (0.007) -0.157*** (0.007) -0.158*** (0.007) -0.155*** (0.007)</td>
<td>-0.069*** (0.005) -0.075*** (0.005) -0.075*** (0.005) -0.073*** (0.005)</td>
</tr>
<tr>
<td>$r_{t-1}$</td>
<td>0.019*** (0.006) 0.030*** (0.006) 0.020*** (0.006) 0.012*** (0.006)</td>
<td>0.059*** (0.011) 0.063*** (0.011) 0.065*** (0.011) 0.062*** (0.011)</td>
<td>0.009 (0.007) 0.013* (0.007) 0.015** (0.007) 0.011*** (0.007)</td>
</tr>
<tr>
<td>$BBD_{t-1}$</td>
<td>0.130*** (0.043) 0.129* (0.078)</td>
<td>0.261*** (0.078)</td>
<td>0.261*** (0.078)</td>
</tr>
<tr>
<td>$AML_{t-1}$</td>
<td>0.073*** (0.037)</td>
<td>0.012 (0.139)</td>
<td>0.275** (0.140)</td>
</tr>
<tr>
<td>$AMH_{t-1}$</td>
<td>0.076*** (0.002)</td>
<td>-0.020 (0.042)</td>
<td>0.008 (0.031)</td>
</tr>
<tr>
<td>$PSP_{t-1}$</td>
<td>0.044*** (0.006)</td>
<td>0.044*** (0.006)</td>
<td>0.073*** (0.021)</td>
</tr>
<tr>
<td>Within-$R^2$</td>
<td>0.24 0.24 0.24 0.24</td>
<td>0.41 0.41 0.41 0.41</td>
<td>0.40 0.39 0.39 0.39</td>
</tr>
<tr>
<td>Observations</td>
<td>92240 92801 92790 92663</td>
<td>27531 27371 27290 27333</td>
<td>39312 39421 39252 39386</td>
</tr>
<tr>
<td>Premium (bps)</td>
<td>5.1 0.8 0.3 5.7</td>
<td>6.4 0.1 -0.7 1.0</td>
<td>7.4 2.1 0.7 6.6</td>
</tr>
</tbody>
</table>
the magnitudes of liquidity premia. Table 1.6 summarizes the estimation results. Regardless of which illiquidity measure is used in the estimation, liquidity premia are higher post crisis than pre crisis. For example, the premium associated with $BBD$ is 5.1bp pre crisis, but 7.4bp post crisis. The increase in the dispersion of stock illiquidity clearly contributes to those higher premia. In addition, illiquidity premia may be higher today because institutional investors now change positions more frequently, or possibly because it is harder to locate liquidity in today’s more fragmented markets, especially in the presence of HFTs who front run, stealing liquidity.

V Conclusion

Massive regulatory changes and technology advances have transformed U.S. equity markets in the new millennium dramatically altering trading strategies. In today’s high frequency trading markets, individual transactions do not reflect institutional trading decisions in the sense of classical models of microstructure, and traditional measures of liquidity no longer describe liquidity considerations of institutional investors.

Our first contribution is to develop an aggregation approach that groups successive individual transactions into objects (trade sequences) that are economically meaningful and empirically tractable. Our stock-specific measure of trading activity distinguishes the effects of trading volume from those of trading activity. Using this measure, we show that higher trading activity is associated with greater trade sizes and trade imbalances. While classical models of speculation attribute high trading activity and large imbalances to informed trading, these patterns may alternatively reflect endogenous choices to consume liquidity when depth at good prices is high, and to supply

\[^{33}\text{We isolate the financial crisis period to avoid mixing the impact of Reg NMS with the crisis.}\]
liquidity when such depth is lower. Our findings suggest that endogenous consumption of liquidity largely underlies trading activity variations—for small and mid-sized stocks, price impacts of fixed dollar positions decline uniformly as trading activity rises, and for large stocks, price impacts peak at intermediate activity levels.

We then identify a valid instrument for endogenous trading activity. We establish that trading costs monotonically fall as predicted trading activity rises, indicating that institutional investors can save on trading costs by timing trades accordingly. Further examinations reveals that in recent years, as predicted trading activity rises, trading costs decline by comparable amounts for stocks of all sizes—trading costs of differently-sized stocks have become similarly sensitive to stock-specific trading activity.

Levels of price impacts evolve very differently, however. We use our price impact measure to construct a measure of stock illiquidity, and show that while trading costs of more liquid stocks declined post RegNMS, those of less liquid stocks rose. That is, markets did not become uniformly more liquid. This raises the question—since the costs of trading more and less liquid stocks have diverged, what has happened to the return premia that less liquid assets must offer to attract investors? To answer this question, we first verify that our measure of illiquidity is priced in an augmented four-factor CAPM, and that it outperforms standard illiquidity measures. We conclude our analysis by showing that illiquidity premia in post financial-crisis years have risen past pre-crisis levels.
Chapter 2

Intraday Market Dynamics

I Introduction

Recent revolutionary changes in the design of financial markets have radically altered trading strategies and outcomes. Tiny tick sizes have led to bid-ask spreads of pennies and vast reductions in limit book depth near inside quotes. Advances in computing power and programming have raised trading speeds beyond human capacities, making algorithmic trading feasible, and resulting in exploding trading volume. Together with the implementation of RegNMS, these changes have given rise to a new market participant, the high frequency trader (HFT), who exploits information-processing and trading speed advantages in fragmented markets to front-run large investors (Hirschey (2013)). As a result, institutional investors, when establishing or unwinding positions now finely split orders over time and across exchanges, mixing liquidity-taking and liquidity-making orders.

In such an environment, intra-day trading and price dynamics consist of more than stochastic fluctuations, price discovery near open, or the well-understood U-shaped time-of-day trading intensity patterns. In this paper, I study those dynamics.

I ask and answer important fundamental questions: what are the consequences of this revolution for intraday trading activity, trading and pricing dynamics? what systematic dynamic
relationships, if any, exist? How have the dynamic properties of the market been affected by the evolving trading environment, financial crisis, and so on? Concretely, when activity rises, what happens to the correlation structure of returns, and signed trade imbalances? when activity falls, what happens? how do correlation structures hinge on the time of day? and how has the advent of HFTs altered intraday market dynamics? Perhaps due to the daunting data challenges involved, the profession knows little about the answers to these questions.

To assess intra-day market dynamics properly, one must first identify appropriate measures of the fundamentals—activity, trading volumes, price changes, and signed trade imbalances. The strong intertemporal dependence of trades complicates measurements. Trading decisions are made at lower frequencies than the tick-by-tick frequencies observed in the data. Excessively high-frequency measurements, e.g., trade-by-trade or minute-by-minute, are subject to biases associated with the temporal dependence of trades driven by dynamic strategic order-splitting by institutional traders, and the shredding of marketable orders against the book (see e.g., the concerns expressed by Gouriéroux, Jasiak, and Le Fol (1999)). Conversely, low-frequency measurements, e.g. hourly, are likely to mix different market conditions—a one-hour time interval may span very different trading activity levels. Further, if volume varies with the extent of measured trading activity—e.g., if trading activity is measured over constant time periods—then distinguishing the distinct effects of trading activity from those of trade volume becomes problematic.

This leads me to create a trade-based measure of stock-specific market activity that groups consecutive trades together into a sequence of trades with minimum cumulative dollar value (concretely, $80,000 +0.025\%$ of the stock’s market capitalization). Once a stock’s target dollar value is reached, I start over, grouping the next set of consecutive trades until their cumulative dollar value reaches its target value, continuing in this way iteratively over a year. By choosing the levels of the
fixed dollar (e.g., $80,000) and fixed fraction of market capitalization (e.g., 0.025%) appropriately, for any sized firm, I can address concerns about aggregating over multiple activity levels due to excessively large targets, and address concerns about mis-attributing variation in activity due to intertemporal dependence of trades when target dollar values are too small.

Crucially, this approach controls for the underlying relevant economic variable, dollar volume, making the fundamentals, like returns, comparable across different trade sequences. I then use the time durations of these trade sequences to measure trading activity—shorter time durations for the execution of comparable dollar volumes reflect higher activity.

My first step is to establish that trading activity matters for returns: I document that, on average, the return of a trade sequence in more active markets (shorter time durations) is positive, but the average return in less active markets is negative—less active markets seem to reflect reduced “market interest”.

This leads naturally to the question: Does it matter how the market reaches a given level of activity? As a first pass to answering this question, I group the sample of trade sequences into pairs of successive sequences, and then decompose the sample according to whether the second trade sequence of a pair represents an *increase* or *decrease* in activity, relative to the first trade sequence. Finally, I estimate the stock-level return correlation structure when the concluding trading activity in the second trade sequence represents an increase in activity, and compare it to the return correlation structure for decreases in trading activity. I establish that relative *increases* in activity are associated with strikingly high *return persistence*, but relative *reductions* in trading activity are associated with sharp *price reversals*. These strong patterns are qualitatively independent of stock attributes and time. By analyzing dynamics at the stock-year level, I preclude heterogeneity-
driven biases such as small or illiquid stock effects; by limiting my analysis to intra-day market evolutions, I exclude open-to-open or close-to-close phenomena and by basing statistical inference on the empirical distributions of stock-level point estimates, I ensure that the findings are robust to arbitrary heterogeneity in stock-specific error structures.

At one level, these findings bear a surface similarity to those in studies that establish predictable price dynamics at very low (e.g., daily or monthly) frequencies. For example, Jegadeesh and Titman (1993) and Jegadeesh, and Lakonishok (1996) document return persistence, attributing them to under-reaction to news; while Lehmann (1990), Lo and MacKinlay (1990), and Jegadeesh (1990) find price reversals. My results also bear a surface similarity to findings of Connolly and Stivers (2003) who study weekly portfolio return auto-correlations. Examining the correlation between returns realized over two consecutive weeks, they document that returns persist when trading activity is abnormally high in the latter week, but that prices revert when the latter week’s activity is abnormally low. They attribute these variations in return auto-correlation to news and Macroeconomic shocks. However, the driving forces underlying the phenomena that I uncover are very different—my findings reflect purely market microstructure forces at the individual stock level.

The systematic strong return persistence/reversals that I find given particular trading activity dynamics also connect my study to the literature that seeks to predict high frequency trading activity. Barardehi, Bernhardt, and Davies (2014) provide strong evidence that proxies of past order book states predict activity, as measured in this study. Engle and Russell (1998) and Engle (2000) develop models of autoregressive conditional duration (ACD) to model the time duration between individual trades. These models estimate/predict the expected cost and time to execute a single order. Such models can be applied to the measures of trading activity used here.

1 My empirical design excludes the first and last few minutes of a trading day.
2 In ongoing work, I use ACD models to provide sophisticated predictions of trading activity. I consider multi-asset
The relationship between the nature of the change in trading activity and the consequences for return persistence or reversals beg for sound explanations. I consider two alternative explanations for the core finding of price reversals when activity falls. It could be that episodes of high activity reflect intense private-information driven order flow. The magnitude of information content may be unknown to the market, so prices may revert when activity declines and “the market” concludes that the information content was less than it could have been. Alternatively, a reduction in activity could reflect a decline in demand for liquidity. If so, higher previous liquidity demand may have driven prices further from fundamentals as liquidity providers are compensated for their services, and prices move toward their fundamental values when activity declines. Such liquidity-based price impacts are at the heart of the classical Chriss-Almgren (2001) model that underlie the trading strategies of many HFTs.

So, too, the return persistence of rising activity could reflect investors aggressively trading when they acquire substantial private information, driving up trading activity and persistently moving prices in one direction. Alternatively, the increased trading activity could reflect continued endogenous consumption of liquidity when the market is unusually deep on the other side.

Studies attributing price and activity dynamics to liquidity provision include Campbell et al. (1993) and Conrad, Hameed and Niden (1994) who argue that an increase in trading activity that is coupled with a price change may reflect non-informational trade whose subsequent price level is expected to revert. Jegadeesh and Titman (1995) argue that short-term return reversals reflect dealer-inventory effects. Some studies use price reversals to measure liquidity: Pástor and

---

3 Kraus and Stoll (1972) were among the first to notice that prices deviate from fundamental values when institutional traders make large trades. They suggest that reversals reflect the cost of institutional trades. Empirical facts about traditional intermediary’s market-making behavior match dynamic inventory control models (Ho and Stoll (1981)). Hasbrouck and Sofianos (1993) document inventory mean reversions. Hendershott and Seasholes (2007) and Hendershott and Menkveld (2013) document short-run price reversals following intermediary’s non-zero net positions at close.
Stambaugh (2003) exploit average daily return reversals to construct a market-wide liquidity factor and explore liquidity risk. Nagel (2012) builds on models that imply a negative association between volatility and market maker’s funding constraints (Brunnermeier and Pedersen (2009)) to show that returns to liquidity provision predictably rise in times of financial turmoil.

Other researchers put forth informational-based arguments for pricing dynamics at high frequencies. Dufour and Engle (2000) use the time duration between consecutive trades to measure activity; They find faster price adjustments in times of high trading intensity, and interpret it as informed trades boosting trading intensity. Bianco, Corsi, and Renò (2009) document intra-day positive relation between volatility and return auto-correlations.

I establish that the price reversals and price persistence phenomena that I find are liquidity driven. I find that price reversals of falling activity are strongest in least active trading markets, even when conditioning on starting activity level. This result is striking because the lower is the starting activity level, the smaller is the feasible subsequent reduction in trading activity. This is strong evidence against information-based explanations, which predict that informed trade is more likely when activity is higher (e.g., Kyle (1985), Glosten (1994)). In contrast, it is natural to presume that the starting activity level is lower because less liquidity is available; and with less competition to provide liquidity, liquidity demand drives prices further away from fundamentals.

I further establish that the strength of price reversals is stronger when signed trade is more balanced, not less. This is inconsistent with information-based explanations—informed agents will pay to consume liquidity to make purchases before their information leaks out; but consistent with a scenario in which an institutional trader responds to limited liquidity provision by altering the composition of his orders toward taking less liquidity and providing more. I also establish that
return persistence of rising trading activity is strongest when signed trade imbalance is highest, and that signed trade imbalances rise with activity. These patterns highlight the endogeneity of trading activity—when liquidity is unusually high, traders will aggressively consume it, giving rise to active markets—and emphasize the necessity of controlling for this endogeneity in analyses of price impacts (see e.g., Easley et al. (2012)).

My findings indicate that intra-day price reversals reflect rewards to short-term liquidity provision that can involve taking and leaving positions at intra-day frequencies. As such, my paper in connected to the literature that recognizes market-making role of HFTs. Hendershott, Jones, Menkveld (2011), Menkveld (2013), and Hagström and Nordén (2013) document various market-making properties of HFTs such as a negative association between mid-quote price and net positions, the desire to hold net zero position at open and close, and increasing market-making activity in tick size.

These findings lead me to investigate how intraday market dynamics vary over the course of a trading day. Reflecting the U-shaped pattern of trading activity over a trading day, relative increases in activity are least (most) likely early (late) in the day; and relative decreases are most (least) likely early (late) in the day. This might lead to a conjecture that, because reductions in trading activity are most likely earlier in the day, reversals will be greater then. In fact, the price reversals associated with declining activity rise over the trading day. This result strongly suggests that liquidity providers have targets for closing positions, and that since providing liquidity later in the day makes it more difficult to attain these positions, they must receive increased compensation, resulting in greater reversals. I then establish that the return persistence associated with increasing trading activity is greatest earlier in the day, and almost vanishes in the last two hours of a trading day. These results suggest that increases in activity early in the day reflect previously-planned
unwinding by institutional traders of large positions that increase trading activity with enduring price impacts.

I conclude by exploring how the changing market structure and advent of HFTs who have supplied greater shares of extant liquidity in recent years have altered the structure of intraday market dynamics. HFTs differ from the specialists whom they crowded out in many ways. In particular, HFTs have more limited capital and supply liquidity to many more stocks. This suggests that HFTs have far shorter rebalancing horizons than historical market makers. I provide evidence that as time has passed, the rebalancing horizons of liquidity providers has gotten shorter and shorter, systematically altering intraday market dynamics, even though the extent of return reversals in declining activity markets and return persistence in rising activity markets have not.

I first provide an aggregate picture of intra-day price dynamics that could suggest structural changes over time. I investigate whether stocks with stronger price reversals associated with falling activity also feature stronger return persistence associated with rising activity, and how this relationship has evolved over time. I find that these two phenomena are in fact strongly negatively related in early years of my sample. However, this commonality vanishes post 2009.

Motivated by the secular reduction in commonality, I then explore whether the dynamics of aggressive trading have changed between the earlier (2001-2006) and later (2007-2012) time periods in a way indicative of the shorter rebalancing horizons of HFTs. Controlling for concluding trading activity levels, I establish that in the 2001–2006 period, aggressive trading moves in the opposite directions of past price impacts: when trading activity falls after positive returns, the proportion of buyer-initiated trade falls; and after negative returns, the proportion of buyer-initiated trades rises. I show that the exact opposite holds true in later years: when activity falls after positive
returns, the proportion of buyer-initiated trade rises; and after negative returns, the proportion of buyer-initiated trades falls. These effects are especially pronounced later in the trading day for mid-sized and large stocks, where HFTs are most active. These pronounced changes indicate that liquidity providers rebalance positions more and more quickly after providing liquidity in recent years, highlighting the impact of HFTs who provide increasing shares of liquidity over time, and the importance HFTs place on balancing positions by market close.

II Reconstructing comparable trade sequences

To begin, I detail how I construct trade sequence in a way that controls for variations in dollar volume. I then explain how I measure different fundamentals over trade sequences. I number sequential trades of stock \( j \) in a year by index \( n_j \). Corresponding to trade \( n_j \), (a) \( \tau_j(n_j) \) denotes the time of the trade measured in seconds from the beginning of the year, (b) \( Q_j(n_j) \) denotes the number of shares traded, and (c) \( P_j(n_j) \) denotes the transaction price. To group trades into comparable dollar-volume buckets, I identify successive trade sequences with cumulative value of at least \( V_{j,t} = \bar{V} + \theta M_{j,t-1} \), where \( \bar{V} \) is a fixed dollar amount, \( \theta \) is a fixed proportion, and \( M_{j,t-1} \) is the value of shares outstanding at the end of the previous month.

To identify trade sequences, I only need to find the last trade that belongs to a sequence. Each year, I set \( n_j^0 = 0 \) for each stock \( j \). I then iteratively solve equation (2.1) to identify the last trade of sequence \( k \) of stock \( j \), for \( k = 0, 1, 2, \ldots \).

\[
  n_j^k = \arg\min_{n^*} \left\{ \sum_{n=n_j^{k-1}+1}^{n^*} P_j^c(n) \times Q_j(n) \bigg| \sum_{n=n_j^{k-1}+1}^{n^*} P_j^c(n) \times Q_j(n) \geq V_{j,t} \right\} , \tag{2.1}
\]

To ensure that current price evolutions do not alter identification of sequences, I use the closing
price on the previous trading day \((P_j^C(n))\) to calculate dollar trading volumes.

A: Dollar volume

B: Cumulative dollar volume and trade sequences

Figure 2.1: Calculations of trade sequences and returns for successive trade sequences.

Figure 3.1 illustrates how I identify trade sequences and their corresponding fundamentals. I measure three fundamentals over each trade sequence \(k\) for stock \(j\):

1. **Stock-specific market activity:** To capture trading activity, I use the time duration of trade sequence \(k\):

   \[
   D_j(k) = \tau_j(n_k^j) - \tau_j(n_{k-1}^j). \tag{2.2}
   \]

   Shorter time durations reflect higher activity levels.

2. **Returns and price impacts:** The return realized for the fixed dollar volume traded in trade

4I use the midpoint at close if the closing price is not available.
sequence \( k \) is

\[
 r_j(k) = \frac{P_j(n_j^k)}{P_j(n_j^{k-1})} - 1. \tag{2.3}
\]

I use the absolute value of a return to approximate the price impact of trading the fixed dollar position.

3. **Signed trade imbalance**: I use the Lee and Ready (1991) algorithm based on contemporaneous mid-quote prices to classify trades according to whether they are buyer- or seller-initiated. In each trade sequence, I then weight signed trade \( n \) by its corresponding dollar trading value, \( P_j(n)Q_j(n) \). I define the signed trade imbalance for a trade sequence as the proportion of buyer- or seller-initiated dollar volume in the sequence (whichever is highest).\( ^5 \) I sometimes further distinguish the split between the proportions of buyer- and seller-initiated trades.\( ^6 \)

I construct trade sequences that span multiple trading days. However, I exclude them from my sample to avoid including overnight effects and to avoid ambiguity about which day’s closing price is used to calculate the relevant dollar volume used in the construction of a trade sequence. This filter contributes in several ways: (a) it removes the effect of overnight price adjustments that are not a focus of this paper interests; (b) it randomizes the starting (ending) time of the first (last) trade sequences of trading days, reducing the systematic market-at-open (market-at-close) effects;\( ^5 \) The total dollar volume traded over a sequence \( k \) is

\[
 DVOL_j(k) = \sum_{n = n_j^{k-1}+1}^{n_j^k} P_j(n) \times Q_j(n), \tag{2.4}
\]

which usually slightly exceeds \( V_{j,t} \) since the last trade size in a sequence typically slightly exceeds the quantity required to deliver the target cumulative value.

\( ^6 \) Carrion and Kolay (2014) use NASDAQ HFT data to validate the Lee-Ready algorithm in fast trading markets. They find that Lee-Ready algorithm correctly classifies more than 85% of trades, and that its performance improves as trade sizes rise. Concerns about omission of tiny odd-lot trades (O’Hara, Yao, and Ye (2014)) are addressed by weighting signed trades by dollar volume, so they would comprise a small portion of any aggregate dollar volume; that is, my measure of signed trade imbalance benefits from the better performance of the Lee-Ready algorithm in signing larger trades. Moreover, aggregation across trades, as Easley, Lopez de Pardo, and O’Hara (2012) point out, further reduces classification errors. Finally, any classification/measurement errors would seem largely just serve to add noise that would weaken the magnitudes of the observed relationships between stock-specific market activity and signed trade imbalances.
and (c) it tends to exclude the first few trades at open where most price discovery occurs.

In some of the analysis, I sort trade sequences according to a fundamental in order to control for or capture relevant variations. I always sort observations stock-by-stock on a monthly basis to account for monthly target dollar volume updates. Reflecting varying purposes and sample size considerations, the sorts can be as coarse as quintiles, or as fine as percentiles. These sorts facilitate cross-stock comparisons/aggregations by standardizing variations across stocks and time.

A key component of my analysis is the choice of the trading dollar volume target for stock \( j \) in month \( t \), \( V_{j,t} \). I present results for the base case where \( \bar{V} = 80,000 \) and \( \theta = 0.025\% \). However, qualitatively similar results obtain with remarkably different parameter values, e.g., \( \bar{V} = 50,000 \) and \( \theta = 0.015\% \), or \( \bar{V} = 0 \) and \( \theta = 0.03\% \). The choice of the level and composition of \( V_{j,t} \) reflect several considerations: (a) the fixed dollar component of \( V_{j,t} \) reduces the noise in trade sequences for small firms, while the fixed market capitalization component reduces the noise in trade sequences for large firms; (b) the level of \( V_{j,t} \) delivers positions that are large enough to be relevant for institutional traders both for small and for large stocks, (c) the resulting sequences include enough trades to control for dynamic order-splitting and shredding of orders against the book, (d) trade sequences are short enough to minimize the extent of aggregating over different market conditions, and (e) the specification roughly fixes target dollar volume over time, so that despite the explosion of trading volumes in the past decade, trade sequences remain comparable over time.

