From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Economics

2022

Essays on Institutional Trading and Liquidity

Schmickler, Simon

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01k3569755p

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

Essays on Institutional Trading and Liquidity

Simon Niklas Maria Schmickler

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Economics
Adviser: Motohiro Yogo

September 2022

© Copyright by Simon Niklas Maria Schmickler, 2022.

All Rights Reserved

Abstract

This dissertation shines light on financial market liquidity from both sides: its supply by market makers and its demand by investors.

Chapter 1, co-authored with Pedro Tremacoldi-Rossi, investigates how exchanges and regulators can improve the liquidity and stability of modern financial markets through liquidity provision obligations and incentives. We exploit two market maker programs as natural experiments using unique message-level trade and quote data from the Brazilian stock exchange that reveal market participants' identities. We find combining obligations and incentives improves and stabilizes liquidity which attracts volume and lifts asset prices. In normal times, these positive effects are driven by the incentives, while tight obligations constrain market makers and decrease market quality. In crises, however, the results flip: stocks with larger incentives experience worse liquidity dry-ups because voluntary liquidity providers withdraw; in contrast, tight obligations mitigate liquidity dry-ups because mandatory intermediaries step in as the liquidity providers of last resort. Overall, our results suggest exchanges and regulators should combine incentives with countercyclical liquidity provision obligations.

Chapter 2 proposes a new method to isolate a plausibly exogenous component of mutual fund flows to estimate the price impact of fire sales. The method addresses a potential reverse causality problem: instead of mutual fund outflows inducing fire sales, which drive down prices, poor stock returns reduce mutual fund returns, which in turn trigger outflows. The solution is to construct a new instrument from high-frequency surprise flows. Using surprise flows, I find equity markets are deeper than previously suggested.

Chapter 3 uses the reinvestment of corporate payouts by financial institutions as a nonfundamental shock to asset prices to estimate the slope of the demand curve for stocks and the real effects of stock returns on corporate financing and investment. I find dividends in particular generate payment date price pressure but no announcement date news spillover effects, suggesting that dividend-induced trading is plausibly exogenous to fundamentals. Using this natural experiment, I provide a new asset demand elasticity estimate and document a releveraging market feedback effect: firms respond to an exogenous stock price increase by issuing debt and use the funds to invest.

Acknowledgements

For his constant guidance, support, and encouragement throughout my graduate studies, I am grateful to my advisor Motohiro Yogo. I am also indebted to Markus Brunnermeier and Wei Xiong, the other members of my dissertation committee, for many instructive conversations as well as for their advice and backing on the job market. I would also like to thank my placement director Gianluca Violante for his help and support during the job market. Further, I would like to thank all Princeton faculty who taught and inspired me. I am also grateful for the support of my friends and of my family.

For helpful comments and discussions, I thank Yacine Ait-Sahalia, Caio Almeida, Mathias Büchner, Jasmin Gider, Moritz Lenel, Jonathan Payne, Adrien Matray, Karsten Müller, Don Noh, Christopher Sims, Yannick Timmer, Christian Westheide and seminar participants at the Brazilian stock exchange B3, Yale SOM, Columbia GSB, Northwestern Kellogg, Berkeley Haas, INSEAD, Dartmouth Tuck, Vanderbilt Owen, Georgetown McDonough, Johns Hopkins Carey, Washington University Olin, University of Maryland Smith, University of Miami Herbert, University of Virginia Darden, Federal Reserve Board, PIMCO, the BlackRock Applied Research Award, Young Economist Symposium, and Princeton University. I thank the Brazilian stock exchange B3 for providing data. I also thank Ben Bernanke, Markus Brunnermeier, and the Bendheim Center for Finance for supporting my research with the inaugural Ben Bernanke Prize in Financial and Monetary Economics.