This last feature—that I fix dollar volume for a stock when aggregating—is a central distinguishing feature of my approach from those adopted by other researchers. Some form of aggregation over trades is necessary to deal with the temporal dependence of trades. Amihud (2002) aggregates trading volume over time (days, but one could construct an intraday analogue of Amihud’s mea-
sure by aggregating over shorter time periods (e.g., half hour or hour). Alternatively, Easley et al. (2012) fix the number of trade sequences in a trading day over time and across stocks (e.g., they construct 50 sequences (“bundles”) each day with identical dollar volumes).

With both of these approaches, the dollar value of bundles will vary radically across trading days and over time due to the tremendous variations in daily trading volumes. Such varying dollar volumes would preclude the analyses I do—one would not be able to determine whether prices move due to higher trading activity or due to higher trading volume. Concretely, the price impacts of fixed dollar volumes tend to fall as activity rises (Barardehi, Bernhardt and Davies (2014)); but price impacts and volatility rise with trading volume (see e.g., Anderson and Bondarenko (2014)). As important, measures of activity would not be comparable across days—moderate dollar volume would represent high trading activity on a low volume day, but low trading activity on a high volume day—much less over time, due to the explosion in trading volumes over the past decade. Moreover, the relevant object for institutional traders is the price impact of the given dollar volume associated with the position that they are trying to establish or unwind.

III Data

The sample period spans January 1, 2001 to December 31, 2012. At the beginning of each year, I form a sample of U.S.-based stocks maintaining a minimum daily closing price at least $1. I extract previous month’s market capitalization $M_{j,t-1}$ for each stock from the CRSP monthly stock file. Market capitalization is defined as the product of the shares outstanding and the closing price on the last trading day of the previous month. I use beginning of the year market-capitalizations to construct size deciles in order to control for size heterogeneities. The previous day’s closing price
used to calculate intra-day dollar volumes is obtained from the CRSP daily stock file. Finally, stock years with fewer than 1000 trade sequences are excluded.

I extract tick-by-tick trade volumes, prices, and time stamps from the monthly TAQ database and best quoted prices from the TAQ National Best Bids and Offers (NBBO). I exclude trades that do not occur between 9:30AM and 4:00PM, except for those representing delayed reporting of market-on-close orders. I filter to remove trade sequences with realized absolute returns (|r_j(k)|) that exceed 10% (average durations are under 40 minutes, so such high returns are either data errors or reflect large news arrival).

IV Results

Each year, I measure the median time duration of trade sequences, stock by stock. Each stock is assigned to a size decile based on the beginning of the year market-capitalizations; size decile 1 contains the smallest stocks, and higher size deciles contain larger stocks. To illustrate how stock-specific market activity has evolved over time for different sized firms, I average the stock-specific median time durations across all stocks in a size decile group on an annual basis. Figure 2.2 shows that the average time duration is a ∪-shaped function of stock size and that average time duration falls over time, reflecting the large increases in trading volume, especially for mid-sized and larger stocks. These time durations are as long as 34 minutes in 2001 for the smallest stocks and as short as 8 minutes in 2008 for the relatively large stocks in size decile 8. Reductions in average time duration between 2001 and 2012, range from 5% for the decile of smallest stocks to 50% for mid-sized firms—with an average decline of about 40%.

Next, I derive how returns realized over a trade sequence vary with stock-specific market activity,
Figure 2.2: **Time duration of trade sequences by year and by firm size.** The figure reports the average median duration (in minutes) of a sequence of trades with a cumulative aggregate value of at least $80,000 plus 0.04% of a firm’s market capitalization. The median duration is first calculated on a stock-by-stock basis, and then the average is taken across stocks within a size decile in a year. Size decile 1 contains the smallest firms; decile 10 contains the largest.

where I emphasize that the dollar value of these trade sequences is fixed for a given stock year. To do this, every month, I sort trade sequences of each stock by time duration to create ten deciles of trade sequences, where higher activity deciles have shorter time durations. For each stock, I calculate the mean return for each activity decile, annually. To capture the relationship between stock-specific market activity and average returns of fixed dollar positions, I average these stock-specific mean returns across time and stocks by activity group for each size decile.

**Result 1: Returns are negative in less active markets, but positive in active markets.**

The top row in Figure 2.3 shows that returns of fixed dollar positions tend to rise with stock-specific market activity. Indeed, returns are **negative** in less active markets, becoming positive only once activity rises past the 35th percentile for smaller stocks and the 40th percentile for larger stocks.
Figure 2.3: **Average return by activity group.** Average return trade sequences with cumulative value of $80,000 plus 0.025% of market capitalization in 2001–2012 versus activity level by stock size categories. Reported average return is the average of firm-specific annual mean returns (realized over trade sequences), computed per activity group. The average of means is taken across all firms in a given size decile, over the entire sample period by activity group.
For small and mid-sized stocks, returns tail off in very active markets; and for small stocks, returns are very negative in very inactive markets. These results suggest that more active markets tend to reflect greater investor interest, and hence rising share prices; while less active markets tend to reflect waning investor interest and hence declining prices. The extent to which magnitudes of returns vary with trading activity is surprising given the high frequency levels at which I am measuring activity; as is the extent of negative returns in low activity markets, given that unconditionally firm returns are positive.

The bottom row of Figure 2.3 shows that these patterns are predictable: conditioning only on past levels of trading activity reveals that mean returns rise with lagged activity. In particular, for small- and mid-sized firms, more active past markets are associated with positive future returns, while less active past markets are associated with negative future returns. Indeed, roughly two-thirds of the variation in mean returns with activity remains when I use predicted activity rather than activity itself.

A Changes in stock-specific market activity and the correlation structure of returns

To begin my main analysis, I investigate whether the evolution of prices over successive trade sequences of a given stock depends on how the market reaches a particular level of activity. More specifically, I explore the correlation structure of returns realized over successive (two) trade sequences conditioning on different patterns of trading activity dynamics: increases versus decreases in activity. Successive trade sequences represent an increase in trading activity if the time duration in seconds of the second sequence is shorter than the first; and they represent a decrease in activity otherwise. I estimate the correlation structure of returns realized over two successive trade
sequences of each stock $j$ in each year $y$:

\[ r_{jy}(k) = \alpha_{jy} + \rho_{jy} r_{jy}(k - 1) + \epsilon_{jy}(k). \]  

(2.5)

Here $r(k)$ and $r(k - 1)$ represent returns realized over two successive trade sequences on the same trading day. I begin by decomposing the sample into pairs of successive trade sequences according to whether there was an increase or decrease in stock-specific market activity. I then estimate model (2.5) for each sub-sample.

I now show that increases in activity give rise to radically different return dynamics than decreases in activity. To distinguish between the two sets of estimates, I use $\hat{\rho}$ to denote the estimated correlation coefficient on returns for the subsample of increases in activity and $\tilde{\rho}$ for the subsample of activity reductions. To make statistical inferences, I use the empirical distributions of these estimates across stocks and years.

**Result 2: Relative increases in trading activity are associated with persistence in returns.** The top row of Figure 2.4 presents the empirical distributions of the correlation parameter $\tilde{\rho}_{jy}$ estimates for increases in activity. I present the estimates separately for small (bottom 30%), medium (middle 40%), and large (top 30%) sized stocks, both for the 2001–2006 and 2007-2012 time periods, where the market capitalizations are calculated at the beginning of each year. The bottom row presents the empirical distributions of the corresponding t-statistics that show the stock-level precisions of the $\tilde{\rho}_{jy}$ estimates.

The estimates reveal that increases in trading activity are associated with extensive persistence

---

7My estimation strategy does not have typical time-series properties because (1) returns are measured over irregularly-spaced points in time; and (2) the decomposition of the sample to isolate increases from decreases in activity is inherently discontinuous in nature.
Figure 2.4: Empirical distributions of correlation parameter estimates and the corresponding t-statistics, isolating increases from decreases in stock-specific market activity. Model 2.5 is estimated year-by-year, stock-by-stock, on subsamples representing increases in activity (yielding $\rho_{jy}$) and decreases in activity (yielding $\hat{\rho}_{jy}$). The top row reports the empirical distributions of $\tilde{\rho}$ and $\hat{\rho}$, separately for stocks falling in the bottom 30%, middle 40%, and top 30% of beginning-of-the-year market-capitalizations, for the periods 2001–2006 and 2007–2012. The bottom row exhibits the corresponding t-statistics with vertical lines standing for the 95% critical values for relevant one-sided significance tests. The cumulative trade value over each trade sequence is $80,000 plus 0.025\% of market-capitalization.
of returns—price changes in successive trade sequences tend to have the same sign. The average of all individual estimates of \( \hat{\rho}_{jy} \) is 0.053 with a t-statistic of 121.46, and 67% of the individual t-statistics exceed the 95% upper-tail critical value.\(^8\) As the empirical distributions indicate, this result holds across all stock sizes and time periods. Panel A of Table \( \text{2.1} \) presents sample means and 99% confidence intervals for \( \hat{\rho} \), based on the empirical distributions, by size decile in the 2001–2006 and 2007–2012 time periods. Returns exhibit significantly positive persistence in all size deciles in both time periods. Mean return persistence ranges between 0.037 and 0.08. In both periods, the magnitude of the persistence in returns with relative increases in trading activity exhibits a \( \cup \)-shaped relationship with stock size. These results indicate that if the current level of trading activity exceeds the past level then, on average, returns of fixed dollar positions move in the same direction. Table \( \text{2.2} \) presents means and 99% intervals for \( \hat{\rho} \) by year, showing that qualitative findings do not vary over time.

\textbf{Result 3: Relative decreases in trading activity are associated with reversals in returns.} The top row of Figure \( \text{2.4} \) presents the empirical distributions of the correlation parameter \( \hat{\rho}_{jy} \) estimates for decreases in activity. The bottom row presents the empirical distributions of the corresponding t-statistics that show the stock-level precisions of the \( \hat{\rho}_{jy} \) estimates. The correlation parameter estimates associated with decreases in activity are even larger than those associated with increases, especially in more recent years, but they have the \textit{opposite} sign—the mean of all individual estimates of \( \hat{\rho}_{jy} \) is -0.069 with a t-statistic of -152.89, and 75% of individual t-statistics fall below the 95% lower-tail critical value. Price reversals are found for all stock sizes in both time periods. Panel B in Table \( \text{2.1} \) reports the sample means and 99% confidence intervals for \( \hat{\rho} \) by size.

\(^8\)I do not make assumptions about the stock-specific error term structures. One should therefore not rely on individual t-statistics for statistical inference. Instead, base statistical inference on the empirical distributions of point estimates; as such, inference is robust to the structure of stock errors.
decile and time period. The sample means of statistically significant price reversals fall between $-0.106$ and $-0.055$. These results indicate that if the current level of trading activity is less than the past level then, on average the returns of the successive fixed dollar positions move in opposite directions. Table 2.2 reports means and 99% confidence intervals of $\hat{\rho}$ by year, revealing that the extent of these price reversals do not vary systematically over time.

Together, Results 2 and 3 indicate that for return dynamics, it matters how we get to a given level of stock-specific market activity.\footnote{I replicate this exercise for more extreme change in activity. When I focus on changes that exceed ten activity percentiles (defined in the text), the mean correlation parameter estimate for falling activity, $\hat{\rho}$, rises by over one-third in magnitude to $-0.098$, and that for rising activity, $\hat{\rho}$, goes up by over one-third to $0.069$.}

**Result 4: The correlation structure of returns does not vary with the sign of the return in the initial trade sequence.** Unconditionally, returns rise with the level of trading activity. One might therefore wonder whether the persistence in returns found when trading activity rises, and the return reversals found when activity falls may evolve asymmetrically according to the sign of the return in the initial trade sequence.

To investigate the robustness of the findings to the initial direction of price changes, I control for both the direction of change in activity and sign of the return realized over the first trade sequence (\(r_{jy}(k - 1)\)) in model 2.5. Figure 2.5 shows that the empirical distributions do not vary with the sign of the initial return. Interval estimates of $\hat{\rho}$ and $\hat{\rho}$ given opposite signs of past returns reveal that even though the two confidence intervals of each parameter do not quite overlap, they differ only marginally: the 99% confidence intervals of $\hat{\rho}$ given positive and negative past returns are $(0.037, 0.040)$ and $(0.031, 0.034)$, respectively, while those for $\hat{\rho}$ are $(-0.086, -0.082)$ and $(-0.095, -0.091)$.\footnote{Note that sample selection (positive versus negative past returns) biases estimates of the correlation parameter in predictable ways. Past positive returns will bias the intercept upward, while past negative returns will bias the}
Table 2.1: Sample mean and confidence intervals of correlation parameters, isolating increases from reductions in stock-specific market activity. Model 2.5 is estimated each year, stock by stock, on sub samples representing increases in activity (yielding $\hat{\rho}_{jy}$) and decreases in activity (yielding $\hat{\rho}_{jy}$). Sample means and 99% confidence intervals of correlation parameters ($\hat{\rho}$ in panel A and $\hat{\rho}$ in panel B) are reported by size group in periods 2001–2006 and 2007–2012. The cumulative trade value over each trade sequence is $80,000 plus 0.025% of market-capitalization.

<table>
<thead>
<tr>
<th>Size decile</th>
<th>$\hat{\rho}$ 2001–2006</th>
<th>99% CI</th>
<th>$\hat{\rho}$ 2007–2012</th>
<th>99% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.080</td>
<td>(0.073 , 0.087)</td>
<td>0.069</td>
<td>(0.063 , 0.075)</td>
</tr>
<tr>
<td>2</td>
<td>0.056</td>
<td>(0.050 , 0.063)</td>
<td>0.054</td>
<td>(0.049 , 0.060)</td>
</tr>
<tr>
<td>3</td>
<td>0.043</td>
<td>(0.038 , 0.049)</td>
<td>0.049</td>
<td>(0.044 , 0.053)</td>
</tr>
<tr>
<td>4</td>
<td>0.037</td>
<td>(0.031 , 0.042)</td>
<td>0.044</td>
<td>(0.040 , 0.048)</td>
</tr>
<tr>
<td>5</td>
<td>0.043</td>
<td>(0.038 , 0.048)</td>
<td>0.046</td>
<td>(0.043 , 0.050)</td>
</tr>
<tr>
<td>6</td>
<td>0.045</td>
<td>(0.040 , 0.050)</td>
<td>0.049</td>
<td>(0.045 , 0.053)</td>
</tr>
<tr>
<td>7</td>
<td>0.043</td>
<td>(0.038 , 0.048)</td>
<td>0.047</td>
<td>(0.043 , 0.051)</td>
</tr>
<tr>
<td>8</td>
<td>0.047</td>
<td>(0.042 , 0.051)</td>
<td>0.054</td>
<td>(0.050 , 0.058)</td>
</tr>
<tr>
<td>9</td>
<td>0.053</td>
<td>(0.049 , 0.058)</td>
<td>0.057</td>
<td>(0.053 , 0.062)</td>
</tr>
<tr>
<td>10</td>
<td>0.065</td>
<td>(0.060 , 0.070)</td>
<td>0.067</td>
<td>(0.063 , 0.072)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Size decile</th>
<th>$\hat{\rho}$ 2001–2006</th>
<th>99% CI</th>
<th>$\hat{\rho}$ 2007–2012</th>
<th>99% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.055</td>
<td>(-0.063 , -0.047)</td>
<td>-0.072</td>
<td>(-0.079 , -0.066)</td>
</tr>
<tr>
<td>2</td>
<td>-0.063</td>
<td>(-0.070 , -0.056)</td>
<td>-0.065</td>
<td>(-0.070 , -0.060)</td>
</tr>
<tr>
<td>3</td>
<td>-0.074</td>
<td>(-0.081 , -0.068)</td>
<td>-0.068</td>
<td>(-0.072 , -0.063)</td>
</tr>
<tr>
<td>4</td>
<td>-0.080</td>
<td>(-0.085 , -0.074)</td>
<td>-0.064</td>
<td>(-0.069 , -0.060)</td>
</tr>
<tr>
<td>5</td>
<td>-0.069</td>
<td>(-0.075 , -0.064)</td>
<td>-0.064</td>
<td>(-0.068 , -0.060)</td>
</tr>
<tr>
<td>6</td>
<td>-0.064</td>
<td>(-0.069 , -0.059)</td>
<td>-0.061</td>
<td>(-0.065 , -0.057)</td>
</tr>
<tr>
<td>7</td>
<td>-0.065</td>
<td>(-0.069 , -0.060)</td>
<td>-0.061</td>
<td>(-0.065 , -0.057)</td>
</tr>
<tr>
<td>8</td>
<td>-0.064</td>
<td>(-0.068 , -0.059)</td>
<td>-0.067</td>
<td>(-0.071 , -0.064)</td>
</tr>
<tr>
<td>9</td>
<td>-0.069</td>
<td>(-0.073 , -0.065)</td>
<td>-0.080</td>
<td>(-0.083 , -0.077)</td>
</tr>
<tr>
<td>10</td>
<td>-0.079</td>
<td>(-0.083 , -0.075)</td>
<td>-0.106</td>
<td>(-0.109 , -0.102)</td>
</tr>
</tbody>
</table>

These findings are very strong, statistically. The stock-specific estimates allow for arbitrary stock-level heterogeneity in the correlation structure of returns. The empirical distributions of these estimates reveal that the impact of stock attributes (stock size) on the correlation structure is modest. Thus, the remarkably strong statistical dominance of return persistence estimates over intercept downward. In all cases, this biases estimates of the correlation coefficient downward; return persistence weakens and price reversals become stronger when I distinguish past positive returns from past negative returns.

\footnote{This is similar to introducing stock-year fixed effects.}

73
Table 2.2: Sample mean and confidence intervals of correlation parameters, isolating increases from reductions in stock-specific market activity by year. Model 2.5 is estimated each year, stock by stock, on subsamples representing increases in activity (yielding \( \tilde{\rho}_{jy} \)) and decreases in activity (yielding \( \hat{\rho}_{jy} \)). Sample means and 99\% confidence intervals of correlation parameters (\( \tilde{\rho} \) and \( \hat{\rho} \)) are reported by year. The cumulative trade value over each trade sequence is $80,000 plus 0.025\% of market-capitalization.

<table>
<thead>
<tr>
<th>Size decile</th>
<th>mean ( \tilde{\rho} ) 99% CI</th>
<th>mean ( \hat{\rho} ) 99% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>0.079 (0.074 , 0.084)</td>
<td>-0.047 (-0.053 , -0.042)</td>
</tr>
<tr>
<td>2002</td>
<td>0.064 (0.060 , 0.068)</td>
<td>-0.055 (-0.060 , -0.050)</td>
</tr>
<tr>
<td>2003</td>
<td>0.044 (0.040 , 0.048)</td>
<td>-0.066 (-0.070 , -0.061)</td>
</tr>
<tr>
<td>2004</td>
<td>0.054 (0.050 , 0.058)</td>
<td>-0.066 (-0.070 , -0.061)</td>
</tr>
<tr>
<td>2005</td>
<td>0.038 (0.034 , 0.042)</td>
<td>-0.082 (-0.086 , -0.078)</td>
</tr>
<tr>
<td>2006</td>
<td>0.037 (0.034 , 0.041)</td>
<td>-0.085 (-0.088 , -0.082)</td>
</tr>
<tr>
<td>2007</td>
<td>0.059 (0.055 , 0.062)</td>
<td>-0.074 (-0.077 , -0.071)</td>
</tr>
<tr>
<td>2008</td>
<td>0.051 (0.048 , 0.055)</td>
<td>-0.080 (-0.084 , -0.076)</td>
</tr>
<tr>
<td>2009</td>
<td>0.053 (0.050 , 0.057)</td>
<td>-0.067 (-0.071 , -0.064)</td>
</tr>
<tr>
<td>2010</td>
<td>0.062 (0.059 , 0.066)</td>
<td>-0.073 (-0.077 , -0.069)</td>
</tr>
<tr>
<td>2011</td>
<td>0.055 (0.052 , 0.058)</td>
<td>-0.067 (-0.071 , -0.063)</td>
</tr>
<tr>
<td>2012</td>
<td>0.042 (0.039 , 0.045)</td>
<td>-0.064 (-0.068 , -0.060)</td>
</tr>
</tbody>
</table>

Figure 2.5: Empirical distributions of correlation parameter estimates for signed past returns, isolating increases from decreases in stock-specific market activity. I estimate model 2.5 each year, stock by stock, on subsamples representing increases in activity (yielding \( \tilde{\rho}_{jy} \)) and decreases in activity (yielding \( \hat{\rho}_{jy} \)), conditioning on the sign of return realized over the first trade sequence, \( r_{jy}(k - 1) \). The cumulative trade value over each trade sequence is $80,000 plus 0.025\% of market-capitalization. Reversals estimates highlighted by the distinct kernel density estimates in Figure 2.4 reflect fundamentally different price dynamics. Moreover, the statistical significance of these estimates holds regardless of stock attributes and time period; they do not hinge on small/illiquid stock anomalies.
or the radical revolution that occurred in stock market design and algorithmic trading. Further, by working with the empirical distributions of stock-specific estimates, I circumvent the necessity of specifying a particular stock-specific error term structure. This means that I do not need to worry about mis-specifying the error structure, which can result in spuriously small standard error estimates at the stock level. Finally, since I restrict the sample to pairs of successive trade sequences that fall on the same trading day, the documented price movements do not reflect open-to-open or close-to-close phenomena.  

B Information or liquidity?

In this section, I investigate and try to understand what drives the striking differences in dynamics. I ask: what are the primitives driving the two types of trading activity and price dynamics? Are these systematic price movements stronger in some market environments that in others? If so, can such variations provide insights into the nature of the underlying market dynamics?

A priori, two plausible alternative explanations suggest themselves for the persistence in returns found as trading activity rises. It could be that investors who have new information about risky fundamentals continue to trade aggressively as long as the value of their information exceeds the cumulative price impact caused by their trades; as a result, in this scenario, trading activity increases while prices persistently move in the direction of aggressive trading. Alternatively, it could be that traders aggressively consume liquidity when markets feature unusually high depth at “good prices”; in this scenario, increases in trading activity could reflect improvements in liquidity

12Recall that, constructing trade sequences, I drop the first few trades after open and the last few trades before close of each trading day (there are on average sixty trades in a trade sequence for the smallest firms and about 800 for the largest).

13In this context, aggressive trading involves submission of orders, primarily marketable orders, that have high probabilities of being executed quickly prior to information leaking out.
that lead to more aggressive consumption of liquidity via marketable orders and persistent price impacts (returns).

So, too, there are two plausible competing explanations for why reductions in trading activity are associated with price reversals. It may be that a relative reduction in activity follows a phase of higher activity that reflects informed trading. Because “the market” does not know the magnitude of this information (when agents are informed), the market does not know when the informed agents will stop trading. Thus, prices may revert when activity falls, as this decline suggests that some of the order flow was not information based and that any private information was smaller than it could have been. Alternatively, it may be that the previous relatively more active phase reflected traders demanding liquidity, and paying for that liquidity demand in the form of greater price impacts. In this scenario, the reduction in trading activity reflects a reduction in the demand for liquidity, and the subsequent price reversals reflect the previous compensation for providing that liquidity, and possibly the unwinding at better prices of the positions that the liquidity providers took on.

I next test to distinguish between these competing scenarios, before investigating more nuanced implications. To distinguish between these scenarios, I investigate how price dynamics vary with the level of stock-specific market activity, exploiting the different predictions of the two scenarios for the magnitudes of price reversals in different market conditions. The first observation that I exploit is that informed trading is more likely in active markets than inactive ones. For example, classical models of speculation (Kyle (1985) and Glosten (1994)), which presume invariant liquidity provision, predict that institutional investors trade more aggressively, establishing larger positions when their private information is more substantial. More generally, informed agents face pressure to establish their positions before their information leaks out, so their order composition is weighted toward marketable orders that consume liquidity and give rise to shorter durations. Thus, one
can attribute short-term price movements to information-driven trading if both return persistence (when activity rises) and price reversals (when activity falls) should be more pronounced in more active markets. In particular, for any reduction in activity, lower initial trading activity levels (less likely due to informed trading activity, and smaller private information) should be associated with smaller price reversals.

Instead, if price dynamics are driven by state-varying liquidity provision, liquidity providers will be able to extract greater compensation for its provision when there is less competition, i.e., in less active markets. As a result, in less active markets characterized by less competition to provide liquidity (and with less likely liquidity demand the opportunity cost of providing liquidity in that market rather than in another market rises), informationless liquidity demand will drive prices further from their fundamental values, as in Campbell, Grossman, and Wang (1993). That is, a given liquidity demand shock induces greater deviations from the fundamental value in less liquid environments, and less active markets feature lower liquidity. Subsequently, prices will return toward their fundamentals, delivering greater price reversals in markets where prices were driven further away.

**Are the price reversals observed as stock-specific market activity fall liquidity or information driven?**

To distinguish between liquidity and information-based scenarios for price reversals, I analyze how the magnitude of price reversals associated with reductions in activity vary with the starting activity level, i.e., the level of trading activity at the first trade sequence, where concluding activity level is the level of trading activity at the second trade sequence. I show that price reversals are greater when the starting activity level is lower. This test is very conservative because
greater declines in activity—possibly due to greater reductions in liquidity demand—are possible when the starting level of activity is higher. Such selection biases estimates away from finding stronger reversals in less active markets.\textsuperscript{14} As such, this represents strong evidence that price dynamics are driven by state-varying levels of liquidity provision rather than information.

To standardize measurement across stocks, instead of working directly with time durations, I create trading activity percentiles. Each month, I sort trade sequences by time duration from longest to shortest.\textsuperscript{15} I then assign activity percentile statistics such that the trade sequence with longest time duration is assigned the smallest, and that with the shortest time duration is assigned the largest activity percentile statistic. For example, if there are 200 trade sequences in a month, after sorting them by time duration, I assign activity percentile 0.005 to the longest time duration, 0.01 to the next to the longest, ..., and 1 to the shortest time duration. The change in activity percentile over two successive trade sequences quantifies activity dynamics: positive (negative) changes reflect increases (decreases) in trading activity.

To capture different levels of general trading activity, for each stock-month I sort trade sequences by time duration into five equally-sized groups. The “Low” activity group contains the longest 20% of trade sequences in a month, and the shortest 20% comprise the “High” activity group. These activity quintiles control for starting/concluding activity levels. For each stock-year, I estimate model (2.5) by activity quintile, focusing on increases versus decreases in activity. Thus, conditioning on the starting or concluding activity level, for each stock-year, I estimate five $\hat{\rho}$’s for increases and five $\hat{\rho}$ for decreases in trading activity. Figure\textsuperscript{2.6} presents the empirical distributions of cor-

\textsuperscript{14}To mitigate this selection, I sometimes restrict the sample to changes in activity of less than 20 percentiles. Such trading activity changes represent roughly two-thirds of the observations.

\textsuperscript{15}Note that monthly sorts reflect two important considerations: (a) the target dollar value $V_{j,t}$ is updated every month, and one wants to identify different levels of activity for a specific fixed-dollar position; (b) by identification of activity variations on a monthly basis, I avoid mixing temporal effects (e.g., the log-term downward trend of activity reflected in Figure\textsuperscript{2.2} with those of stock-specific market activity.
relation parameter estimates given different starting trading activity levels; Table 2.3 presents the corresponding sample means and 99% confidence intervals based on these empirical distributions. To highlight the effects of selection, I also present estimation results for subsamples that exclude increases and decreases in activity that exceed twenty activity percentiles. This more homogeneous subsample of changes in activity facilitates apples-to-apples comparisons across different starting activity levels.

Result 5: Decreases in trading activity are associated with significantly larger reversals when starting activity levels are lower. The empirical densities of estimated price reversals ($\hat{\rho}_{jy}$’s) in the right panel of Figure 2.6 shift towards increasingly negative values as starting activity level falls; and the shifts are stronger when large decreases are excluded. Table 2.3 shows that mean reversals decrease from $-0.052$ to $-0.096$ as starting activity level falls from its highest to lowest levels. When I control for selection by excluding large activity reductions, which biases findings against liquidity-based arguments, the average reversals fall even more sharply from $-0.008$ to $-0.096$ as starting activity goes from its highest level to its lowest. The associated confidence intervals are all very tight and mutually exclusive, implying statistically significant variations. This pattern is the exact opposite of what would be predicted if price dynamics as trading activity declines were driven by information—I can decisively attribute these reversals to rewards for liquidity provision.

Result 6: Increases in trading activity are associated with significantly greater return persistence when starting trading activity levels are higher. The left panel of Figure 2.6 shows that the empirical distributions of estimated return persistence associated with increases in trading activity ($\hat{\rho}_{jy}$’s) shift toward increasingly positive magnitudes as starting activity level
Unrestricted activity changes

\( \hat{\rho}; \) increase in activity

\( \hat{\rho}; \) decrease in activity

Activity changes restricted to less than twenty percentile

\( \hat{\rho}; \) increase in activity

\( \hat{\rho}; \) decrease in activity

Figure 2.6: Empirical distributions correlation parameter estimates that isolate increases from reductions in activity by starting stock-specific market activity level. Model 2.5 is estimated each year stock by stock, for each starting activity quintile, on subsamples representing increases in activity (yielding \( \hat{\rho}_{iy} \)) and decreases in activity (yielding \( \hat{\rho}_{by} \)). The top row displays the empirical kernel densities of estimated correlations with no restrictions on activity changes. The bottom row reports similar densities whose estimates are based on activity changes that are less than twenty activity percentiles.

The bottom left panel reveals that these shifts are greater when restrict the magnitudes of increases in trading activity, implying that selection reduces variation return persistence across starting activity levels. Consistently, Table 2.3 indicates that mean return persistence goes from 0.029 to 0.090 as starting activity level goes from lowest to highest; when I screen for large increases in activity, average return persistence is 0.009 to 0.090 at lowest and highest starting activity levels.
respectively. Remarkably, the tight confidence intervals at different starting activity levels are all mutually exclusive, affirming the statistical significance of these variations. These findings of increasing momentum at higher starting activity levels when trading activity rises, is once more consistent with the persistent impact on prices of increased liquidity demand.

Table 2.3: Sample means and confidence intervals of correlation parameter estimates, isolating increases from reductions in activity by starting stock-specific market activity level. Model $\hat{\rho}_{jy}$ is estimated year-by-year, stock-by-stock, on subsamples representing increases in activity (yielding $\hat{\rho}_{jy}$) and decreases in activity (yielding $\hat{\rho}_{jy}$). Sample means and 99% confidence intervals of the correlation parameter are reported by starting activity group. Panel A reports the sample mean confidence intervals correlation parameters given no restrictions on activity changes. Panel B reports sample means and confidence intervals for activity changes that are less than twenty activity percentiles.

<table>
<thead>
<tr>
<th>Starting activity level</th>
<th>$\hat{\rho}$ mean</th>
<th>99% CI</th>
<th>$\hat{\rho}$ mean</th>
<th>99% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0.029</td>
<td>(0.027 , 0.030)</td>
<td>-0.096</td>
<td>(-0.100 , -0.093)</td>
</tr>
<tr>
<td>Mid-low</td>
<td>0.045</td>
<td>(0.043 , 0.047)</td>
<td>-0.082</td>
<td>(-0.084 , -0.080)</td>
</tr>
<tr>
<td>Medium</td>
<td>0.060</td>
<td>(0.058 , 0.062)</td>
<td>-0.074</td>
<td>(-0.076 , -0.072)</td>
</tr>
<tr>
<td>Mid-high</td>
<td>0.076</td>
<td>(0.073 , 0.078)</td>
<td>-0.063</td>
<td>(-0.065 , -0.061)</td>
</tr>
<tr>
<td>High</td>
<td>0.090</td>
<td>(0.087 , 0.093)</td>
<td>-0.052</td>
<td>(-0.054 , -0.050)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Starting activity level</th>
<th>$\hat{\rho}$ mean</th>
<th>99% CI</th>
<th>$\hat{\rho}$ mean</th>
<th>99% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>-0.009</td>
<td>(-0.011 , -0.007)</td>
<td>-0.096</td>
<td>(-0.100 , -0.093)</td>
</tr>
<tr>
<td>Mid-low</td>
<td>0.009</td>
<td>(0.007 , 0.011)</td>
<td>-0.066</td>
<td>(-0.068 , -0.063)</td>
</tr>
<tr>
<td>Medium</td>
<td>0.035</td>
<td>(0.033 , 0.038)</td>
<td>-0.038</td>
<td>(-0.041 , -0.035)</td>
</tr>
<tr>
<td>Mid-high</td>
<td>0.071</td>
<td>(0.068 , 0.073)</td>
<td>-0.017</td>
<td>(-0.020 , -0.014)</td>
</tr>
<tr>
<td>High</td>
<td>0.090</td>
<td>(0.087 , 0.093)</td>
<td>-0.008</td>
<td>(-0.011 , -0.005)</td>
</tr>
</tbody>
</table>

Robustness Test 1: As a first robustness check, I replicate the exercises above, controlling for concluding activity levels rather than starting activity levels. This investigation sheds light on whether the relationships I documented above hinge on the starting level of activity rather than the general trading activity level.

Figure 2.7 presents the empirical distributions of $\hat{\rho}_{jy}$ and $\hat{\rho}_{jy}$ for five different concluding levels
of trading activity. The top panels present results for unrestricted changes in activity, and the bottom panels focus on subsamples where activity changes are less than twenty activity percentiles. Results are qualitatively similar to when I control for starting activity level in the sense that (1) the extent of return reversals associated with declining activity weaken monotonically when the concluding activity level rises; (2) the return persistence associated with increases in trading activity strengthens in more active market; and (3) these result continue to hold when large changes in trading activity are excluded. Table 2.4 reports sample means and 99% confidence intervals for at different levels of concluding activity with and without screening for large relative activity changes. I find similar variations in strength of return persistence/reversals across different levels of concluding versus starting levels of activity. In particular, when I filter out large changes in activity, the difference between the average $\hat{\rho}$ at "High" and "Low" concluding trading activity levels is 0.097; while the analogous difference is 0.082 when I control for starting trading activity levels.

**Robustness Test 2:** Estimates of the correlation parameters could conceivably reflect the impact of outliers. I now confirm similar patterns hold for probability based tests that are robust to outliers. I explore the conditional probability that the returns realized over two successive trade sequences are of the same sign given increases versus decreases in activity. That is, I derive the impact of different relative changes in trading activity on the likelihood of persisting or reverting returns: investigating whether larger relative increases (declines) of trading activity are associated with greater probabilities of persisting (reverting) returns? These tests provide further evidence that it is the relative changes in trading activity and not the level of activity that drive these intra-day price movements.

To do this, I compute the conditional probabilities at the stock level each year, and then average
across stocks and time. For each stock-year, to address selection, I focus on pairs of successive trade sequences whose associated changes in trading activity are less than twenty activity percentiles. I use activity quintiles to distinguish five activity levels. Within each activity level, observations are classified into four groups according to the magnitudes and signs of the associated changes in trading activity percentile: −20 to −10 and −10 to 0 for large and small reductions; 0 to 10 and
Table 2.4: Sample means and confidence intervals of correlation parameter estimates, isolating increases from reductions in activity by concluding trading activity level. Model [2.5] is estimated year-by-year, stock-by-stock, for each activity quintile, on subsamples representing increases in activity (yielding $\hat{\rho}_{ij}$) and decreases in activity (yielding $\check{\rho}_{ij}$). Sample means and 99% confidence intervals of the correlation parameter are reported by concluding activity group. Panel A reports the sample mean confidence intervals correlation parameters given no restrictions on activity changes. Panel B reports sample means and confidence intervals for activity changes that are less than twenty activity percentiles.

<table>
<thead>
<tr>
<th>Panel A: Unrestricted activity changes</th>
<th>Panel B: Activity changes restricted to less than twenty percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Concluding activity level</td>
</tr>
<tr>
<td>Low</td>
<td>-0.023</td>
</tr>
<tr>
<td>Mid-low</td>
<td>0.016</td>
</tr>
<tr>
<td>Medium</td>
<td>0.042</td>
</tr>
<tr>
<td>Mid-high</td>
<td>0.068</td>
</tr>
<tr>
<td>High</td>
<td>0.090</td>
</tr>
</tbody>
</table>

10 to 20 activity percentiles for small and large increases in activity. I then compute the relative frequency of successive returns of the same sign within each of the twenty categories to measure stock-year level figures. Lastly, I average relative frequencies across stocks and time for each group to draw the aggregate picture. Note, however, that limiting the size of change in trading activity to twenty percentiles still does not eliminate selection concerns: by construction, increases (decreases) that lead to to the highest (lowest) concluding activity quintiles cannot exceed twenty activity percentiles. To emphasizes the caution one should use when comparing these estimates with the others, I attach an asterisk (*) to the conditional probability estimates in these categories.

Table 2.5 presents average likelihood of return persistence for three stock size groups, controlling

---

16In addition to excluding observations where activity changes by more than 20 percentiles, I exclude the few observations where the return was literally zero, as then the meaning of persistence or reversal is not clear.
Table 2.5: Conditional likelihood of return persistence as a function of stock-specific market activity level and the size of activity change. For each stock-year, pairs of successive trade sequences are grouped into quintiles of starting/concluding activity. For each activity quintile, observations are classified into groups of changes in activity percentile: percentile changes are categorized into (−20,−10), (−10,0), (0,10), and (10,20) to reflect large decreases, small decreases, small increases, and large increases, respectively. Relative frequencies of successive returns of same sign are computed at the stock level. Reported conditional likelihoods reflect averages of such relative frequencies across stocks and time by category. (*) symbolizes the categories whose results are subject to potential selection biases.

<table>
<thead>
<tr>
<th>Activity level</th>
<th>By starting activity</th>
<th>By concluding activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Activity percentile change</td>
<td>Activity percentile change</td>
</tr>
<tr>
<td></td>
<td>(−20,−10) (−10,0) (0,10) (10,20)</td>
<td>(−20,−10) (−10,0) (0,10) (10,20)</td>
</tr>
<tr>
<td>Low</td>
<td>0.44* 0.45* 0.46 0.48</td>
<td>0.44 0.45 0.46* 0.49*</td>
</tr>
<tr>
<td>Mid-low</td>
<td>0.44 0.46 0.47 0.48</td>
<td>0.45 0.46 0.47 0.48</td>
</tr>
<tr>
<td>Medium</td>
<td>0.45 0.46 0.48 0.49</td>
<td>0.45 0.46 0.47 0.48</td>
</tr>
<tr>
<td>Mid-high</td>
<td>0.45 0.47 0.49 0.50</td>
<td>0.44 0.47 0.48 0.49</td>
</tr>
<tr>
<td>High</td>
<td>0.42 0.45 0.50* 0.51*</td>
<td>0.39* 0.45* 0.50 0.51</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Activity level</th>
<th>By starting activity</th>
<th>By concluding activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Activity percentile change</td>
<td>Activity percentile change</td>
</tr>
<tr>
<td></td>
<td>(−20,−10) (−10,0) (0,10) (10,20)</td>
<td>(−20,−10) (−10,0) (0,10) (10,20)</td>
</tr>
<tr>
<td>Low</td>
<td>0.43* 0.45* 0.46 0.48</td>
<td>0.44 0.45 0.46* 0.48*</td>
</tr>
<tr>
<td>Mid-low</td>
<td>0.44 0.46 0.47 0.48</td>
<td>0.45 0.46 0.47 0.48</td>
</tr>
<tr>
<td>Medium</td>
<td>0.45 0.46 0.48 0.49</td>
<td>0.45 0.46 0.47 0.48</td>
</tr>
<tr>
<td>Mid-high</td>
<td>0.45 0.47 0.48 0.50</td>
<td>0.44 0.47 0.48 0.49</td>
</tr>
<tr>
<td>High</td>
<td>0.43 0.45 0.49* 0.50*</td>
<td>0.40* 0.45* 0.49 0.50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Activity level</th>
<th>By starting activity</th>
<th>By concluding activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Activity percentile change</td>
<td>Activity percentile change</td>
</tr>
<tr>
<td></td>
<td>(−20,−10) (−10,0) (0,10) (10,20)</td>
<td>(−20,−10) (−10,0) (0,10) (10,20)</td>
</tr>
<tr>
<td>Low</td>
<td>0.42* 0.44* 0.47 0.49</td>
<td>0.43 0.44 0.47* 0.49*</td>
</tr>
<tr>
<td>Mid-low</td>
<td>0.44 0.46 0.47 0.49</td>
<td>0.45 0.46 0.47 0.49</td>
</tr>
<tr>
<td>Medium</td>
<td>0.45 0.47 0.49 0.50</td>
<td>0.46 0.47 0.48 0.49</td>
</tr>
<tr>
<td>Mid-high</td>
<td>0.46 0.47 0.50 0.52</td>
<td>0.45 0.47 0.49 0.50</td>
</tr>
<tr>
<td>High</td>
<td>0.44 0.47 0.51* 0.52*</td>
<td>0.43* 0.47* 0.51 0.52</td>
</tr>
</tbody>
</table>

for the concluding level of trading activity and the size of change in activity percentile. Most average likelihoods are below 0.5, partially reflecting Roll (1984)’s observation that a positive price tick, even of one penny, raises the likelihood of reversals, unconditionally. The most important point to extract is that regardless of firm size or the activity level conditioned upon (or whether starting or concluding activity levels are used to condition), greater increases in activity (or smaller decreases) make persistence more likely (i.e., reversals is less likely). That is, returns are more likely to persist
as trading activity rises by greater amounts; and they are more likely to revert when declines in activity are greater. Thus, I establish the qualitative robustness of my regression-based results.

I establish additional robustness properties in the Appendix, showing that the relationship between the level of trading activity and (a) price reversals associated with increases in activity and (b) return persistence associated with activity reductions is qualitatively unaffected by firm size, controlled for at the stock size decile level.

What do signed trade imbalance reveal?

So far, I have established that price reversals of falling activity reflect rewards to liquidity provision. I next exploit this finding to identify whether higher signed trade imbalances reflect informed trading or endogenous consumption of liquidity. Information-based theories of trade predict that investors submit more marketable orders as they acquire more substantial private information before their information becomes public. In contrast, a liquidity-based explanation that allows for state-varying liquidity and the endogenous choice of whether to consume or provide liquidity, would predict that investors would submit more marketable orders when the other side of the market has significant depth at “good” prices. If this is so, then trade sequences suggesting more consumption of liquidity—trade sequences with greater signed trade imbalances—should feature smaller returns to liquidity provision, and hence smaller price reversals. If, instead, the other side of the market is “less liquid” with less depth near “good” prices, then a trader will alter the composition of his orders toward consuming less liquidity and submitting more liquidity-making orders. Less competition to provide liquidity in such environments should present itself in the form of a more balanced mix of signed trades (the endogenous composition response), and stronger price reversals. I now document precisely such findings.
I investigate the extent of the significance of marketable orders by signed trade imbalance. I first classify trades into buyer- and seller-initiated using Lee and Ready (1991) algorithm. I then find the total values of buyer- and seller-initiated trades comprising a trade sequence. I finally divide the maximum of these two quantities by the realized total dollar volume of a trade sequence to obtain the percent signed trade imbalance. This ratio measures variations in trading via marketable orders at the stock level for fixed dollar positions. Each month, I sort trade sequences by signed trade imbalance into five equally-sized “imbalance” quintiles. Finally, for each stock-year, I estimate Model (2.5) by imbalance quintile in the subsamples of increases and decreases in trading activity—I exclude increases and decreases that exceed twenty activity percentiles. To sharpen the visual contrast, I present empirical distributions of \( \hat{\rho}_{j,y} \) estimates given highest and lowest signed trade imbalance levels. Qualitatively similar findings obtain for less extreme levels of signed trade imbalance.

Result 7: Price reversals associated with reductions in trading activity are weaker when signed trade imbalances are higher.

Result 8: The return persistence associated with increases in trading activity are stronger when signed trade imbalances are higher.

Figure 2.8 presents the empirical distributions of \( \hat{\rho}_{j,y} \) and \( \tilde{\rho}_{j,y} \) when signed trade imbalances are highest, for samples of small, medium and large firms, while Figure 2.9 presents the empirical distributions when signed trade imbalances are lowest. Regardless of stock size and whether one controls for starting or concluding levels of signed trade imbalance, price reversals are stronger when trade is more balanced, but return persistence is stronger when trade is less balanced. Table 2.6

17Findings are robust to relaxing this restriction.
Figure 2.8: Empirical distributions correlation parameter estimates that isolate decreases in activity by starting/concluding signed trade imbalance level for stock size groups. Model (2.5) is estimated year-by-year, stock-by-stock, controlling for concluding/starting imbalance quintiles for the subsample of trade sequences associated with a decrease in activity (yielding $\hat{\rho}_{12}$). Decreases in activity that exceed twenty percentiles are excluded. The top (bottom) row reports the empirical density estimates at the highest and lowest starting (concluding) signed trade imbalances.
Figure 2.9: Empirical distributions of correlation parameter estimates that isolate increases in activity by starting/concluding signed trade imbalance level for stock size groups. Model (2.5) is estimated year-by-year, stock-by-stock, controlling for concluding/starting imbalance quintiles for the subsample of trade sequences associated with a increase in activity (yielding $\hat{\rho}_{ij}^y$). Increases in activity that exceed twenty percentiles are excluded. The top (bottom) row reports the empirical density estimates at highest and lowest starting (concluding) signed trade imbalance.
Table 2.6: Sample means and confidence intervals of correlation parameter estimates, isolating increases from reductions in activity at highest and lowest levels of signed trade imbalance by stock size decile. Model 2.5 is estimated year by year, stock by stock, for the highest and lowest starting/concluding imbalance quintiles, on subsamples representing increases in activity (yielding $\hat{\rho}_{ijy}$) and decreases in activity (yielding $\hat{\rho}_{ijy}$). Increases and decreases in activity that exceed twenty percentiles are excluded. Sample means and 99% confidence intervals of the correlation parameter are reported by stock size decile. Panel A (Panel B) reports the sample means and confidence intervals of correlation parameters at different starting (concluding) signed trade imbalance levels.

### Panel A: By starting signed trade imbalance level

<table>
<thead>
<tr>
<th>Size decile</th>
<th>Low starting imbalance</th>
<th>High starting imbalance</th>
<th>Low starting imbalance</th>
<th>High starting imbalance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \hat{\rho} )</td>
<td>99% CI</td>
<td>( \hat{\rho} )</td>
<td>99% CI</td>
</tr>
<tr>
<td>1</td>
<td>.042 (.032 , .053)</td>
<td>.096 (.082 , .108)</td>
<td>-.026 (-.037 , -.014)</td>
<td>.013 (-.001 , .027)</td>
</tr>
<tr>
<td>2</td>
<td>.029 (.020 , .037)</td>
<td>.076 (.065 , .087)</td>
<td>-.038 (-.048 , -.029)</td>
<td>-.007 (-.019 , .004)</td>
</tr>
<tr>
<td>3</td>
<td>.021 (.014 , .028)</td>
<td>.060 (.051 , .070)</td>
<td>-.056 (-.064 , -.047)</td>
<td>-.014 (-.024 , -.005)</td>
</tr>
<tr>
<td>4</td>
<td>.014 (.007 , .020)</td>
<td>.056 (.047 , .066)</td>
<td>-.056 (-.063 , -.049)</td>
<td>-.018 (-.027 , -.009)</td>
</tr>
<tr>
<td>5</td>
<td>.019 (.013 , .025)</td>
<td>.062 (.053 , .071)</td>
<td>-.051 (-.057 , -.044)</td>
<td>-.012 (-.021 , -.004)</td>
</tr>
<tr>
<td>6</td>
<td>.026 (.020 , .032)</td>
<td>.057 (.048 , .066)</td>
<td>-.044 (-.050 , -.038)</td>
<td>-.012 (-.022 , -.009)</td>
</tr>
<tr>
<td>7</td>
<td>.019 (.013 , .025)</td>
<td>.060 (.052 , .068)</td>
<td>-.047 (-.053 , -.040)</td>
<td>-.009 (-.017 , -.001)</td>
</tr>
<tr>
<td>8</td>
<td>.021 (.016 , .027)</td>
<td>.058 (.050 , .066)</td>
<td>-.058 (-.064 , -.052)</td>
<td>-.005 (-.013 , .003)</td>
</tr>
<tr>
<td>9</td>
<td>.020 (.014 , .025)</td>
<td>.060 (.060 , .077)</td>
<td>-.067 (-.073 , -.062)</td>
<td>-.018 (-.025 , -.011)</td>
</tr>
<tr>
<td>10</td>
<td>.029 (.023 , .035)</td>
<td>.077 (.070 , .085)</td>
<td>-.082 (-.088 , -.076)</td>
<td>-.038 (-.045 , -.031)</td>
</tr>
</tbody>
</table>

### Panel B: By concluding signed trade imbalance level

<table>
<thead>
<tr>
<th>Size decile</th>
<th>Low concluding imbalance</th>
<th>High concluding imbalance</th>
<th>Low concluding imbalance</th>
<th>High concluding imbalance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \hat{\rho} )</td>
<td>99% CI</td>
<td>( \hat{\rho} )</td>
<td>99% CI</td>
</tr>
<tr>
<td>1</td>
<td>.030 (.018 , .041)</td>
<td>.091 (.080 , .103)</td>
<td>-.040 (-.051 , -.029)</td>
<td>.027 (.010 , .043)</td>
</tr>
<tr>
<td>2</td>
<td>.023 (.015 , .031)</td>
<td>.063 (.053 , .073)</td>
<td>-.053 (-.062 , -.045)</td>
<td>.014 (.002 , .027)</td>
</tr>
<tr>
<td>3</td>
<td>.011 (.004 , .018)</td>
<td>.046 (.037 , .054)</td>
<td>-.061 (-.069 , -.053)</td>
<td>-.003 (-.014 , .009)</td>
</tr>
<tr>
<td>4</td>
<td>.011 (.004 , .017)</td>
<td>.040 (.032 , .048)</td>
<td>-.065 (-.072 , -.058)</td>
<td>-.003 (-.014 , .007)</td>
</tr>
<tr>
<td>5</td>
<td>.014 (.008 , .020)</td>
<td>.047 (.039 , .055)</td>
<td>-.059 (-.065 , -.052)</td>
<td>-.009 (-.020 , .001)</td>
</tr>
<tr>
<td>6</td>
<td>.017 (.011 , .023)</td>
<td>.044 (.036 , .052)</td>
<td>-.054 (-.060 , -.048)</td>
<td>-.006 (-.017 , .006)</td>
</tr>
<tr>
<td>7</td>
<td>.020 (.014 , .025)</td>
<td>.049 (.041 , .057)</td>
<td>-.055 (-.061 , -.049)</td>
<td>-.009 (-.019 , .001)</td>
</tr>
<tr>
<td>8</td>
<td>.018 (.013 , .024)</td>
<td>.050 (.043 , .058)</td>
<td>-.057 (-.063 , -.052)</td>
<td>-.009 (-.019 , .000)</td>
</tr>
<tr>
<td>9</td>
<td>.019 (.013 , .024)</td>
<td>.057 (.049 , .064)</td>
<td>-.071 (-.076 , -.066)</td>
<td>-.009 (-.018 , .001)</td>
</tr>
<tr>
<td>10</td>
<td>.006 (-.001 , .013)</td>
<td>.078 (.070 , .086)</td>
<td>-.095 (-.100 , -.090)</td>
<td>-.013 (-.023 , -.003)</td>
</tr>
</tbody>
</table>
reports sample means and 99% confidence intervals of the return correlation parameter estimates by size decile when signed trade imbalance is at its lowest level, and when it is highest. For all size deciles, regardless of whether one controls for trading activity using starting or concluding activity levels, the confidence intervals at “Low” imbalance levels do not overlap with those at “High” levels. Reversals are much larger when signed trade imbalances are low than when they are high, indicating that traders respond to limited liquidity by endogenously providing liquidity, reducing signed trade imbalances when the price of liquidity provision is higher; and reversals are smaller when the high signed trade imbalances indicate that the price charged for providing liquidity is low. In contrast, returns exhibit substantially more momentum when activity rises and concluding signed trade imbalances are high than when they are low. That is, when liquidity is inexpensively provided, traders aggressively consume it, and the resulting price reversal is small.

These findings reinforce earlier conclusions about the nature of intra-day price movements, especially price reversals. These findings are consistent with studies that attribute high order flow (signed trade imbalance) to endogenous consumption of liquidity when depth near inner quotes is unusually high (Hollifield, Miller, and Sandas (2004), Goettler, Parlour, and Rajan (2005), Bloomfield, O’Hara, and Saar. (2005), and Hollifield et al. (2006)). This observation motivates further analysis to establish how the extant liquidity affects trading activity level.

If high order flow reflects endogenous consumption of liquidity when markets are deeper near “good” prices, then trading activity should be positively correlated with the extent of aggressive trading. That is, more aggressive consumption of liquidity induces higher trading activity. However, with this higher, liquidity-consuming driven trading, the price impacts of trading fixed dollar

---

18 Moreover, if one believes that informed trading is unlikely when signed trade imbalances are low (indicating little time pressure to make trades), then stronger reversals in environments with lower signed trade imbalance are especially hard to reconcile with information-based story.
positions should be smaller in more active markets. That is, there should be a negative relationship between stock-specific market activity and price impacts of fixed dollar positions. Barardehi, Bernhardt, and Davies (2014) provide extensive evidence of this. They use the absolute return realized over each trade sequence to measure price impacts of fixed dollar positions that are relevant for institutional investors; signed trade imbalances over these trade sequences measure the aggressiveness of the trading. They find strong evidence that signed trade imbalances rise with stock-specific market activity, while price impacts fall. This suggests that institutional investors endogenously respond in the composition of their orders to state-varying levels of liquidity provision.

I next document that, at a given trading activity level, similar relations hold dynamically in ways that are consistent with endogenous consumption of liquidity. I ask: does it matter how we get there? That is, at a given level of activity, do price impacts and signed trade imbalances vary systematically according to whether the previous level of trading activity was higher or lower? In this analysis, one must control for the concluding level of trading activity to avoid conflating the distinct impacts of levels and changes in activity. Conditioning on starting levels of activity, gives rise to different concluding levels, making it difficult to discern whether the concluding level of activity drives the variations or the changes in activity that led to that level of trading activity.

In this analysis, one must control for trading activity levels more finely. As a result, to minimize noise associated with small samples, I pool across stocks by trading activity level. This, in turn, means that I must work with activity percentiles that standardize cross-stock measurements. To control for trading activity levels finely, I sort trade sequences of each stock by time duration into twenty equally-sized groups of activity on a monthly basis. Next, pooling across stocks, I decompose the sample by year, market-capitalization decile at the beginning of the year, and activity group. Thus, I construct 2400 mutually exclusive subsamples, each containing data on the same activity
group of about 110 stocks in a given year. In each subsample, I estimate simple regressions that estimate the associations of past changes in activity percentiles with current signed trade imbalances and absolute returns. I account for the cross-sectional variation in the number of durations across different stocks by weighting each stock’s observations by the inverse of its number of durations in that year (see Solon, Haider, and Wooldridge (2013)). The slope coefficients measure, on average, the amount by which signed trade imbalance or price impacts of fixed dollar positions vary as trading activity changes by 100 percentiles over the two successive trade sequences—sensitivity of signed trade imbalance/price impact to past changes in trading activity. Finally, to uncover how the massive changes in the market micro-structure after implementation of RegNMS affect patterns, I present the estimates for 2001–2006 and 2007–2012 periods separately.

**Result 9:** Controlling for the concluding level of stock-specific market activity, a greater increase in trading activity to get to that concluding level is associated with greater signed trade imbalances. The top row of Figure 2.10 presents estimates of signed trade imbalance sensitivities to activity changes, along with corresponding t-statistics, by concluding trading activity level. The two horizontal dashed lines in the plot of t-statistics stand for the 95% critical values. Save for very low levels of concluding activity, these sensitivities are generally positive and significant; they grow as concluding level of activity rises. I then focus on subsamples of increases and decreases in activity, separately. The bottom row of Figure 2.10 reveals positive sensitivities of signed trade imbalance to increases in trading activity. In contrast, sensitivities are negative given decreases in trading activity. With the exception of estimated sensitivities to decreases in activity, the rest of these findings comply with with contemporaneous relationships between stock-specific

---

19Note that an activity change of 100 percentiles is not feasible; rather, the possible change in activity hinges on the starting/concluding activity level of a trade sequence pair.

20Note that the estimates that follow increases (reductions) in activity are noisy at the least (most) active markets, since there are few observations of increases (decreases) in activity followed by low (high) concluding activity levels.
Figure 2.10: Sensitivity of current signed trade imbalance to past changes in activity by concluding stock-specific market activity level. The sample is decomposed by year, size decile, and concluding activity group—observations are pooled across different stocks. Within each category, signed trade imbalance is regressed on the past change in activity percentile, inversely weighting observations on each stock by its number of trade sequences that year. The point estimates and the associated t-statistics are reported by activity groups and sub-periods (top row). Similar estimates are obtained, isolating samples of increases and decreases in trading activity (bottom row).

market activity and signed trade imbalance. That is, consistent with endogenous consumption of liquidity, they imply that further increases in trading activity are associated with larger signed trade imbalances.

Result 10: Controlling for the concluding level of stock-specific market activity, a greater increase (decrease) in trading activity to get to that concluding level of activ-
Figure 2.11: Sensitivity of current price impact to past changes in activity by concluding stock-specific market activity level. The sample is decomposed by year, size decile, and concluding activity group—observations are pooled across different stocks. Within each category, price impact is regressed on the past change in activity percentile, inversely weighting observations on each stock by its number of trade sequences that year. The point estimates and the associated t-statistics are reported by activity groups and sub-periods (top row). Similar estimates are obtained, isolating samples of increases and decreases in trading activity (bottom row).

**Price impact is associated with greater (smaller) price impacts.** Figure 2.11 presents the distributions of point estimates and t-statistics of price impact sensitivities (in basis points) to past changes in activity at different concluding trading activity levels. The top row shows that the estimated sensitivities of price impacts to past changes in activity are overwhelmingly significantly negative. The negative sensitivities suggest that when the concluding activity level represents a greater increase or a smaller decrease from the starting activity level, then, on average, the price impacts are smaller.
This is precisely my simple secondary prediction.

The decomposition of the sample into increases and decreases in activity in the bottom panel of Figure 2.11 suggest more subtle considerations. Note that in the top panel, at low concluding levels of activity, estimates are driven primarily by reductions in activity; while at high concluding levels of activity, estimates are driven primarily by increases, reflecting that there are few observations of increases in activity at low concluding levels of activity, and few observations of decreases at high concluding levels of activity. This also means that when one decomposes the sample by increases and decreases in activity, estimates are imprecise at low levels of concluding activity that represent increases in activity, and at high levels of concluding activity that represent decreases. Subject to that important caveat, one sees that decreases in activity are uniformly associated with a negative sensitivity of price impacts to the change in activity; but that this simple predicted pattern only shows up for increases in activity at relatively high levels of concluding activity, where estimates are quite precise.\footnote{In Chapter \ref{ch2}, I present the analogues of these figures conditioning on starting activity level, and the predicted patterns all show up sharply there—but interpretation is complicated by the conflating of levels of trading activity with changes in trading activity due to using starting activity levels for conditioning.} One must be careful to avoid over-interpreting these results.

Results 9 and 10 together suggest that the endogenous responses of investors to the extant level of liquidity drive several more nuanced features of market dynamics of signed trade imbalance and price impacts, even after controlling for the implications contained in the level of concluding trading activity—one sees that how one reaches a particular level of trading activity matters for these market dynamics.\footnote{In Chapter \ref{ch2}, I replicate this exercise, controlling for starting level of activity. Although findings are robust, but one should interpret them with caution. As discussed earlier, they may mix level effects with those of changes.}

Having documented the fundamental market features that (1) rising trading activity is associated with persistence in returns, and (2) falling trading activity is associated with price reversals, and...
then providing details on the primitives driving these features, I next investigate what *intra-day* dynamics reveal about trading strategies.

### C Time-of-day and economic decisions

An important observation about short-term liquidity providers is that they have strong incentives to re-establish optimal (plausibly net-zero) positions by close every trading day. These incentives reflect desires to avoid exposure of non-zero positions to significant overnight idiosyncratic and market level risks such as after-hours announcements, or even just easier assessments of trading performances given predetermined target balances. As a result, the willingness of these short-term traders to provide liquidity reflects their prospects of re-balancing by the end of the same trading day. In particular, liquidity providers may demand a greater premium for supplying liquidity late in the trading day, when it is more difficult for them to subsequently rebalance. If so, price reversals should be more pronounced late in the trading day. Accordingly, I next investigate market dynamics in different time-of-day windows.

To control for time of day, I first identify three time windows over each trading day: early, 9:30:00AM–11:30:00AM; mid-day, 11:30:00AM–2:00:00PM; and late, 2:00:00PM–4:00:00PM. Each pair of successive trade sequences is assigned one of these time windows if its *second trade sequence begins within that window*. The well-established U-shaped pattern of trading activity over a trading day manifests itself first in the form of shorter average durations of trade sequences in trading hours near open and close of the trading day. This makes it important to establish how my findings of return persistence associated with rising trading activity and price reversals associated

---

23 This time-of-day assignment rule requires that at least one of the two successive trade sequences fully spans within a time window.
with declining activity relate to time-of-day phenomena, and that, in particular, they are not driven by the \( \cup \)-shaped trading activity pattern over a trading day.

I next show that, as one would expect, the \( \cup \)-shaped pattern of trading activity manifests itself in the form of more frequent reductions of activity in the “early” time window, and more frequent increases in activity during the “late” window. To show this, for each stock-year, I separately compute the empirical probability distributions of increases and decreases in trading activity over the three time windows—I find relative frequencies of activity increases (decreases) across the trading day. To provide the aggregate picture, I average these relative frequencies across stocks and years by time window and size decile. Table 2.7 shows that relative increases in activity are most likely near close, and least likely near open. In contrast, the likelihood of relative activity reductions fall monotonically over the course of a trading day. These patterns hold for all stock sizes.

Table 2.7: Mean relative frequency of increases/decreases in trading activity at different time-of-day windows by size decile. For each stock year, pairs of successive trade sequences are categorized by the start time of the second sequence into three time windows: early, 9:30:00AM–11:30:00AM; mid-day, 11:30:00AM–2:00:00PM; and late, 2:00:00PM–4:00:00PM. The relative frequencies of increases and decreases in activity are computed across time windows for each stock-year. Relative frequencies are averaged across stocks and years by size decile.

<table>
<thead>
<tr>
<th>Size decile</th>
<th>Trading activity increase</th>
<th>Trading activity decrease</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Early Mid-day Late</td>
<td>Early Mid-day Late</td>
</tr>
<tr>
<td>1</td>
<td>0.22 0.32 0.45</td>
<td>0.38 0.38 0.25</td>
</tr>
<tr>
<td>2</td>
<td>0.23 0.30 0.47</td>
<td>0.37 0.35 0.28</td>
</tr>
<tr>
<td>3</td>
<td>0.24 0.29 0.46</td>
<td>0.36 0.34 0.30</td>
</tr>
<tr>
<td>4</td>
<td>0.25 0.29 0.46</td>
<td>0.36 0.33 0.31</td>
</tr>
<tr>
<td>5</td>
<td>0.26 0.29 0.45</td>
<td>0.36 0.32 0.32</td>
</tr>
<tr>
<td>6</td>
<td>0.26 0.28 0.45</td>
<td>0.36 0.32 0.31</td>
</tr>
<tr>
<td>7</td>
<td>0.26 0.28 0.45</td>
<td>0.36 0.32 0.31</td>
</tr>
<tr>
<td>8</td>
<td>0.27 0.28 0.44</td>
<td>0.37 0.32 0.31</td>
</tr>
<tr>
<td>9</td>
<td>0.27 0.28 0.45</td>
<td>0.38 0.32 0.29</td>
</tr>
<tr>
<td>10</td>
<td>0.25 0.27 0.48</td>
<td>0.42 0.33 0.24</td>
</tr>
</tbody>
</table>

If the positive (negative) relation between strength of return persistence (reversals) are driven
by the time-of-day variations in stock-specific market activity levels, then I should find the weakest (strongest) return persistence (reversals) in the middle of the day when activity, on average, is lowest. This observation leads me to estimate Model (2.5) by time-of-day window, distinguishing trade sequence pairs that represent an increase in trading activity from those reflecting a decline.

Table 2.8: Sample mean and 99% confidence intervals of of \( \rho_{j,y} \) estimates that distinguish increases from decreases in activity at different time-of-day windows by stock size decile

Model (2.5) is estimated year-by-year, stock-by-stock by time-of-day window: early, 9:30:00AM–11:30:00AM; mid-day, 11:30:00AM–2:00:00PM; and late, 2:00:00PM–4:00:00PM. Model 2.5 is estimated year-by-year, stock-by-stock, for each time window, on subsamples representing decreases in activity (yielding \( \hat{\rho}_{jy} \), Panel A) and increases in activity (yielding \( \tilde{\rho}_{jy} \), Panel B). Activity changes are less than twenty percentiles. Stocks are sorted by their market-capitalizations at the beginning of each year into size deciles. Sample means and confidence intervals are computed by stock size decile.

**Panel A: Price reversals of falling activity**

<table>
<thead>
<tr>
<th>Size decile</th>
<th>Early ( \hat{\rho} ) 99% CI</th>
<th>Mid-day ( \hat{\rho} ) 99% CI</th>
<th>Late ( \hat{\rho} ) 99% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.013 (.023 , 0.004)</td>
<td>-0.035 (.054 , 0.026)</td>
<td>-0.043 (.055 , 0.032)</td>
</tr>
<tr>
<td>2</td>
<td>-0.020 (.027 , 0.013)</td>
<td>-0.046 (.053 , 0.039)</td>
<td>-0.051 (.061 , 0.041)</td>
</tr>
<tr>
<td>3</td>
<td>-0.027 (.033 , 0.020)</td>
<td>-0.055 (.061 , 0.048)</td>
<td>-0.068 (.077 , 0.060)</td>
</tr>
<tr>
<td>4</td>
<td>-0.033 (.038 , 0.027)</td>
<td>-0.056 (.062 , 0.051)</td>
<td>-0.068 (.075 , 0.060)</td>
</tr>
<tr>
<td>5</td>
<td>-0.029 (.034 , 0.024)</td>
<td>-0.051 (.057 , 0.046)</td>
<td>-0.067 (.073 , 0.060)</td>
</tr>
<tr>
<td>6</td>
<td>-0.025 (.031 , 0.020)</td>
<td>-0.049 (.055 , 0.044)</td>
<td>-0.055 (.061 , 0.048)</td>
</tr>
<tr>
<td>7</td>
<td>-0.029 (.034 , 0.024)</td>
<td>-0.050 (.055 , 0.045)</td>
<td>-0.059 (.065 , 0.053)</td>
</tr>
<tr>
<td>8</td>
<td>-0.031 (.036 , 0.027)</td>
<td>-0.056 (.061 , 0.051)</td>
<td>-0.061 (.066 , 0.055)</td>
</tr>
<tr>
<td>9</td>
<td>-0.040 (.044 , 0.036)</td>
<td>-0.069 (.074 , 0.065)</td>
<td>-0.067 (.073 , 0.061)</td>
</tr>
<tr>
<td>10</td>
<td>-0.054 (.059 , 0.050)</td>
<td>-0.088 (.092 , 0.083)</td>
<td>-0.093 (.101 , 0.086)</td>
</tr>
</tbody>
</table>

**Panel B: Return persistence of rising activity**

<table>
<thead>
<tr>
<th>Size decile</th>
<th>Early ( \hat{\rho} ) 99% CI</th>
<th>Mid-day ( \hat{\rho} ) 99% CI</th>
<th>Late ( \hat{\rho} ) 99% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.081 (.072 , .090)</td>
<td>.043 (.035 , .052)</td>
<td>.039 (.031 , .048)</td>
</tr>
<tr>
<td>2</td>
<td>.062 (.055 , .070)</td>
<td>.032 (.026 , .039)</td>
<td>.019 (.011 , .026)</td>
</tr>
<tr>
<td>3</td>
<td>.052 (.046 , .058)</td>
<td>.021 (.015 , .027)</td>
<td>.006 (.000 , .012)</td>
</tr>
<tr>
<td>4</td>
<td>.045 (.039 , .051)</td>
<td>.020 (.015 , .026)</td>
<td>-.001 (-.007 , .004)</td>
</tr>
<tr>
<td>5</td>
<td>.047 (.042 , .053)</td>
<td>.024 (.018 , .029)</td>
<td>.006 (.001 , .011)</td>
</tr>
<tr>
<td>6</td>
<td>.052 (.047 , .057)</td>
<td>.027 (.022 , .032)</td>
<td>.008 (.003 , .013)</td>
</tr>
<tr>
<td>7</td>
<td>.046 (.041 , .051)</td>
<td>.024 (.019 , .029)</td>
<td>.008 (.002 , .013)</td>
</tr>
<tr>
<td>8</td>
<td>.051 (.047 , .056)</td>
<td>.026 (.021 , .030)</td>
<td>.008 (.003 , .012)</td>
</tr>
<tr>
<td>9</td>
<td>.058 (.054 , .063)</td>
<td>.026 (.021 , .031)</td>
<td>.006 (.001 , .011)</td>
</tr>
<tr>
<td>10</td>
<td>.077 (.072 , .082)</td>
<td>.028 (.023 , .033)</td>
<td>.003 (-.002 , .008)</td>
</tr>
</tbody>
</table>

To avoid selection, I focus on increases and decreases that are less than twenty percentiles.
Figure 2.12: Empirical distributions of $\rho_{j,y}$ estimates that distinguish increases from decreases in activity at different time-of-day windows by stock size group. Model (2.5) is estimated year-by-year, stock-by-stock by time-of-day window: early, 9:30:00AM–11:30:00AM; mid-day, 11:30:00AM–2:00:00PM; and late, 2:00:00PM–4:00:00PM. Model (2.5) is estimated year-by-year, stock-by-stock, for each activity quintile, on subsamples representing increases in activity (yielding $\tilde{\rho}_{j,y}$, top row) and decreases in activity (yielding $\hat{\rho}_{j,y}$, bottom row). Activity changes are less then twenty percentiles. Stocks are sorted by their market-capitalizations at the beginning of each year into bottom 30%, middle 40%, and top 30% size groups.
Result 12: Price reversals associated with declining trading activity become progressively larger over the course of a trading day. The top row of Figure 2.12 shows that regardless of firm size, price reversals associated with decreases in activity are weakest early in a trading day, and strongest near close—the empirical distributions of $\hat{\rho}_{jy}$ shift right (toward less negative/more positive values) at earlier trading hours. Consistently, the top panel of Table 2.8 shows that when activity falls, prices revert most strongly during late trading hours, although the differences between mean reversals at “mid-day” and “late” time windows are statistically insignificant for some mid-sized stocks. These findings provide strong empirical support for the premise that short-term providers are more willing to provide liquidity early in the trading day where they have more time left over which to unwind their positions and return to their desired end-of-day positions. As a result, the “price” charged for providing liquidity rises monotonically over the day, manifesting itself in larger price reversals when activity declines.

Result 13: The return persistence of rising activity weakens over the course of a trading day. The bottom row of Figure 2.12 shows that the empirical distribution of $\hat{\rho}_{jy}$ shifts towards left (toward more negative/less positive values) in later trading hours. Consistently, the bottom panel of Table 2.8 shows that when activity rises, prices exhibit uniformly more momentum earlier in the day than later in the day, and the differences are all statistically significant regardless of firm size.

Collectively, Results 12 and 13 indicate that my findings regarding price dynamics and trading activity levels do not simply reflect time of day effects. In Table 2.9 I show that time of day properties of market dynamics strongly persist over time. Indeed, the contrast in the patterns of price dynamics that associate increases and decreases in activity with those in the likelihoods of trading activity increases/decreases at different time windows of a trading day hint at interesting
Table 2.9: Sample mean and 99% confidence intervals of of $\rho_{y,t}$ estimates that distinguish increases from decreases in activity at different time-of-day windows over time Model (2.5) is estimated year-by-year, stock-by-stock by time-of-day window: early, 9:30:00AM–11:30:00AM; mid-day, 11:30:00AM–2:00:00PM; and late, 2:00:00PM–4:00:00PM. Model (2.5) is estimated year-by-year, stock-by-stock, for each time window, on subsamples representing decreases in activity (yielding $\rho_{y,t}$, Panel A) and increases in activity (yielding $\rho_{y,t}$, Panel B). Sample means and confidence intervals are computed year-by-year.

<table>
<thead>
<tr>
<th>Year</th>
<th>Early $\hat{\rho}$ 99% CI</th>
<th>Mid-day $\hat{\rho}$ 99% CI</th>
<th>Late $\hat{\rho}$ 99% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>0.019 (0.010, 0.028)</td>
<td>-0.031 (-0.040, -0.021)</td>
<td>-0.038 (-0.052, -0.024)</td>
</tr>
<tr>
<td>2002</td>
<td>0.000 (-0.009, 0.008)</td>
<td>-0.038 (-0.047, -0.030)</td>
<td>-0.057 (-0.069, -0.045)</td>
</tr>
<tr>
<td>2003</td>
<td>-0.014 (-0.021, -0.006)</td>
<td>-0.054 (-0.062, -0.046)</td>
<td>-0.074 (-0.086, -0.063)</td>
</tr>
<tr>
<td>2004</td>
<td>-0.019 (-0.026, -0.012)</td>
<td>-0.049 (-0.056, -0.042)</td>
<td>-0.072 (-0.081, -0.063)</td>
</tr>
<tr>
<td>2005</td>
<td>-0.046 (-0.052, -0.039)</td>
<td>-0.067 (-0.073, -0.060)</td>
<td>-0.082 (-0.090, -0.075)</td>
</tr>
<tr>
<td>2006</td>
<td>-0.047 (-0.053, -0.041)</td>
<td>-0.077 (-0.083, -0.072)</td>
<td>-0.081 (-0.088, -0.075)</td>
</tr>
<tr>
<td>2007</td>
<td>-0.039 (-0.044, -0.034)</td>
<td>-0.063 (-0.068, -0.057)</td>
<td>-0.053 (-0.059, -0.047)</td>
</tr>
<tr>
<td>2008</td>
<td>-0.048 (-0.054, -0.042)</td>
<td>-0.076 (-0.082, -0.069)</td>
<td>-0.063 (-0.070, -0.056)</td>
</tr>
<tr>
<td>2009</td>
<td>-0.035 (-0.040, -0.030)</td>
<td>-0.043 (-0.049, -0.037)</td>
<td>-0.083 (-0.091, -0.076)</td>
</tr>
<tr>
<td>2010</td>
<td>-0.040 (-0.045, -0.035)</td>
<td>-0.061 (-0.067, -0.056)</td>
<td>-0.047 (-0.057, -0.036)</td>
</tr>
<tr>
<td>2011</td>
<td>-0.042 (-0.048, -0.036)</td>
<td>-0.049 (-0.055, -0.044)</td>
<td>-0.045 (-0.053, -0.037)</td>
</tr>
<tr>
<td>2012</td>
<td>-0.033 (-0.040, -0.026)</td>
<td>-0.041 (-0.047, -0.036)</td>
<td>-0.057 (-0.065, -0.049)</td>
</tr>
</tbody>
</table>

**Panel B: Return persistence of rising activity**

<table>
<thead>
<tr>
<th>Year</th>
<th>Early $\hat{\rho}$ 99% CI</th>
<th>Mid-day $\hat{\rho}$ 99% CI</th>
<th>Late $\hat{\rho}$ 99% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>0.089 (0.080, 0.098)</td>
<td>0.044 (0.034, 0.053)</td>
<td>0.020 (0.010, 0.030)</td>
</tr>
<tr>
<td>2002</td>
<td>0.076 (0.067, 0.085)</td>
<td>0.042 (0.034, 0.050)</td>
<td>0.013 (0.005, 0.021)</td>
</tr>
<tr>
<td>2003</td>
<td>0.059 (0.051, 0.067)</td>
<td>0.018 (0.011, 0.026)</td>
<td>-0.010 (-0.018, -0.002)</td>
</tr>
<tr>
<td>2004</td>
<td>0.064 (0.056, 0.071)</td>
<td>0.031 (0.024, 0.038)</td>
<td>-0.006 (-0.014, 0.001)</td>
</tr>
<tr>
<td>2005</td>
<td>0.036 (0.029, 0.044)</td>
<td>0.025 (0.019, 0.031)</td>
<td>-0.015 (-0.021, -0.009)</td>
</tr>
<tr>
<td>2006</td>
<td>0.046 (0.040, 0.053)</td>
<td>0.016 (0.010, 0.022)</td>
<td>-0.013 (-0.018, -0.007)</td>
</tr>
<tr>
<td>2007</td>
<td>0.059 (0.053, 0.066)</td>
<td>0.028 (0.022, 0.034)</td>
<td>0.028 (0.022, 0.033)</td>
</tr>
<tr>
<td>2008</td>
<td>0.047 (0.041, 0.053)</td>
<td>0.013 (0.007, 0.019)</td>
<td>0.028 (0.022, 0.034)</td>
</tr>
<tr>
<td>2009</td>
<td>0.064 (0.058, 0.070)</td>
<td>0.038 (0.032, 0.043)</td>
<td>-0.009 (-0.015, -0.004)</td>
</tr>
<tr>
<td>2010</td>
<td>0.058 (0.053, 0.064)</td>
<td>0.022 (0.017, 0.027)</td>
<td>0.045 (0.038, 0.053)</td>
</tr>
<tr>
<td>2011</td>
<td>0.051 (0.045, 0.057)</td>
<td>0.039 (0.034, 0.044)</td>
<td>0.042 (0.036, 0.047)</td>
</tr>
<tr>
<td>2012</td>
<td>0.052 (0.046, 0.059)</td>
<td>0.021 (0.016, 0.027)</td>
<td>0.010 (0.006, 0.015)</td>
</tr>
</tbody>
</table>

economic decisions lying behind them. Even though increases in trading activity are relatively less likely near open, they are associated with the strongest extent of return persistence. This strong return persistence likely reflects that institutional investors often decide on their positions overnight, and then begin establishing these large positions at open. To the extent that markets are more liquid earlier in the day because liquidity providers have more time to re-balance within the day, the incentives of institutional traders to begin establishing these positions earlier is reinforced, giving
rise to more momentum. In contrast, as market close approaches, more of the liquidity demands may have been unanticipated; and the price reversals may reflect the greater compensation liquidity providers demand later in the day when they risk closing the day with a sub-optimal position, i.e., as the opportunity cost of liquidity provision rises.

D RegNMS and the correlation structure of price dynamics

Having thoroughly documented that decreases in trading activity are associated with reversals, and increases in activity are associated with momentum, I next investigate whether there are systematic cross-sectional relationships between them. For example, do stocks that tend to have larger reversals when activity falls, also tend to have greater momentum when activity rises? To this end, I investigate the cross-sectional correlations between these two correlation parameter estimates. To deliver more observations, I estimate the correlation return parameters on a quarterly basis, rather than annual. To study the temporal evolution of momentum-reversals commonalities, I follow the quarterly cross-stock correlation between these two correlation parameters. Figure 2.13 shows that the commonality between the two types of price dynamics secularly weakens over time. In earlier years, the extent of price reversals associated with decreases in trading activity is negatively related to the extent of momentum associated with increases in trading activity. This cross-stock correlation almost vanishes in the last four years of the sample period, indicating that the common forces that drive these two types of price dynamics have deteriorated in recent years.

The dramatic reduction in the commonality of return persistence and price reversals overlaps with the advent of high frequency traders (HFTs) whose market share rose sharply after implementation of RegNMS in early 2007. The composition of market participants was radically altered to the extent that HFTs crowded out the specialists who provided traditional market-making services.
Figure 2.13: The temporal evolution of the commonality in the return correlations associated with increases and decreases in stock-specific market activity. Every quarter, I estimate the correlation structure of returns realized over two successive trade sequences for the sample of increases in activity and for the sample of reductions in activity, stock-by-stock. I then estimate the cross-stock correlations of these two correlation parameters quarterly.

Relative to traditional market makers, HFTs effectively have more limited capital and generally operate in markets of many stocks, which shortens the horizons over which they seek to re-establish optimal positions.

The fundamentally different rebalancing considerations of HFTs and traditional liquidity providers lead me to investigate the dynamics of aggressive trading. Liquidity consumption often involves relatively high intensity of marketable orders. For instance, high liquidity demand on the buy side is associated with a high intensity of buy-side marketable orders, and a high proportion of buyer-initiated trades. When liquidity demand falls, trading activity falls and prices revert.

I next explore what the evolving correlation structure of signed marketable order intensities reveals about the impact of HFTs. I measure the intensity of signed marketable orders using the proportion of buyer-initiated trade in a sequence, i.e., using the value of buyer-initiated trade in a trade sequence divided by the total dollar volume of that sequence. For the subsample of activity reductions that are less than twenty activity percentiles, I distinguish pairs of trade sequences with
negative initial (past) returns from those with positive initial returns. Increases in the proportion of buyer-initiated trade reflect reductions in the intensity of marketable sell orders and increases in the intensity of marketable buy orders. Each year, I compute the mean change in the proportion of buyer-initiated trade in the subsample of activity reductions by the sign of initial price impact, stock-by-stock. I then average these means across stocks by year and stock size group (bottom 30%, middle 40%, and top 30% of market-capitalization).

Figure 2.14 shows how outcomes changed radically over the 2001–2012 period. In earlier years, regardless of firm size, reductions in trading activity after negative (positive) returns are, on average, associated with higher (lower) proportions of buyer-initiated trade. That is, as trading activity declines following a liquidity demand shock, the intensity of marketable orders moves in the opposite direction of the initial price impact; the proportion of buyer-initiated trade rises after negative price impacts, but it falls after positive price impacts. These patterns secularly weaken over time to the extent that they revert in the last few years of the sample. Strikingly, in the later years, the intensity of marketable orders moves in the same direction of past price impacts as trading activity declines: when activity falls, past positive returns are followed by increases in buyer-initiated trade, and negative returns are followed by decreases in buyer-initiated trade.

These findings provide stark evidence of the very different rebalancing horizons of traditional market-makers and HFTs. Traditional market makers are less concerned about rebalancing quickly, so reductions in stock-specific market activity after episodes of high liquidity demand are associated with reductions in the intensity of marketable orders, due to the fading of the liquidity demand

\footnote{Restricting the activity reductions to less that twenty percentiles, limits my analysis to about 70% of observations of activity reduction. Relaxing this restriction gives rise to selection biases, since it weights results further in favor of radical reduction in activity. However, qualitative identical findings result. Of note, I’ve already shown in Table 2.8 that reversals of falling activity and return persistence of rising activity estimated based on the same restriction strongly persist over the sample period.}
Negative past returns

Positive past returns

Figure 2.14: The temporal evolution of mean change in proportion of buyer-initiated trading when stock-specific market activity falls by signed past returns. Samples of successive trade sequence pairs that represent reductions in activity are decomposed by signing the return realized over the first trade sequence of each pair (past return). To avoid selection bias, the reduction in activity must be smaller than twenty activity percentiles. For each stock, the mean change in the proportion of buyer-initiated trades, associated with a decline in activity after signed returns, is computed annually. Averages of those means are taken across stocks of a size group (bottom 30%, middle 40%, and top 30% of market-capitalizations at the beginning pf each year) year-by-year.

shock. Thus, after positive price impacts, an activity decline is associated with a relative reduction in marketable buy orders; and after negative price impacts, reductions in activity lead to decreases in the intensity of marketable sell orders. In contrast, after HFTs provide liquidity, their short rebalancing horizons leads them to re-establish their optimal positions quickly. As a result, when a liquidity demand fades, they begin to trade aggressively on the same side to which they just
provided liquidity. Consequently, in recent years the goals of HFTs to unwind their liquidity-providing trades to reestablish their optimal positions raise the intensity of marketable orders in the direction of previous liquidity demand shock.

I next investigate this correlation structure of signed marketable orders when activity declines varies with (concluding) trading activity in early and late years. Focusing on activity declines of less than twenty percentiles, for each stock, I calculate the mean change in the proportion of buyer-initiated trade, conditioning on the sign of the past return, by concluding activity decile. I then average those means across stocks within a firm size group over years in early (2001–2006) and later (2007-2012) periods to distinguish the temporal changes.

The top row of Figure 2.15 shows that in the earlier period 2001–2006, aggressive trading becomes less extreme in the direction of past returns (price impacts) as activity falls, especially when concluding trading activity levels are higher. The average change in the proportion of buyer-initiated trade more than doubles as concluding activity goes from its lowest level to its highest. This feature holds for all stock sizes. I’ve already shown that prices revert under such circumstances, and that these reversals are stronger in less active markets. This phenomena indicates that liquidity providers receive reduced liquidity premia in more active markets where reversals are weakest. Indeed, I find an inverse relationship between the absolute average change in the proportion of buyer-initiated trades and the size of price reversals, which is consistent with an “equilibrium” order flow pricing schedule.

The bottom row of Figure 2.15 reveals that patterns are strikingly different in more recent years. In the 2007–2012 period, the exact opposite pattern in the extent of aggressive trading emerges: aggressive trading is more extreme as trading activity declines. For mid-sized and large firms,
Figure 2.15: The relation between concluding stock-specific market activity level and the change in the proportion of buyer-initiated trades as trading activity declines. Samples of successive trade sequence pairs that represent reductions in activity are decomposed by signing the return realized over the first trade sequence of each pair (past return). To avoid selection bias, the reduction in activity must be smaller than twenty activity percentiles. For each stock, the mean change in the proportion of buyer-initiated trades, associated with a decline in activity after signed returns, is computed by activity decile, annually. Averages of those means are taken across stocks of a size group (bottom 30%, middle 40%, and top 30% of market-capitalizations at the beginning of each year) and over time within the periods 2001–2006 and 2007–2012 by activity decile. Solid and dashed curves signify positive (negative) past returns, respectively.
where HFTs are most active, the changes are sharply magnified when markets are more active. The previously-found inverse relationship between the size of price reversals and the concluding activity level, coupled with greater changes in the proportion of buyer-initiated trades in more active markets, indicate that order flow is more expensive in less active markets. These findings provide additional insight into why the short-term price reversals primarily reflect rewards to liquidity provision rather than reflecting information driven trading.

I next search for evidence that the desire of HFTs to clear their positions by close gives rise to time-of-day effects in the dynamics of the intensity of signed marketable orders. That is, late in the trading day, HFTs have less time to re-establish optimal positions so the impact on the proportion of buyer-initiated trades after reductions in trading activity should be magnified near close. To test this, I distinguish the last two hours of a trading day (2:00PM–4:00PM) from the rest of the day. I then replicate the analysis above controlling for the time window of the concluding trade sequence for pairs that represent reductions in stock-specific market activity. I perform the exercise for the 2007–2012 period only, where dynamics seem to reflect the behavior of HFTs. Figure 2.16 provides evidence of such time-of-day effects for mid-sized and large stocks where HFTs are most active. This result reinforces our understanding of how the trading horizons of HFTs have so radically changed the correlation structure of the intensity of signed marketable orders. This also highlights the importance of the time of day for the rebalancing horizons of liquidity providers.

V Conclusion

Since 2001, post decimalization, daily trading volumes in U.S. stock markets have exploded, and the nature of market participants and their trading strategies have radically changed. The role of
specialized market makers in providing liquidity has been replaced by HFTs who have exploited advances in computing power and programming to implement advanced algorithmic trading strategies, and trade speeds are measured in nanoseconds. Institutional traders, facing tiny bid-ask spreads, minimal depth and fragmented markets, now establish and unwind their large positions by finely dividing their orders over time and across trading venues.

The contribution of my paper is to build a framework in which one can analyze core dynamic properties of today’s ultra high frequency stock trading markets. Before analyzing intra-day market dynamics, one must first confront the fact that individual trades are now temporally-dependent; they are likely to represent small components of a bigger position that a trader intends to establish intertemporally, or parts of a larger order that have been shredded against the order book. This temporal dependence complicates measurements over very short horizons (e.g., tick-by-tick); while
longer horizons are likely to mix different market environments, e.g., different stock-specific market activity levels. I develop new stock-level measures of trading activity by identifying successive trade sequences each worth a fixed dollar volume. Fixing the underlying economic variable—dollar volume—successive trade sequences deliver comparable measurements of fundamentals. I measure three fundamentals of each trade sequence: (a) its time duration, which measures the level of stock-specific market activity, (b) its return, which measures the price impact of fixed-dollar positions; and (c) signed trade imbalance, which measures the intensity of consumption of liquidity.

As a first step, I show that trading activity matters for returns: on average, the return of a trade sequence in more active markets (shorter time durations) is positive, but the average return in less active markets is negative—less active markets seem to reflect reduced “market interest”. I next investigate whether it matters how the market gets to a given level of trading activity. To do this, I estimate the correlation structure of returns realized over two successive trade sequences at the stock level, every year, to investigate the correlation structure of returns. I distinguish between subsamples of increases and increases in activity. I make statistical inferences using the empirical distributions of stock-specific correlation parameter estimates—ensuring that my findings are robust to arbitrary heterogeneity in stock-specific error structures.

I establish that relative increases in stock-specific market activity are associated with return persistence, but relative activity reductions are associated with price reversals. In particular, for a given level of concluding trading activity (i.e., for the second of successive trade sequences), if its activity level exceeds the previous level of activity, then its return is positively correlated with the previous return; but if the concluding activity level is less than the previous level, its return is negatively correlated with the previous return. Further, the extent of reversal or persistence rises with the size of the change in trading activity. Ceteris paribus, price reversals are higher in
less active markets, but return persistence is higher in more active markets. These reversal and persistence findings also vary systematically with the time of day—even though activity tends to be highest at open and lowest mid-day, I find that the extent of price reversals rises secularly over the trading day, while the extent of persistence in returns falls. In short, how the market gets to a given level of trading activity matters for the dynamics of returns.

These findings are remarkably robust. In particular, they do not vary systematically with firm size—the results do not reflect small/illiquid stock anomalies. They also do not vary systematically over time—results do not reflect impacts of the financial crisis, much less the broader yearly performance of the market. My findings also cannot be attributed to outliers, or to biases caused by mis-specified error-term structures. In sum, they beg for sound explanations.

I consider two alternative explanations for the core finding of price reversals when trading activity falls. It could be that episodes of high activity reflect intense private-information driven order flow. The magnitude of information content may be unknown to the market, so prices may revert when activity declines and “the market” concludes that the information content was less than it could have been. Alternatively, a reduction in stock-specific market activity could reflect a decline in demand for liquidity. If so, higher previous liquidity demand may have driven prices further from fundamentals as liquidity providers are compensated for their services, and prices move toward their fundamental values when activity declines. Such liquidity-based price impacts are at the heart of the classical Chriss-Almgren (2001) model that underlie the trading strategies of many HFTs.

So, too, the return persistence of rising activity could reflect investors aggressively trading when they acquire substantial private information, driving up trading activity and persistently
moving prices in one direction. Alternatively, the increased trading activity could reflect continued endogenous consumption of liquidity when the market is unusually deep on the other side.

My analysis indicates that both phenomena are liquidity driven, not information driven. In particular, price reversals of falling activity are strongest in least active trading markets, even when I condition on the starting level of trading activity. However, information-based explanations predict that informed trade is more likely when activity is higher. In contrast, a liquidity-based explanation presents itself—the starting activity level is lower because less liquidity is available, and with less competition to provide liquidity, liquidity demand drives prices further away from fundamentals. Reinforcing these conclusions, I establish that the strength of price reversals is stronger when signed trade is more balanced, not less. This is inconsistent with information-based explanations—informed agents will pay to consume liquidity to make purchases before their information leaks out; but consistent with a scenario in which an institutional trader responds to limited liquidity provision by altering the composition of his orders toward taking less liquidity and providing more. I also establish that return persistence of rising activity is strongest when signed trade imbalance is highest, and that signed trade imbalances rise with stock-specific market activity. These patterns reflect that when liquidity is unusually high, traders will aggressively consume it.

I find that the price reversals associated with declining activity rise secularly over the trading day. This finding sheds light on the trading horizons of liquidity providers. It suggests that liquidity providers, HFTs, in particular, have targets for closing positions, and that it is more difficult for them to supply liquidity later in the day and meet those targets. As a result, the prices charged for providing liquidity rise later in the trading day. I argue that the return persistence earlier in the day reflects institutional traders who determine overnight their desired positions, and then begin trading to establish those positions early in the trading day.
I then document a puzzling feature in the temporal evolution of the cross-stock correlation structure of price reversals and return persistence. I investigate whether stocks associated with stronger price reversals when trading activity falls, on average, feature stronger or weaker return persistence of rising activity—i.e., is the correlation coefficient negative or positive? I find that in early years of the sample, correlations are positive and significant: stocks with stronger reversals feature weaker return persistence. By the second quarter of 2009, this correlation vanishes, possibly reflecting the effects of RegNMS in 2007 and other institutional changes. The paradox is that although the correlation structure has changed systematically, the levels of return persistence associated with increases in activity, and the levels of return reversals associated with decreases have not.

Motivated by these structural changes I conclude my analysis by exploring changes in the dynamics of aggressive trading in the subsample of trading activity reductions. In particular, I identify fundamentally different dynamics in 2001–2006 and 2007–2012 periods. Conditioning on concluding stock-specific market activity levels, which filters level effects, I document the change in the proportion of buyer-oriented trades for positive and negative past returns. In the 2001–2006 period, aggressive trading moves in the opposite directions of past price impacts: when trading activity falls after positive returns, the proportion of buyer-initiated trade falls; and after negative returns, the proportion of buyer-initiated trades rises. The exact opposite holds true in the later years of my sample: when trading activity falls after positive returns, the proportion of buyer-initiated trade rises; and after negative returns, the proportion of buyer-initiated trades falls. This change is consistent with liquidity providers rebalancing positions more and more quickly after providing liquidity in more recent years, which may be attributed to the shorter trading horizons of HFTs who provide increasing shares of liquidity in more recent years.
Chapter 3

ACD Models in Modern Markets

I Introduction

The seminal work of Engle and Russell (1998) introduced a new technique to analyze high-frequency time series data. Motivated by predictions of classical microstructure models, e.g., Easley and O’Hara (1992), about the information content of the time between successive transactions, Engle and Russell viewed the transactions irregularly-spaced time series. As such, they developed Autoregressive Conditional Duration (henceforth ACD) models to estimate and predict the time durations between successive transactions (henceforth inter-transaction durations). ACD models describe inter-transaction durations as a product of a conditional mean with GARCH structure and an error term (a.k.a standardized duration). Standardized durations are assumed to be i.i.d and to come from a distribution with a positive support.

ACD models of a inter-transaction durations have received significant attention and have been used for applications in market microstructure research (e.g. Dufour and Engle (2000)). However, revolutionary changes in financial markets challenge the validity of how we interpret estimations of these models in today’s markets. Trading strategies and outcomes are very different than what classical models of market microstructure were designed to describe. These models treat an individual transaction as the outcome of an economic decision, while this interpretation about a transaction
does not hold in modern markets.

Massive regulatory and technological changes have transformed financial markets in many ways: (a) markets are fragmented offering arbitrage possibilities at ultra high frequencies across many trading venues—two successive transaction may simply reflect arbitrage; (b) anyone can provide liquidity in today’s markets by posting orders on the limit order book, not only a designated market maker, and the learning process is different than what classical models usually assume; (c) with tiny bid-ask spreads and modest depth near good prices, institutional investors have moved from submitting a few large orders to finely split orders over time and across trading venues to minimize the price impacts associated with large intended trade sizes (Hendershott et al. (2011)); and (d) algorithmic trading gives rise to mechanical transaction clusters since algorithms follow instructions given round second start times and time grids (see Hasbrouck and Saar (2013) and Muravyev and Picard (2014)).

These features largely question the information content of inter-transaction durations that originally motivated ACD models. Today, individual transactions and the time durations between them may not reflect information about the economic driving forces of trading; they are largely determined by executions algorithms and high frequency traders’ behaviors.

Yet another set of challenges are posed by the extent of data accuracy and complications. With ultra high frequency trading and discrete time measurements, successive transactions that virtually take place in different points in time may be recorded as “simultaneous” events. Braccini (2014) reports that post 2003 more than 25% of transactions coincided with some other transaction(s), even if time is measured in milliseconds. Utilizing TAQ data, he considers a mixture model with a point mass at zero to account for simultaneous events. Two facts cast doubt on validity of his
approach: (a) inter-transaction durations and the corresponding parameter estimates may not be economically interpretable; and (b) TAQ data set excludes transactions with sizes less than 100 shares while, as reported in O’Hara et al. (2012), these trades make up, on average, 7% of dollar volume and 35% of price discovery.

These concerns lead me to employ in my ACD analysis a novel measure of trading activity developed by Barardehi, Bernhardt, and Davies (2014). This stock-specific metric presents an aggregation approach that fixes a dollar value $V$ and measures the time it takes for a sequence of successive transactions with aggregate value $V$ (a transaction sequence) to trade. Starting at a point in time, e.g., the beginning of the year, mutually exclusive time durations of this sort are computed sequentially. A transaction event occurs when the aggregate value of successive individual transactions reaches $V$—higher trading activity implies shorter time durations between successive events. By choosing $V$ large but not too large, this technique deals with the temporal dependence and systematic clusters of individual transactions preserving the information content of both inter-transaction durations and transaction sizes. It also controls for dollar volume by measuring how fast a fixed dollar-position is traded. Finally, it largely eliminates issues associated with simultaneous events without either excluding or directly handling “coinciding” transactions.

I use this measure of trading activity as an input to ACD models, and estimate stock-year-specific ACDs using a large sample of stocks. This sample includes common NYSE-listed stocks in 2007–2012 that maintain a minimum daily closing price of at least $1$ and deliver at least 1000 transaction sequences a year given my stock-specific choices of $V$. I analyze the empirical distributions of stock-year-specific ACD fits.

I first document that the GARCH structures of ACD fits are relatively invariant across stocks,
especially in earlier years of the sample. Empirical distributions of point estimates are bimodal for some parameters, revealing a two-type GARCH structure. I find that, for all stock-year sets of observations, the shape parameter of the Weibull distributions of standardized durations exceed 1; that is, the hazard rate monotonically increases in the length of time durations, and the likelihood of a transaction event monotonically rises as time elapses after the previous event. Next, I document a monotonically increasing shape parameter size in stock size while the annual number of transaction sequences takes a ∩-shaped pattern in stock size. Since distributions of time durations become less skewed as stock size increases, the pattern of shape parameter size in stock size highlights how the shape of durations’ distribution matter for the shape parameter. More consistent trading activity is associated with disproportionately greater increases in hazard rate as standardized durations grow.

I then use some procedures to verify that the realized standardized durations (error terms) are close to being i.i.d. To do this, I first compare the autoregressive structures of trade durations with those of standardized durations; I show that the conditional mean removes a significant portion of the autoregressive structure. I then contrast the predictability of trade durations with that of standardized durations using levels of lagged trading activity; I verify that durations, unlike standardized durations, are highly predictable. Finally, I compare the empirical hazard rates with those predicted by the model; I find that implied hazard rates tolerably deviate from the predictions of the postulated model.

As for future work, two paths seem reasonable:

1. **Taking the classical ACD structure as the correct specification:** I can employ the stock-specific model estimates to perform trading activity forecasts and evaluate forecast accuracies. I can move one step further and relate this work to that of Barardehi (2014).
He finds systematic but strikingly different price dynamics when trading activity rises versus when it falls. To this end, I can examine the price movements in times when ACD estimates correctly forecast the general direction of trading activity.

2. **Improve ACD specification to improve forecast:** Another path is to alter the ACD specification in order to improve accuracy of ACD forecast performances. One can give more flexibility to the conditional mean process of the ACD by adding in a random component in the sense of Bauwens and Veredas (2004). However, an interesting improvement is to relax their distributional assumption that imposes normality of the random component. To do this, one might be able to use the non-linear improvement that Greenshtein et al. (2014) develop for Kalman Filter using Empirical Bayes methods.

Next, I explain, in detail, how transaction sequences are constructed; section [III] describes the data and provides some summary statistics; section [IV] describes the statistical framework; section [V] presents the results; and section [VI] details future work.

## II Durations

Instead of the time durations between successive transactions, I use a novel notion of duration that fits the features of today’s equity markets. For a given stock $j$ in month $m$, I compute successive execution times of transaction sequences with a fixed dollar position $V_{jm}$. This fixed position is a function of the previous month’s market capitalization; in this draft, $V_{jm}$ equals the sum of $\$80k$ and 0.025% of previous month’s market capitalization. This notion offers various virtues:
• **It deals with the temporal dependence of individual transactions:** instead of directly modeling the temporal dependence of individual transactions, I aggregate across successive transactions, setting the aggregation horizon inversely related to trading activity. As a result, unlike the inter-transaction durations that are highly affected by the behaviors of execution algorithms (order-splitting, etc.), time durations of these transaction sequences are more likely to represent the underlying fundamental trading decisions.

• **It delivers comparable random objects:** information content of successive transaction sequences are comparable since they all correspond to dollar volume \( V_{jm} \)—as opposed to the inter-transaction durations that possess different information contents since successive individual transactions could differ in size. As a result, one obtains parameter estimates that are easier to interpret. More specifically, fitting inter-transaction durations into an ACD structure largely ignores the transaction sizes, and solely focuses on durations; in contrast, my approach indirectly incorporates the size dimension via aggregation across successive transactions.

• **It resolves the “simultaneous events” problem:** As noted by Braccini (2014), due to discreteness of measurement, some transactions appear to happen at the *exact* same time. This feature adds another challenging dimension to modeling the inter-transaction durations, and one must directly deal with it, especially in recent years. Our approach accounts for such events though aggregation; in fact, for reasonable choices of \( V_{jm} \) simultaneous events (time durations of 0) may not exist.

Figure 3.1 shows how successive transaction sequences with a fixed cumulative dollar volume \( V \) are calculated for a typical stock. An *execution event* occurs when the cumulative dollar volume reaches the cutoff \( V \). I denote the series of durations for stock \( j \) by \( x_{jt} \) where \( t = \{1, \ldots, T_j\} \) indexes
the order of transaction events (durations). Accordingly, I denote the series of diurnally-adjusted
series of durations by \( \tilde{x}_{jt} \).

Figure 3.1: Calculations of successive transaction sequences of a given stock.

A: Dollar volume

B: Cumulative dollar volume and trade sequences

Although motivated and constructed similarly, this measure of trading activity is fundamentally
different from those in Gouriéroux et al. (1999): (a) they construct overlapping trade durations by
re-initiating a trade sequence for every transaction, while my approach delivers mutually exclusive
time durations; (b) they use current prices to compute transactions values which, unconditionally,
biases down the time durations when prices are rising, and up when prices are falling—I address
this concern by using the the previous day’s closing price to compute transaction values; and (c) I
set $V$ a function of the stock’s market-capitalization at the end of the previous month to facilitate interpretation of cross-sectional and temporal variations.

III Data

I use NYSE TAQ (Consolidated Trades) trade-by-trade data to obtain the trade sizes and time stamps of intra-day transactions. I extract the daily closing prices from daily CRSP and use the previous day’s closing price to compute current trade-by-trade dollar volumes; thus, I make the identification of transaction sequences independent of simultaneous price movements. I obtain monthly market capitalizations from monthly CRSP and use them to compute the stock-specific cutoffs $V_{jm}$. The sample contains NYSE-listed common stocks whose daily closing price (or mid-quote at close) exceeds $\$1$ throughout the year. Sample period spans Jan 2007–Dec 2012. Finally, after identifying the transaction sequences, stocks with fewer than 1,000 trade sequences are excluded. I compute the trade sequences that span over two successive trading days, but I flag them in the analysis to avoid any systematic bias.

To provide an overview of transaction sequence durations, I compute the median duration of each stock’s transaction sequences year-by-year. I then explore the distributional properties of these medians across stocks and years. Table 3.1 shows that the cross stock average of median duration takes a $\cup$-shaped pattern in stock size. The 25th and the 75th percentiles of these medians

---

1Easley et al. (2012) propose similar measures by dividing each trading day’s transaction history into a fixed number of volume bundles. Their approach do not control for volume (dollar volume) since trading volume massively varies across trading days and over time, and volume bundle sizes massively vary across trading days accordingly.

2These transaction sequences are different from the rest of the observations for various reasons: they commence near close and end after open the next day—they mix very different market conditions; the trading volumes of transactions in such sequences are computed based on two different preceding closing prices—transactions belong to two different trading days and hence two different preceding closing prices are applied; finally, once in a month, such overnight trade sequences experience a cutoff ($V_{jm}$) adjustment once I go from month $m − 1$ to month $m$. I control for these differences using day and month dummies.
Table 3.1: **Overview of transaction sequence duration:** The median time duration of each stock’s transaction sequences is computed annually. Each year, stocks are grouped into size deciles based on market-capitalizations at the end of the previous year; size decile 1 contains stocks with smallest market-caps and size decile 10 contains those with largest. The second column of the table reports the cross-stock averages of stock-specific median durations, the third column reports the inter-quartile range of stock-specific median durations, and the last column reports the cross-stock mean annual number of transaction sequences.

<table>
<thead>
<tr>
<th>Size decile</th>
<th>Average (Inter-quartile range)</th>
<th># of sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>21.7 (13.7, 29.3)</td>
<td>3323</td>
</tr>
<tr>
<td>2</td>
<td>16.8 (10.0, 22.2)</td>
<td>5108</td>
</tr>
<tr>
<td>3</td>
<td>13.7 (7.4, 18.4)</td>
<td>7052</td>
</tr>
<tr>
<td>4</td>
<td>12.1 (6.7, 15.4)</td>
<td>8059</td>
</tr>
<tr>
<td>5</td>
<td>10.4 (5.7, 13.3)</td>
<td>9595</td>
</tr>
<tr>
<td>6</td>
<td>10.3 (5.4, 13.2)</td>
<td>10047</td>
</tr>
<tr>
<td>7</td>
<td>9.6 (5.6, 12.1)</td>
<td>10324</td>
</tr>
<tr>
<td>8</td>
<td>9.1 (5.4, 11.2)</td>
<td>11403</td>
</tr>
<tr>
<td>9</td>
<td>9.5 (6.1, 12.1)</td>
<td>10033</td>
</tr>
<tr>
<td>10</td>
<td>13.0 (8.4, 16.0)</td>
<td>7909</td>
</tr>
</tbody>
</table>

display similar patterns indicating a fundamental association between stock size and durations. Consistently, the average stock-year-specific number of transaction sequences follows a \(\cap\)-shaped pattern in stock size reflecting that the cutoff \(V_{jt}\) is reached faster for medium-large stocks.

### IV Model

Let \(\psi_{jt} = E(\tilde{X}_{jt}|\tilde{x}_{j1}, \ldots, \tilde{x}_{jt-1})\) be the conditional mean of diurnally-adjusted durations at time \(t\).

Then an Autoregressive Conditional Duration model of order \((r, s)\) (denoted ACD\((r, s)\)), proposed by Engle and Russell (1998), is specified as Model 3.1

\[
\begin{align*}
\tilde{x}_{jt} &= \psi_{jt}\epsilon_{jt}, \quad \epsilon_{jt} \overset{i.i.d.}{\sim} f(\epsilon, \phi) \\
\psi_{jt} &= \omega_j^t + \sum_{i=1}^r \alpha_i^j \tilde{x}_{jt-i} + \sum_{i=1}^s \beta_i^j \psi_{jt-i}
\end{align*}
\]

\(^3\)qualitatively identical results obtain if I focus on stock-specific mean durations instead of median durations.
All parameters of the model, including φ, are estimated using quasi-maximum likelihood which requires assertion of the \( f(\epsilon, \phi) \) distribution. To assure positive expected durations, \( \epsilon \) must come from a positive-valued distributions. I focus on Weibull distribution, while exponential and gamma distributions are other standard alternatives. WACD\((r, s)\) denotes an ACD\((r, s)\) with Weibull-distributed \( \epsilon \). The standardized Weibull density function takes the following form, with \( \gamma \) referred to as the shape parameter of the gamma distribution.

\[
f(\epsilon|\gamma) = \begin{cases} \alpha \left[ \Gamma(1 + \frac{1}{\gamma}) \right]^\gamma \epsilon^{\gamma-1} \exp \left\{ - \left[ \left(1 + \frac{1}{\gamma}\right) \epsilon \right]^\gamma \right\} & \text{if } \epsilon \geq 0 \\ 0 & \text{otherwise.} \end{cases}
\]

(3.2)

While the hazard function is assumed to be constant in \( x \) with exponential distribution, Weibull distribution allows the hazard function to monotonically increase (decrease) in \( x \) when \( \gamma \) is greater (less than) unity.

As followed in the literature, I do not model fit raw time durations of transaction sequences using ACD models; there are nontrivial systematic effects that I need to filter out in the outset. The diurnal patterns are known to exist in intra-day stock trade data; trade intensity is higher near open and close, compared to middle of the trading day. This is translated to shorter durations near open and close. In addition, one should distinguish the overnight transaction sequences and those that span different months from other durations. I begin by filtering diurnal patterns using simple splines, as suggested by Tsay (2005), and overnight effect using dummy variables.

Let \( \tau_{jt} \) represent the time of day at which duration \( t \) of stock \( j \) ends (event time) in terms of the number of seconds from open. For example, if duration \( x_{jt} \) ends at 9:33:25AM, then \( \tau_{jt} = 205 \), given market opens at 9:30:00AM, and the duration takes 3 minutes and 25 seconds. Following Engle and Russell (1998) I define 1:00:00PM as the middle of the trading day: it is 3.5 hours from
open. Accordingly, 1:00:00PM corresponds to 12600 seconds from open, and close (4:15:00PM) corresponds to 24300 from open. Next, I define the deterministic functions to capture the \( \cap \)-shaped patterns in durations over the course of a trading day.

\[
O(\tau_{jt}) = \begin{cases} 
\tau_{jt} & \text{if } \tau_{jt} < 12600 \\
0 & \text{otherwise}
\end{cases}
\]

\[
C(\tau_{jt}) = \begin{cases} 
24300 - \tau_{jt} & \text{if } \tau_{jt} \geq 12600 \\
0 & \text{otherwise.}
\end{cases}
\] (3.3)

To capture the over night fixed effects, I define a dummy variable \( ON_{jt} \) that takes 1 if duration \( x_{jt} \) spans two trading days, and zero otherwise. Similarly, the indicator variable \( M_{jt} \) that captures the monthly cutoff adjustment takes 1 if duration \( x_{jt} \) spans two months and zero otherwise. I estimate the following OLS regression, per stock-year set of observations

\[
\ln(x_{jt}) = \lambda_0 + \lambda_1 O(\tau_{jt}) + \lambda_2 C(\tau_{jt}) + \lambda_3 ON_{jt} + \lambda_4 M_{jt,t} + u_{jt}
\] (3.4)

\( \hat{x}_{jt} = \exp(\hat{u}_{jt}) \) represent trade durations after controlling for the “time-of-day” and overnight anomaly effects.

V Results

I commence my estimation analysis by examining the heterogeneity of the GARCH structure fits across stocks and over time. I fit diurnally-adjusted durations into WACD(2,2) models on a stock-by-stock basis every year. Thus, each year, I estimate the set of parameters \( \{\alpha_1, \alpha_2, \beta_1, \beta_2, \omega, \gamma\} \).

Thereby, I provide a rich picture of trading activity dynamics of stocks, using a reasonably large

\[\text{footnote}{4} \text{I find qualitatively similar results for WACD}(1,1) \text{ as well as when I assume exponentially-distributed } \epsilon \text{ (EACD}(1,1) \text{ and EACD}(2,2)).} \]

125
sample and allowing for stock-year level parameter heterogeneity.

Table 3.2: Number of insignificant estimates: WACD(2,2)s are fitted on a stock-year basis. Reported are the number of cases where a particular parameter in an ACD fit is insignificant at 5% level. Durations measure the time it takes for a sequence of transactions worth $80,000 plus 0.025% of market capitalization. Durations are adjusted for diurnal patterns and overnight regime changes according to 3.4.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>α1</th>
<th>α2</th>
<th>β1</th>
<th>β2</th>
<th>ω</th>
<th>γ</th>
</tr>
</thead>
<tbody>
<tr>
<td># of insignificant</td>
<td>4</td>
<td>134</td>
<td>131</td>
<td>2072</td>
<td>569</td>
<td>0</td>
</tr>
</tbody>
</table>

I fit 7032 stock-year-specific WACD(2,2)s to estimate \( \{ \alpha_j^1, \alpha_j^2, \beta_j^1, \beta_j^2, \omega_j, \gamma_j \} \); in 4655 cases, parameters are all significant at 5% level. In other cases, a subset of parameters are insignificant; however, a stock-year estimate without a complete set of significant parameters most likely includes an insignificant \( \beta_2 \). Table 3.2 shows that a notable bulk of \( \beta_2 \) estimates are insignificant at 5% level while other parameters are significant in most stock-year estimates. A closer look at the distribution of insignificant \( \beta_2 \)s reveals that the corresponding point estimates are very close to zero; the 10th percentile, median, and 99th percentile of these insignificant estimates are \(-0.088\), \(-0.006\), and \(0.103\), respectively. This indicates a strong association between the weight of \( \beta_2 \) in explaining trading activity dynamics and its statistical significance, confirming merits of an ACD(2,2) specification for the conditional mean.

The GARCH structure of the conditional mean is safely interpretable only if it constitutes a stationary process. To inspect the stationarity of the processes, one should first rewrite Model (3.1) as an ARMA\((l, l)\) process. Let \( \tilde{x}_{jt} = \psi_{jt} + \eta_{jt} \) where \( \eta \) is a mean-zero process. Then Model (3.1) can be written as the following ARMA process.

\[
\tilde{x}_{jt} + \sum_{i=1}^{l} (\alpha^j + \beta^j) x_{jt-i} = \omega_j + \eta_{jt} + \sum_{i=1}^{l} \beta^j \eta_{jt-i} \tag{3.5}
\]

The conditional mean process in Model (3.1) is stationary only if both AR\((l)\) and MA\((l)\) processes
Figure 3.2: **Empirical distributions of parameter estimates**: WACD(2,2)s are fitted on a stock-year basis. The histograms reflect the empirical distributions of stock-year-specific estimates of parameters \{α_1, α_2, β_1, β_2, ω, γ\}. Durations measure the time it takes for a sequence of transactions worth $80,000 plus 0.025% of market capitalization. Durations are adjusted for diurnal patterns and overnight regime changes according to 3.4.
in equation (3.5) are stationary. Given the stock-year-specific estimates of \( \{\alpha_j^1, \alpha_j^2, \beta_j^1, \beta_j^2\} \), I verify that in more than 87% of cases the conditional mean process in WACD(2,2) is stationary. That is, all roots of both AR(2) and MA(2) characteristic polynomials fall outside the unite circle.

Figure 3.2 provides further information by presenting the empirical distributions of the parameter estimates. \( \beta_2 \) is the only parameter with a significant bulk of point estimates that are close to zero. The other parameter estimates exhibit minimal mass of empirical distributions around zero. Furthermore, despite the significant variation in median durations across and within size categories, parameter estimates are remarkably clustered. For example, I find no strong association between stock size and parameter estimates save for the shape parameter \( \gamma \), suggesting a generally invariant GARCH structure with respect to stock characteristics. Finally, among the parameters of conditional mean component, unlike \( \alpha_1 \), empirical distributions of the other parameters exhibit distinctly bimodal distributions. This suggests that a two-fold conditional mean specification might explain dynamics of trading activity of all stocks.

Theory suggests that the parameters of the conditional mean should be correlated. Given the bimodal empirical densities of \( \alpha_j^1, \beta_1^j, \) and \( \beta_2^j \), an interesting question is whether a two-fold structure of conditional mean is associated with a two-fold correlation structure of parameters. I investigate this possibility by simply plotting the point estimates of different parameter pairs in scatter plots. Figure 3.3 shows the empirical pair-wise association between parameters. The two-type structure that was detected at parameter levels presents itself as a two-fold correlation structure, too. In particular, the association between \( \alpha_1 \) and \( \alpha_2 \) converts from positive to negative as I mover from the cluster of positive-valued \( \hat{\alpha}_1^j \)'s to the cluster comprised of negative \( \hat{\alpha}_1^j \)’s; similarly, while \( \beta_1 \) and \( \beta_2 \) seem to negatively correlate for positive \( \hat{\beta}_1^j \)’s, the association becomes positive for negative \( \hat{\beta}_1^j \)’s.
Figure 3.3: Pair-wise correlations of parameters: WACD(2,2)s are fitted on a stock-year basis. The scatter plots reflect the pair-wise associations between stock-year-specific estimates of parameters \( \{\alpha_1^j, \alpha_2^j, \beta_1^j, \beta_2^j, \omega^j, \gamma^j\} \). Durations measure the time it takes for a sequence of transactions worth $80,000 plus 0.025% of market capitalization. Durations are adjusted for diurnal patterns and overnight regime changes according to 3.4.
The other important observation in Figure 3.2 is the distant center of empirical distribution of the shape parameter $\gamma$ from 1. Virtually, very few $\hat{\gamma}_j$’s are close to, or less than, unity; the point estimates average at 1.68 and, as Table 3.2 indicates, are all significantly different from zero. More importantly more that 99% of stock-yr-specific $\gamma$ estimates are significantly greater than 1. This finding highlights the crucial importance of a specification that assumes Weibull, but not Exponential distribution, to describe standardized durations, and allows the hazard rate to vary with durations. $\gamma$ estimates are mostly greater that unity, implying that the hazard function monotonically increases in $\tilde{x}$, and that assuming Exponentially-distributed standardized durations can greatly bias estimates and predictions of the model.

As mentioned earlier, shape parameter estimates $\hat{\gamma}$’s significantly vary with stock size. To show this, I sort stocks from smallest to largest at the beginning of each year by market-capitalization at the end of the previous year. I assign normalized rank statistics that reflect stock size percentiles on an annual basis, where a percentile statistic 1 reflects the largest market-capitalization that year. Figure 3.4 shows that the shape parameter estimates monotonically rise with stock size, and that the rate of increase also rises in stock size. However, log-number of durations takes a $\cap$-shaped pattern in stock size. This observations underlies the fact that liquidity is “slower” for larger stocks; that is, trading activity is more consistent for these stocks. Indeed, I find that the skewness of duration distributions sharply declines in stock size; indicating fewer periods of very low volume for larger stocks. Intuitively, this observation should translate to greater shape parameters for larger stocks; as time elapses after an event, larger stocks experience disproportionately larger increases of the hazard rate—a transaction event becomes much more likely. As a result, the size pf $\gamma$ depends on the consistency of trading activity, not the number of trade sequences.

Following Engle and Russell (1998), I next turn to examining how close to i.i.d are the realized
Figure 3.4: **Shape parameter estimates and log-number of durations against stock size:** Stock-year specific shape parameter estimates $\gamma_i$ and log-number of durations are plotted against stock size percentile. Size percentiles reflect annual normalized rank statistics that sort stocks from smallest to largest by market-capitalization each year—market-capitalizations are measured at the end of the previous year. Durations measure the time it takes for a sequence of transactions worth $80,000 plus 0.025\%$ of market capitalization.

Standardized durations—recall that the error term $\epsilon$ in Model 3.1 must be *i.i.d.* For each stock-year set of observations, I use the fitted ACD model’s parameters to construct a series of realized standardized durations $\epsilon_{jt}$. I then employ three distinct strategies in my analysis. I first note that the realized standardized durations $\epsilon_{jt}$ in each period must be independent of those in preceding periods. To this end, I explore the autoregressive structure of diurnally-adjusted durations before and after controlling for the ACD’s GARCH structure. Next, I contrast the predictability of standard durations ($\epsilon_{jt}$) with that of diurnally-adjusted durations ($\tilde{x}_{jt}$) using past levels of diurnally-adjusted trading activity; theory suggests that an *i.i.d* process should be unpredictable. Finally, for each stock, I examine how well empirical hazard rates at different levels of standardized duration match the corresponding theoretical hazard implied by a Weibull distribution given the fitted parameters.

I consider an AR(2) specification $y_{jt} = \delta_0 + \delta_1 y_{jt-1} + \delta_2 y_{jt-2} + u_{jt}$ to contrast the autoregressive structure of diurnally-adjusted durations with that of standardized durations. Stock-year-specific
AR(2) structures of $\tilde{x}_{jt}$ and $\epsilon_{jt}$ are estimated separately, so that the corresponding coefficients that best describe the two series can be contrasted. I highlight the massive contribution of the conditional mean in removing temporal dependence by showing much weaker persistence in $\epsilon_{jt}$ compared to $\tilde{x}_{jt}$.

Figure 3.5 shows empirical distributions of the AR(2) persistence parameter point estimates and their t-statistics. Both $\delta_1$ and $\delta_2$ take much smaller magnitudes as I move from the series of diurnally-adjusted durations (for whom empirical distributions of $\hat{\delta}_1$ and $\hat{\delta}_2$ are shown by dashed kernel curves) to standardized durations (solid kernel curves)—$\hat{\delta}_1$ and $\hat{\delta}_2$ of the $\epsilon_{jt}$ series are, on average, at least five times smaller than those of $\tilde{x}_{jt}$. Similar results obtain once I contrast the corresponding t-statistics. As shown in the bottom row of Figure 3.5, t-statistics associated with the AR(2) parameter estimates dramatically shrink as I move from $\tilde{x}_{jt}$ to $\epsilon_{jt}$ series. This shrinkage is to the extent that a sizable portion of t-statistics in between the critical values at 0.1% significance level, which are shown in t-statistics graphs by the vertical dashed lines. Overall, the conditional mean of the WACD(2,2) model successfully captures the temporal dependence in the time series of diurnally-adjusted durations.

I next employ a simple semi-parametric specification to compare the predictability of $\tilde{x}_{jt}$ and $\epsilon_{jt}$. Our only predictor is the previous level of diurnally-adjusted trading activity $\tilde{x}_{jt-1}$. To avoid strong parametric assumptions, I define indicator variable $I^n$ with $n \in \{2, \ldots, 20\}$ that takes 1 if the preceding diurnally-adjusted duration falls between percentiles $\frac{5(n-1)}{100}$ and $\frac{5(n)}{100}$ of diurnally-adjusted trading activity, but takes 0 otherwise. I rank durations of each stock on a monthly basis and assign percentile statistic so that the shortest durations corresponds to activity percentile.

In unreported results, by examining the sum $\hat{\delta}_1^j + \hat{\delta}_2^j$, I verify the stationarity of both AR(2) processes. I also obtain qualitatively identical results as I contrast the two AR(2) processes examining the empirical distributions of $\delta_1 + \delta_2$. 

132
Figure 3.5: **Empirical distributions of AR(2) coefficients and their corresponding t-statistics:**

\[ y_{jt} = \delta_0 + \delta_1 y_{jt-1} + \delta_2 y_{jt-2} + u_{jt} \]

is estimated using the series of \( \tilde{x}_{jt} \) and \( \epsilon_{jt} \) for each stock-year set of observations. The empirical distributions of the corresponding parameter estimates and t-statistics for the two series types are contrasted. Vertical short-dashed lines in the bottom row represent two-sided critical values at 0.1% significance level. Durations measure the time it takes for a sequence of transactions worth $80,000 plus 0.025% of market capitalization.

Finally, for each stock-year set of observations, I estimate

\[ y_{jt} = \pi_0 + \sum_{i=2}^{20} \pi_i I_{jt}^{\pi_i} + u_{jt} \]

where \( y_{jt} \) can be \( \tilde{x}_{jt} \) or \( \epsilon_{jt} \).

Table 3.3 highlights the remarkably weaker predictability of the \( \epsilon \) series than that of the \( \tilde{x} \) series, on average. For both series, high past activity levels possess relatively greater predictive significance for future duration. However, the weights of past trading activity indicators (\( \pi_i \)) in predicting \( \epsilon_{jt} \) are between two to twenty three times smaller that those in predicting \( \tilde{x}_{jt} \). Consistently, the average
Table 3.3: **Predictability of duration and standardized duration series:** \( y_{jt} = \pi_0^j + \sum_{i=2}^{20} \pi_i^n I_{jt} + u_{jt} \)

is estimated using the stock-year specific series if \( \tilde{x}_{jt} \) and \( \epsilon_{jt} \). Mean point estimates and mean t-statistics of \( \{\pi_2 \ldots \pi_{20}\} \) for the two series \( \tilde{x}_{jt} \) and \( \epsilon_{jt} \) are reported. The last row of the table presents the mean regression F-statistics of the two cases.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( \epsilon ) series</th>
<th>( \tilde{x} ) series</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>t-stat</td>
</tr>
<tr>
<td>( \pi_2 )</td>
<td>0.029</td>
<td>0.55</td>
</tr>
<tr>
<td>( \pi_3 )</td>
<td>0.060</td>
<td>1.16</td>
</tr>
<tr>
<td>( \pi_4 )</td>
<td>0.078</td>
<td>1.51</td>
</tr>
<tr>
<td>( \pi_5 )</td>
<td>0.089</td>
<td>1.74</td>
</tr>
<tr>
<td>( \pi_6 )</td>
<td>0.098</td>
<td>1.94</td>
</tr>
<tr>
<td>( \pi_7 )</td>
<td>0.104</td>
<td>2.06</td>
</tr>
<tr>
<td>( \pi_8 )</td>
<td>0.109</td>
<td>2.18</td>
</tr>
<tr>
<td>( \pi_9 )</td>
<td>0.114</td>
<td>2.30</td>
</tr>
<tr>
<td>( \pi_{10} )</td>
<td>0.119</td>
<td>2.42</td>
</tr>
<tr>
<td>( \pi_{11} )</td>
<td>0.123</td>
<td>2.52</td>
</tr>
<tr>
<td>( \pi_{12} )</td>
<td>0.129</td>
<td>2.63</td>
</tr>
<tr>
<td>( \pi_{13} )</td>
<td>0.134</td>
<td>2.74</td>
</tr>
<tr>
<td>( \pi_{14} )</td>
<td>0.137</td>
<td>2.84</td>
</tr>
<tr>
<td>( \pi_{15} )</td>
<td>0.142</td>
<td>2.95</td>
</tr>
<tr>
<td>( \pi_{16} )</td>
<td>0.146</td>
<td>3.05</td>
</tr>
<tr>
<td>( \pi_{17} )</td>
<td>0.149</td>
<td>3.15</td>
</tr>
<tr>
<td>( \pi_{18} )</td>
<td>0.149</td>
<td>3.21</td>
</tr>
<tr>
<td>( \pi_{19} )</td>
<td>0.144</td>
<td>3.21</td>
</tr>
<tr>
<td>( \pi_{20} )</td>
<td>0.108</td>
<td>2.75</td>
</tr>
</tbody>
</table>

Mean F-stat \( 4.09 \) \( 256.2 \)

t-statistics are also smaller by orders of magnitudes once the predictability regressions are estimated for \( \epsilon \) series. Furthermore, the average F-statistic of individual regressions fall by a factor of 62.6 as I substitute \( \epsilon \) for \( \tilde{x} \) in predictability regressions. These finding provide additional supportive evidence that \( \epsilon_{jt} \) is close to \( i.i.d \) processes.

I next explore the match between empirical and predicted hazard rates (conditional on fitted parameters) on a stock-by-stock basis—this approach verifies the validity of Weibull assumption. For each stock-year WACD(2,2) estimate, I first compute the ratios of empirical to predicted hazard rates given \( \epsilon_{jt} \); I call this ratio *hazard error*. Working with this notion of relative error controls for level effects. I then sort \( \epsilon_{jt} \) into twenty equally-sized categories and find the mean hazard error.
Figure 3.6: Point estimates and t-statistics of mean hazard errors: WACD(2,2)s are fitted on a stock-year basis. Using stock-year-specific estimates, hazard errors (ratios of empirical-to-predicted hazard) are computed for each $\epsilon_{jt}$. Observations on each stock-year are sorted into twenty equally-sized groups of standardized duration ($\epsilon_{jt}$), and mean hazard errors are computed by standardized duration group per stock-year set of observations. The top row presents the averages of mean hazard errors at different levels of standardized duration; the bottom row reports the corresponding t-statistics for a null of average error equal to unity (the two horizontal lines stand for critical values at 0.1% significance level). Averages and t-statistics are calculated across stocks and years, controlling for stock size decile. Size deciles are formed at the beginning of each year by sorting stocks into ten equally-sized groups of market-capitalizations at the end of the previous year.
for each $\epsilon_{jt}$ category. I focus on averages of these mean hazard errors across stocks and over time to explore systematic departures from predictions of the Weibull density at different standardized duration groups—note that a correctly specified model implies mean hazard errors of unity.

The top row of Figure 3.6 shows that while the unconditional mean hazard error is close to zero, empirical hazard systematically departures from predictions of a Weibull distribution once one controls for the levels of $\epsilon_{jt}$. However, save for the highest standardized duration level of small firms, the size of departures fall within 60% of Weibull predictions. At the lowest levels of standardized duration, predictions of Weibull specification are very close to the empirical hazard rates. As $\epsilon_{jt}$ slightly rises, Weibull underestimates hazard rate; but at intermediate and high levels of $\epsilon_{jt}$, Weibull predictions fall above empirical hazard rates. Finally, there is sizable underestimation at the highest levels of standardized durations. Overall, except for the highest level of standardized duration, the extent of hazard errors tend to rise with stock size, indicating that after accounting for the conditional mean, hazard rates of smaller stock durations may be best predicted by Weibull.

The bottom row of Figure 3.6 presents the t-statistics of mean hazard error, testing the null that average of mean hazard errors equals one. In majority of cases, departures are statistically significant at 0.1% level—horizontal lines represent critical values.

My extensive examination of trading activity dynamics using a large sample of stocks over a six-year time period reveals important and undocumented facts about the dynamic structure of intra-day trading activity. Among other things, my findings uncover that there are two distinct types of trading activity dynamics in the sense that the empirical distributions of the conditional mean component’s parameters are bimodal; I also document a corresponding two-type parameter correlation structure. However, quite generally, I establish two invariant conditional mean structures
across stocks. My analysis of the standardized duration highlights the significant contribution of the conditional mean process of the WACD(2,2) model in removing the auto-correlation of diurnally-adjusted durations, and that standardized durations are close to being (but not convincingly) i.i.d processes. The other relevant piece of evidence is the systematic but tolerable departures of the empirical hazard from that of the theoretical hazard (predicted by Weibull) rates.

VI Future work: to believe or to not believe in WACD?

Given what is been documented so far, there are two possible ways to continue depending on whether I trust my WACD specification. Accepting the existing WACD(2,2) as the correct specification of trading activity dynamics directs the analysis towards explaining the patterns I observe in parameter estimates based on stock characteristics. Moreover, one can engage in prediction exercises that could relate the findings of my estimation exercise to systematic price dynamics documented in Barardehi et al. (2014) and Barardehi (2014). In contrast, questioning the validity of a WACD specification calls for re-specification of the model in search for significant improvements in performance and/or prediction. Such re-specifications typically concern the conditional mean of an ACD model, since there quite o few simple alternatives that can describe standardized durations.

A To believe in WACD

Extending the study under this circumstance rely on the insight that in modern financial markets where trading algorithms handle execution processes; as a result, arrival of a typically tiny individual transaction may primarily reflect an execution decision rather than the economic behavior of an investor in response to fundamentals like flow of information. In other words, with extensive use
of order-splitting in today’s financial markets, a series of (not necessarily successive) individual transactions can represent the same trading decision. This feature makes the inter-transaction time durations very hard to model as it grossly violates the i.i.d requirement in complicated ways. Thus, inter-transaction time durations are not the right input to ACD models. In contrast, successive execution times of a fixed dollar position that, to a reasonable extent, preserve the information content of inter-transaction time durations, serves as the right input to ACD models.

Taking this direction suggests two empirical exercises:

1. I need to explain the patterns in empirical distributions of estimated ACD parameters and the pair-wise association between these estimates. What explains the two-type GARCH structure? Do stocks switch between the two types, and if so, what explains it? To answer these questions I employ stock attributes and whether a stock was of a certain type the previous year to explain the probability of the stock falling in the type category to which it belongs this year. This analysis can help us better understand the dynamics of trading activity and provide economic insight into why these dynamics may dramatically differ across stocks. As such, my work could have important implications regarding market quality.

2. At a different level, I can test the prediction performances of these ACD estimates. A relevant benchmark is the probability of the model correctly predicting an increase or a decrease in trading activity. This exercise is related to Barardehi (2014) who finds systematically different price dynamics for increases and decreases in trading activity. It also can provide further insight to predictability of trading activity and the associated price impacts (trading costs) as highlighted by Barardehi, Bernhardt, and Davies (2014). This analysis not only may provide guidance to how practitioners might find my measure of trading activity useful, but it also can

---

6I have discussed the merits of this measure earlier.
shed light on how prediction performance of ACD models might vary with stock attributes.

B To not believe in WACD

Our findings may also reflect some sort of mis-specification of the model. For example, the invariant estimated GARCH structure across stocks despite the fundamental differences between those stocks can raise doubt regarding accuracy of the conditional mean specification of ACD models: the conditional mean is too restrictive, and leads to a spuriously invariant structure across stock-years. This sort of mis-specification can also present itself in (a) standardized durations that do not perfectly follow \( i.i.d \) processes or (b) significant mismatch of empirical and theoretical hazard rates. Thus, another line of future work mainly focuses on improving the conditional mean specification.

There is a significant literature on introduction of useful alternatives for the conditional mean, besides the GARCH structure proposed by Engle and Russell (1998), that can improve the performances of ACD models.\footnote{Among others, Jasiak (1998) develops fractionally integrated ACD models arguing that durations may exhibit decay rates that are similar to long memory processes, Bauwens and Giot (2000) develop ACD models with a logarithmic conditional mean processes to to expand the parameter space, Veredas et al. propose a method that jointly handles the dynamics and diurnal patterns, and Bauwens and Giot (2003) condition the dynamics of duration process on the state of the price process.} One of the improvements to the conditional mean process in ACD models is proposed by Bauwens and Veredas (2004) who introduce a latent random component as a part of the conditional mean process. They upgrade the classical ACD structure to a mixture set-up that contains two random components: the latent variable in the conditional mean and the standardized duration. This way, they allow higher flexibility in the structure of the conditional mean and the hazard function. They assert the model as the following, and call it Stochastic Conditional
Duration (SCD).

\[ x_{jt} = \Psi_{jt}\epsilon_{jt} \] where \( \Psi_{jt} = e^{\psi_{jt}} \) (3.6)

\[ \psi_{jt} = \omega_{i}^{j} + \sum_{i=1}^{r} \alpha_{i}^{j} x_{jt-i} + \sum_{i=1}^{s} \beta_{i}^{j} \psi_{jt-i} + u_{jt} \]

\[ u_{jt} | I_{jt-1} \overset{i.i.d}{\sim} N(0, \sigma_{j}^{2}) \]

\[ \epsilon_{jt} | I_{jt-1} \overset{i.i.d}{\sim} f(\epsilon, \phi) \]

\( f(\epsilon, \phi) \) is any distribution with positive support. \( I_{jt-1} \) is the available information prior to event \( t \), and allows utilization of a Kalman Filter in the estimation process given normality of \( u_{jt} \).

The justification behind this random component could be that flow of information or unexpected liquidity provision, whose dynamic structure may not be known, vary randomly. Despite the greater flexibility implied by this specification, the normality assumption turns out to be pretty restrictive. Importantly, Kalman Filter that employs linear estimators is optimal if the normality assumption holds in practice. relatedly, Bauwens and Veredas (2004) who study the durations of a fixed number of shares, show that the estimated random components do not follow normal distributions for any of the 14 stocks they study. This indicates that the normality assumption might render futile the efforts to make the conditional mean flexible by over-restricting the parameter space. To this end, relaxing the normality assumption on the latent variable in model (3.6) is desirable.

Greenshtein et al. (2014) extend the Kalman Filter to the setting where the variable of interest can follow any distribution. In this case, the traditional Kalman Filter may not be the best predictor. They employ Empirical Bayes methods to develop a non-linear improvement to Kalman Filter that assumes no particular underlying distribution. In a tentative joint work with Jiaying Gu, we aim at combining Greenshtein et al. (2014)’s Empirical Bayes exercise and the SCD structure. Then we
can allow a flexible conditional mean process without imposing a particular distributional structure for the random component $u_{jt}$.

An ultimate goal of estimating the dynamic structure of trading activity is to forecast the future transaction event(s) or trading activity levels. Thus, the goal, besides the regular specification tests, is to investigate the out-of-sample forecast improvements of this mixture set-up vis-a-vis other standard setups. There, we follow Dufour and Engle (2000) who contrast the predictability of the time between successive transactions given different classes of ACD specifications. They develop forecasting models to predict one step-ahead and multi step-ahead. They use different loss functions satisfying smoothness conditions to construct tests on loss differentials across different alternatives to determine the “best” forecast performance.
Appendix A

Does endogenous consumption/provision of liquidity predictably persist? This validation step explores how signed trade imbalances vary with trading activity and predicted trading activity. If the endogenous consumption and provision of liquidity can be predicted in the same way as activity then these patterns should be similar. To investigate, we employ the same estimation approaches to investigate the relationships between signed trade imbalances and both trading activity and predicted trading activity.

Figure A.1 and Table A.1 show that signed trade imbalances still co-vary with predicted (instrumented) activity. This association, however, is much weaker than that with the level of trading activity itself—the relative change in average signed trade imbalance as activity rises from its lowest level to its highest is at least twice what it is as instrumented activity goes from its lowest level to its highest. This finding contrasts with the striking effects that instrumentation has on trading cost/market (predicted) activity relationships shown in the text.

Calendar time relationships. A positive relationship between volatility and activity exists in calendar time. To illustrate, we estimate the return volatility that would result if a given level of activity was maintained for an entire trading day. We first estimate implied return volatilities for a given level of market (instrumented) activity. For a given year \( y \) and stock \( j \), we calculate the

\[ \text{Implied Return Volatility} \]

\[ \text{Implied Return Volatility} = F \text{-statistics above 10.} \]
Figure A.1: **Average signed trade imbalances versus predicted trading activity.** Average signed trade imbalances of trade sequences versus predicted trading activity for the deciles of smaller and larger stocks. The reported average signed-dollar volume is the average of firm-specific annual mean dollar-weighted proportion of buy- (sell-) oriented trades (realized over trade sequences), computed per predicted activity group. The average of means is taken across all firms in a given size decile, over the entire sample period by predicted activity group. Average signed trade imbalances versus trading activity is displayed to highlight differences. Stock-year observations with first stage regression $F$-statistics below 10 are dropped.
Table A.1: Signed trade imbalances versus trading activity and instrumented trading activity.

For each stock-year observation, the absolute return realized over a trade sequence is regressed on the corresponding activity percentile (Panel A) or instrumented activity percentile (Panel B). The empirical distributions of these point estimates within each size decile are employed to draw statistical inference. The sampling distribution is assumed to follow a student $t$ distribution (“Stock level estimates”). Observations are pooled across stocks within a size decile. Absolute return is regressed on the corresponding activity percentile; partial regression outputs are reported (“Size group level estimates”). The point estimates reflect the amount by which trade imbalances change as activity (instrumented activity) rises from lowest to highest percentile.

**Panel A: Sensitivity of trade imbalance to trading activity**

<table>
<thead>
<tr>
<th>Size decile</th>
<th>Stock level estimates</th>
<th></th>
<th></th>
<th>Size group level estimates</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean slope</td>
<td>Std. Err.</td>
<td>95% CI</td>
<td>Estimate</td>
<td>Std. Err.</td>
<td>95% CI</td>
</tr>
<tr>
<td>1</td>
<td>0.140</td>
<td>0.0007</td>
<td>(0.139 , 0.141)</td>
<td>0.140</td>
<td>0.0003</td>
<td>(0.139 , 0.140)</td>
</tr>
<tr>
<td>2</td>
<td>0.139</td>
<td>0.0008</td>
<td>(0.137 , 0.140)</td>
<td>0.138</td>
<td>0.0003</td>
<td>(0.138 , 0.139)</td>
</tr>
<tr>
<td>3</td>
<td>0.135</td>
<td>0.0009</td>
<td>(0.134 , 0.137)</td>
<td>0.135</td>
<td>0.0002</td>
<td>(0.135 , 0.136)</td>
</tr>
<tr>
<td>4</td>
<td>0.131</td>
<td>0.0009</td>
<td>(0.129 , 0.132)</td>
<td>0.130</td>
<td>0.0002</td>
<td>(0.130 , 0.131)</td>
</tr>
<tr>
<td>5</td>
<td>0.127</td>
<td>0.0010</td>
<td>(0.126 , 0.129)</td>
<td>0.127</td>
<td>0.0002</td>
<td>(0.127 , 0.128)</td>
</tr>
<tr>
<td>6</td>
<td>0.123</td>
<td>0.0010</td>
<td>(0.122 , 0.125)</td>
<td>0.123</td>
<td>0.0002</td>
<td>(0.123 , 0.124)</td>
</tr>
<tr>
<td>7</td>
<td>0.125</td>
<td>0.0011</td>
<td>(0.123 , 0.126)</td>
<td>0.125</td>
<td>0.0002</td>
<td>(0.124 , 0.125)</td>
</tr>
<tr>
<td>8</td>
<td>0.119</td>
<td>0.0011</td>
<td>(0.118 , 0.121)</td>
<td>0.119</td>
<td>0.0002</td>
<td>(0.119 , 0.119)</td>
</tr>
<tr>
<td>9</td>
<td>0.112</td>
<td>0.0009</td>
<td>(0.110 , 0.113)</td>
<td>0.112</td>
<td>0.0002</td>
<td>(0.111 , 0.112)</td>
</tr>
<tr>
<td>10</td>
<td>0.103</td>
<td>0.0008</td>
<td>(0.101 , 0.104)</td>
<td>0.103</td>
<td>0.0002</td>
<td>(0.102 , 0.103)</td>
</tr>
</tbody>
</table>

**Panel B: Sensitivity of trade imbalance to instrumented trading activity**

<table>
<thead>
<tr>
<th>Size decile</th>
<th>Stock level estimates</th>
<th></th>
<th></th>
<th>Size group level estimates</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean slope</td>
<td>Std. Err.</td>
<td>95% CI</td>
<td>Estimate</td>
<td>Std. Err.</td>
<td>95% CI</td>
</tr>
<tr>
<td>1</td>
<td>0.047</td>
<td>0.0006</td>
<td>(0.046 , 0.047)</td>
<td>0.047</td>
<td>0.0003</td>
<td>(0.046 , 0.047)</td>
</tr>
<tr>
<td>2</td>
<td>0.054</td>
<td>0.0005</td>
<td>(0.053 , 0.055)</td>
<td>0.054</td>
<td>0.0003</td>
<td>(0.054 , 0.055)</td>
</tr>
<tr>
<td>3</td>
<td>0.058</td>
<td>0.0005</td>
<td>(0.057 , 0.058)</td>
<td>0.058</td>
<td>0.0002</td>
<td>(0.057 , 0.058)</td>
</tr>
<tr>
<td>4</td>
<td>0.059</td>
<td>0.0005</td>
<td>(0.058 , 0.060)</td>
<td>0.059</td>
<td>0.0002</td>
<td>(0.059 , 0.060)</td>
</tr>
<tr>
<td>5</td>
<td>0.061</td>
<td>0.0005</td>
<td>(0.060 , 0.061)</td>
<td>0.061</td>
<td>0.0002</td>
<td>(0.060 , 0.061)</td>
</tr>
<tr>
<td>6</td>
<td>0.060</td>
<td>0.0005</td>
<td>(0.060 , 0.061)</td>
<td>0.060</td>
<td>0.0002</td>
<td>(0.060 , 0.061)</td>
</tr>
<tr>
<td>7</td>
<td>0.062</td>
<td>0.0005</td>
<td>(0.062 , 0.063)</td>
<td>0.063</td>
<td>0.0002</td>
<td>(0.062 , 0.063)</td>
</tr>
<tr>
<td>8</td>
<td>0.062</td>
<td>0.0005</td>
<td>(0.061 , 0.063)</td>
<td>0.062</td>
<td>0.0002</td>
<td>(0.061 , 0.062)</td>
</tr>
<tr>
<td>9</td>
<td>0.059</td>
<td>0.0004</td>
<td>(0.058 , 0.060)</td>
<td>0.059</td>
<td>0.0002</td>
<td>(0.058 , 0.059)</td>
</tr>
<tr>
<td>10</td>
<td>0.048</td>
<td>0.0005</td>
<td>(0.047 , 0.048)</td>
<td>0.048</td>
<td>0.0002</td>
<td>(0.047 , 0.048)</td>
</tr>
</tbody>
</table>

return variance and average duration (in seconds) of trade sequences belonging to (instrumented) activity level $i$. Using these values, and assuming independent price movements across different trade sequences, we estimate the corresponding calendar day return standard deviation for a given
activity level as:

\[ S_i^D(y, j) = \sqrt{\text{Return variance of trade sequences} \times \text{(number of seconds in trading day)}} \]

\[ \text{Average duration in seconds of trade sequences} \]

Size decile estimates are obtained by averaging \( S_i^D(y, j) \) across all stocks in size decile \( x \).

Figure A.2: **Average daily return volatility by activity group.** Average return volatility of 0.04% of market capitalization in 2001–2012 versus activity level for the three deciles of smallest and largest stocks. The reported average daily return volatility is the average of firm-specific annually computed daily standard deviations of returns \( S_i^D(y, j) \), computed per duration group. The average of daily standard deviations is taken across all firms in a given size decile, over the entire sample period, by activity level.

Figure A.2 shows that estimated daily return volatility rises sharply at high levels of activity.\(^2\)

\(^2\)An analogous exercise for yearly subsamples reveals similar patterns. Daily return volatility is generally higher
Daily return volatility, however, only rises with instrumented activity at about 20% of the rate with raw activity. This finding suggests that the strong relation between activity and daily volatility largely reflects the endogenous nature of activity. We also see that average daily return volatility varies with firm size—even using instrumented activity, average daily return volatility is over twice as high for the smallest stocks as for the largest.

**Robustness to Error Structure.** We impose a robust standard error structure with firm level clusters. Our approach follows Petersen (2009) to account for potential cross-firm and cross-time correlations in panel data error terms. Failing to account for such correlations would lead to underestimated standard errors. To account for common fixed firm and time effects, we estimate a fixed-effect model using generalized least squares (GLS) with month dummies. We also consider two other potential error term correlations: non-fixed time and non-fixed firm effects. Non-fixed firm effects are those causing temporal autocorrelations in a given firm’s error term. Such effects can arise when a shock to stock $j$ persists (e.g., the effect of a technology shock on stock’s $j$’s performance, which decays over time). Non-fixed time effects make error terms of different firms related in a given period of time where the correlation differs for different pairs of firms (see Petersen (2009) or Cameron et al. (2008)). These effects can arise when industry-specific shocks affect most firms in an industry similarly, leaving firms in other industries unaffected.

Fama and MacBeth (1973) adjust for non-fixed time effects by running cross-sectional regressions period by period. Point estimates from different cross-sections give the empirical distributions of parameters of interest. Further statistical inference using the Fama-MacBeth method is based on the averages and standard deviations of empirical distributions, and hence requires temporal independence of cross-sectional point estimates to obtain unbiased estimates. Extensive time de-
dependence in our sample precludes using this approach.³

Some estimation strategies call for multi-dimensional clustering when a researcher suspects errors to correlate in multiple dimensions (Cameron, et al. (2008)). However, techniques that allow for multi-dimensional and non-nested clustered error terms have been developed only for ordinary least squares (OLS) estimation. OLS estimation using panel data is equivalent to pooling the information content of different periods about a given stock. Such aggregation is inappropriate, so we employ a fixed-effects GLS approach as our main estimation strategy. In unreported results, we estimate model (1.6) using OLS while clustering by firm, time, and both. Standard errors of estimates given two dimensional clustering are only modestly larger (less than 40% higher) than when we account for firm clusters only. It is also inappropriate to try to control for non-fixed time effects by clustering by time when we use fixed-effects GLS, since the unbalanced nature of the panel makes the clusters non-nested.⁴

These concerns lead us to use industry-specific shocks to control for non-fixed time effects associated with shocks that affect firms within an industry similarly. Using four-digit GIC industry codes, we identify about 65 industries per month. For each firm \( j \), we compute the associated average industry returns on a monthly basis excluding its own return. To de-trend, we find the excess industry return against the equally weighted monthly average return. The first difference of the monthly industry excess return is added to model (1.6), as a proxy of non-fixed time effects. Adding this variable to the model has only modest (less than 20%) effects on all parameter estimates save for lagged return, and standard errors change even less. In particular, estimates of the \( BBD \) coefficient are altered by less than two percent.

³Petersen (2009) highlights downward biases in Fama-MacBeth estimated standard errors in the presence of firm effects.
⁴Nichols and Schaffer (2007) note that with differently-sized clusters, inference could be an issue. We have differently-sized clusters, but with approximately 1400 firm-level clusters, asymptotic properties should hold.
An alternative way to deal with non-fixed time effects is to model the serial correlation of error terms parametrically. Accordingly, we consider error terms that follow an AR(1) process, \( \epsilon_{j,t} = \rho \epsilon_{j,t-1} + \nu_{j,t-1} \), with \( \nu_{it} \sim iid N(0, \Sigma^2_\nu) \). We estimate model (1.6) with this error structure using a fixed-effects GLS strategy. As when we cluster at the firm level, the autoregressive specification of errors does not affect point estimates. Regardless of the liquidity measure employed, the associated standard errors are slightly smaller than those reported in Table 1.3.

Overall, these findings support our modeling approach of allowing for firm clustering, along with stock fixed effects and time dummies. This approach mitigates downward biases in estimated standard errors. One need not cluster in the time dimension since the products of Fama-French factors and their corresponding betas along with time dummies capture most commonalities in the time variation of stock excess returns.

**Temporal comparisons with standard illiquidity measures.** We posited at the outset that standard liquidity measures do a worse job of capturing illiquidity in today’s markets. We now investigate how BBD’s performance in recent years compares with those of standard measures. To proceed, we estimate (1.6) using the residuals \( z_{t-1}^F \) and \( \tilde{z}_{t-1}^F \) from the orthogonal decompositions in the two sub-periods of early years, April 2001 to December 2006, and recent years, January 2007 to December 2012. We compare the version of BBD based on all trade sequences with AML, AMH, and PSP.

Table A.2 reports only the coefficients on the illiquidity variables, because the coefficients of the other pricing variables remain stable across the sub-periods. The table reveals that after filtering out the covariation with these standard liquidity measures, BBD contains significant incremental information about liquidity, especially in more recent years. In the recent 2007–2012 period, the
Table A.2: Illiquidity coefficients in the two sub-periods, after orthogonal decomposition. The table reports panel estimation partial results of (1.6) for two subsamples, April 2001–Dec 2006 and Jan 2007–Dec 2012. Standard errors are clustered at the stock level. Stock fixed effects capture any fixed heteroskedasticity. Month dummies capture common variation caused by market condition changes. The coefficients of $z_{t-1}$ and $\tilde{z}_{t-1}$ are reported, the version of BBD based on all trade sequences is contrasted against the three standard liquidity measures. Observations in the upper 1% tail of the relevant liquidity measures are excluded. Robust standard errors are reported in parentheses. The symbols *, **, and *** denote significance at 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AML</td>
<td>AMH</td>
</tr>
<tr>
<td>$F$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$z_{t-1}$</td>
<td>0.228***</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>(0.087)</td>
<td>(0.037)</td>
</tr>
<tr>
<td>$\tilde{z}$</td>
<td>−0.243**</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(0.103)</td>
<td>(0.002)</td>
</tr>
</tbody>
</table>

BBD residuals are priced, but those from the other three illiquidity measures are not. In the earlier period, percentage bid-ask spreads (PSP) appear to do the best job of capturing illiquidity. These findings indicate that the relative qualities of standard liquidity measures have deteriorated as trading volume exploded, bid-ask spreads shrunk and trading strategies took on dynamic dimensions. This analysis indicates that the BBD trade-time liquidity measures better capture illiquidity concerns in today’s trading environments.
Figure A.3: **Empirical cumulative density functions of duration by size decile.** Durations are computed based on \(\theta = 0.04\) and their CDFs are presented for two subsets of large and small stocks, over the entire period (2001–2012), and the early (2001–2006) and later (2007–2012) sub-periods. To calculate the CDF of durations for stock \(j\) in year \(y\), every month, we first sort durations into 20 buckets, each containing 5% of observations. Then we pool buckets across different months. The median of these durations corresponds to the quantile statistic falling at the midpoint of that duration bucket. For example, the median in the pooled bucket of shortest 5% estimates quantile statistic 2.5%; that for next bucket estimates quantile statistic 7.5%; etc. For each duration bucket, we average the stock-specific quantile statistic estimates over years and stocks in size decile \(x\). Plotting these averages against their relevant percentile points, for each size portfolio, yields the empirical CDF of durations conditional on firm size.
Figure A.4: Average absolute return, return volatility, and average number of trades versus signed trade imbalance. Average absolute return, average return volatility, and mean number of trades for trade sequences versus signed trade imbalance for deciles of smaller and larger stocks. We first sort trade sequences into ten equally-sized groups of percent signed trades, every month. The reported averages are the sample averages of firm-specific annually-computed mean absolute return, return standard deviation, and mean realized depth (all assessed over sequences), computed per signed trade imbalance decile. The averages of means are taken across all firms in a given size decile, over the entire sample period by signed trade imbalance level.
Appendix B

Robustness Test 3: I investigate the extent to which stock attributes such as stock size, introduce heterogeneity in my findings on the variations of price dynamics with stock-specific market activity levels. To this end, I compare the empirical distributions of $\hat{\rho}_{jy}$ and $\tilde{\rho}_{jy}$ coefficients in Model (2.5) at lowest and highest starting/concluding trading activity quintiles, controlling for stock size. I decompose the sample into firms in the top 30%, middle 40% and bottom 30% of market capitalizations at the beginning of each year. To provide conservative estimates against selection-oriented heterogeneity, I discard increases and decreases in activity that exceed twenty activity percentiles.

Figures B.1 and B.2 present the empirical distributions of $\hat{\rho}_{jy}$ and $\tilde{\rho}_{jy}$, respectively, by size group. To shed light on the economic magnitudes in Table B.1, I present corresponding sample means and 99% confidence intervals for the estimates of return persistence and reversals by size decile, given highest and lowest starting/concluding activity levels.

The distinct kernel densities in Figure B.1 show that reversals associated with relative reductions in trading activity are strongest in least active markets. This holds for all firm size groups and is unaffected by whether I control for starting or concluding stock-specific market activity level. Table B.1 shows that in the most active markets the reversals associated with relative activity reductions are tiny, averaging about .001 and are statistically insignificant for most size deciles. Whereas,

1Findings are robust to relaxing this constraint.
Figure B.1: Empirical distributions correlation parameter estimates that isolate decreases in activity by starting/concluding stock-specific market activity level for stock size groups. Every year, model (2.5) is estimated stock by stock, controlling for concluding/starting trading activity quintiles. Estimates are obtained using the sample of trade sequences that are associated with a decrease in activity (yielding $\hat{\rho}_{jy}$). Decreases in activity that exceed twenty percentiles are excluded. The top (bottom) row reports the empirical density estimates at highest and lowest starting (concluding) activity levels.
Figure B.2: Empirical distributions correlation parameter estimates that isolate increases in activity by starting/concluding stock-specific market activity level for stock size groups. Every year, model (2.5) is estimated stock by stock, controlling for concluding/starting trading activity quintiles. Estimates are obtained using the sample of trade sequences that are associated with an increase in activity (yielding $\tilde{\rho}_{yy}$). Increases in activity that exceed twenty percentiles are excluded. The top (bottom) row reports the empirical density estimates at highest and lowest starting (concluding) activity levels.
Table B.1: Sample means and confidence intervals of correlation parameter estimates, isolating increases from reductions in trading activity in most and least active markets by stock size decile. Model 2.5 is estimated year by year, stock by stock, for the highest and lowest starting/concluding activity quintiles, on subsamples representing increases in activity (yielding $\hat{\rho}$) and decreases in activity (yielding $\tilde{\rho}$). Increases and decreases in activity that exceed twenty percentiles are excluded. Sample means and 99% confidence intervals of the correlation parameters are reported by stock size decile. Panel A (Panel B) reports the sample means and confidence intervals of the correlation parameters at different starting (concluding) activity levels.

### Panel A: By starting activity level

<table>
<thead>
<tr>
<th>Size decile</th>
<th>Low starting activity</th>
<th>High starting activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{\rho}$</td>
<td>99% CI</td>
</tr>
<tr>
<td>1</td>
<td>.015</td>
<td>(.005, .025)</td>
</tr>
<tr>
<td>2</td>
<td>-0.004</td>
<td>(-.012, .004)</td>
</tr>
<tr>
<td>3</td>
<td>-0.015</td>
<td>(-.021, -.008)</td>
</tr>
<tr>
<td>4</td>
<td>-0.013</td>
<td>(-.019, -.008)</td>
</tr>
<tr>
<td>5</td>
<td>-0.009</td>
<td>(-.014, -.004)</td>
</tr>
<tr>
<td>6</td>
<td>-0.008</td>
<td>(-.013, -.003)</td>
</tr>
<tr>
<td>7</td>
<td>-0.011</td>
<td>(-.016, -.006)</td>
</tr>
<tr>
<td>8</td>
<td>-0.013</td>
<td>(-.017, -.009)</td>
</tr>
<tr>
<td>9</td>
<td>-0.017</td>
<td>(-.021, -.012)</td>
</tr>
<tr>
<td>10</td>
<td>-0.023</td>
<td>(-.027, -.018)</td>
</tr>
</tbody>
</table>

### Panel B: By concluding activity level

<table>
<thead>
<tr>
<th>Size decile</th>
<th>Low concluding activity</th>
<th>High concluding activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{\rho}$</td>
<td>99% CI</td>
</tr>
<tr>
<td>1</td>
<td>.002</td>
<td>(.001, .016)</td>
</tr>
<tr>
<td>2</td>
<td>-0.020</td>
<td>(-.030, -.010)</td>
</tr>
<tr>
<td>3</td>
<td>-0.028</td>
<td>(-.036, -.020)</td>
</tr>
<tr>
<td>4</td>
<td>-0.023</td>
<td>(-.031, -.016)</td>
</tr>
<tr>
<td>5</td>
<td>-0.020</td>
<td>(-.027, -.014)</td>
</tr>
<tr>
<td>6</td>
<td>-0.018</td>
<td>(-.025, -.012)</td>
</tr>
<tr>
<td>7</td>
<td>-0.021</td>
<td>(-.027, -.015)</td>
</tr>
<tr>
<td>8</td>
<td>-0.026</td>
<td>(-.031, -.020)</td>
</tr>
<tr>
<td>9</td>
<td>-0.031</td>
<td>(-.036, -.025)</td>
</tr>
<tr>
<td>10</td>
<td>-0.041</td>
<td>(-.047, -.035)</td>
</tr>
</tbody>
</table>
similar estimates at lowest activity quintiles average, approximately, at −.098 and are significant for all size deciles. Consistent with previous findings, Figure B.2 highlights that, independent of whether I control for the concluding or starting trading activity level, increases in activity are significantly more likely to feature positive return correlations at “High” activity levels. Table B.1 shows that increases in trading activity associate slight reversals in least active markets; in sharp contrast, they associate strong and significant return persistence in the most active markets. \( \hat{\rho}_{ij} \)’s average at −.016 in least active markets and .089 in most active markets.

**Robustness Test 4:** I estimate sensitivities of signed trade imbalance and price impacts of current sequence to past changes in trading activity, controlling for the starting level of activity. Figure B.3 presents estimated sensitivities of current signed trade imbalance to past changes in activity, controlling for starting levels of activity. The findings are similar to those documented controlling for concluding activity levels. Consistent with the contemporaneous relationships, signed trade imbalances are positively correlated with past changes in activity.

Compared to when I control for concluding levels of stock-specific market activity, sensitivities to past changes are larger and all significant with magnitudes that halve in the 2007-2012 period. The bottom row of Figure B.3 shows that the differences are mainly contributed by the massively large positive correlations that associate increases in trading activity; the positive correlations following reductions in trading activity are relatively much smaller. Again, I note that these correlations pick up the level effects associated with the following higher activity level that correspond to greater signed trade imbalance.

In Figure B.4 consistent with my previous findings, most price impact sensitivities are significantly negative, implying that reductions (increases) in activity are followed by larger (smaller)
Figure B.3: Sensitivity of current signed trade imbalance to past changes in activity by concluding stock-specific market activity level. The sample is decomposed by year, size decile, and starting activity group—observations are pooled across different stocks. Within each category, signed trade imbalance is regressed on the past change in activity percentile, inversely weighting observations on each stock by its number of trade sequences that year. The point estimates and the associated t-statistics are reported by activity groups and sub-periods (first row). Similar estimates are obtained, isolating samples of increases and decreases in trading activity (second row).

price impacts. This feature is preserved when I focus on increases versus decreases in activity, separately. These results are also subject to conflating the level effects with those of changes.
Figure B.4: Sensitivity of current price impact to past changes in activity by concluding stock-specific market activity level. The sample is decomposed by year, size decile, and starting activity group—observations are pooled across different stocks. Within each category, signed trade imbalance is regressed on the past change in activity percentile, inversely weighting observations on each stock by its number of trade sequences that year. The point estimates and the associated t-statistics are reported by activity groups and sub-periods (first row). Similar estimates are obtained, isolating samples of increases and decreases in trading activity (second row).
References


Blume, M.E., D.B. Keim (2012). Institutional investors and stock market liquidity: trends and


164


metrica 65(3), 557–586.