Contents

Abstract 3

Acknowledgements 4

1 In Good Times and in Bad: High-Frequency Market Making Design, Liquidity, and Asset Prices 12

1.1 Introduction 12

1.2 Data and Empirical Strategy 19

1.2.1 Institutional Details 19

1.2.2 The Natural Experiments 20

1.2.3 Observing Market Makers' Trades and Quotes 23

1.2.4 Comparison to Alternative Data Sources 24

1.2.5 Variable Construction 25

1.2.6 Summary Statistics 27

1.3 Results 29

1.3.1 Liquidity Provision Obligations and Incentives in Normal Times 29

1.3.2 Liquidity Provision Obligations and Incentives During Crises 35

1.3.3 How to Assign Market Makers to Assets? 38

1.4 Conclusion 44

Figures 46

Tables 60

2 Identifying the Price Impact of Fire Sales Using High-Frequency Surprise Mutual Fund Flows 77

2.1 Introduction . . . 77

2.2 Empirical Strategy . . . 82

2.2.1 Mutual Fund Flow-Induced Trading . . . 82

2.2.2 Decomposition of the Mutual Fund Flow Instrument . . . 83

2.3 Data . . . 87

2.3.1 Mutual Fund Data . . . 87

2.3.2 Portfolio Holdings Data . . . 88

2.3.3 Stock Level Data . . . 88

2.3.4 Summary Statistics . . . 89

2.4 Results . . . 91

2.4.1 Mutual Fund Flows and Mutual Fund Returns . . . 91

2.4.2 Flow-Induced Trading and Stock Returns . . . 92

2.4.3 The New Instrument - Surprise Flow-Induced Trading . . . 94

2.4.4 Which Mutual Fund Flows Trigger Fire Sales/Purchases? . . . 97

2.4.5 Robustness . . . 98

2.5 Conclusion . . . 100

Figures . . . 101

Tables . . . 109

3 Payout-Induced Trading, Asset Demand Elasticities, and Market Feedback Effects 114

3.1 Introduction . . . 114

3.2 Data . . . 120

3.2.1 Institutional Stock Holdings . . . 120

3.2.2 Firm Payouts . . . 121

3.2.3 Firm Characteristics . . . 122

3.2.4 Summary Statistics . . . 123
3.3 Price Pressure Effects . . . 125
3.3.1 The Demand Shock . . . 125
3.3.2 Identification Strategy . . . 126
3.3.3 The Return Pattern . . . 128
3.3.4 Institutional Heterogeneity . . . 132
3.4 Market Feedback Effects . . . 134
3.4.1 Identification Strategy . . . 135
3.4.2 The Investment Response . . . 136
3.4.3 Expected payout-induced trading . . . 138
3.4.4 Capital Structure Rebalancing Effect . . . 139
3.4.5 Robustness . . . 139
3.5 Conclusion . . . 140
Figures . . . 142
Tables . . . 151

List of Figures

1.1 The Timeline of B3's Market Maker Programs 46
1.2 Institutional Background 47
1.3 Example of Bunching at the Minimum Quantity (at 25) 48
1.4 Quoting and Trading: Mandatory vs. Voluntary Activity 49
1.5 Fraction of Time On Each Price Level: Mandatory vs. Voluntary Activity 50
1.6 Illustration of the Mandatory vs. Voluntary Activity 51
1.7 Dynamic Treatment Effects of DMMs on Liquidity, Volume, and Asset Prices 52
1.8 Dynamic Treatment Effects of Designated Market Makers on Volatilities 53
1.9 The Anatomy of the COVID-19 Crash 54
1.10 Share of Liquidity Provision During the COVID-19 Crash 55
1.11 Effects of Obligations, Effects of Incentives, and Overall Program Effects on Liquidity, Volume, and Asset Prices During the COVID-19 Crash 56
1.12 The Evolution of the Common Intermediation Network (10 Stock Subsample) 57
1.13 Distribution of Program Parameters 58
1.14 Illustration of the Mandatory vs. Voluntary Activity: Quoting of Tucana Bay in Ambev on 1st Day of 2018 Program 59

2.1 Quarterly versus Daily Mutual Fund Flow-Induced Fire Sale Event Study 101
2.2 Illustration of the Reverse Causality Problem 102
2.3 Regression of Daily Mutual Fund Flows on Lags of Flows and Returns 103

2.4 Regression of Daily Flow-Induced Trading on Lags of FIT and Stock Returns 104
2.5 Mutual Fund Trading in Response to Mutual Fund Flows 105
2.6 Regressions of Quarterly Stock Returns on Surprise Flow-Induced Trading for Different Flow Cutoffs 106
2.7 Regression of Daily Mutual Fund Flows on Lags of Flows and Returns Index Funds Only 107
2.8 Regressions of Quarterly Stock Returns on Surprise Flow-Induced Trading for Different Flow Cutoffs - Scaled by Shares Outstanding 108

3.1 Illustration of Payout-Induced Trading Mechanism 142
3.2 Illustration of Identification Strategy 143
3.3 Total Cash Payouts vs. Alternative Sources of Asset Demand Shocks 144
3.4 The Payout-Induced Trading Shock 144
3.5 Ordinary Dividends 145
3.6 Abnormal Returns Around Dividend Announcement and Payment Dates 146
3.7 Abnormal Returns Around M&A Announcement and Payment Dates 147
3.8 Abnormal Returns Around Share Repurchase Announcement Dates 148
3.9 Price Impact of Payout-Induced Trading – Institutional Heterogeneity 149
3.10 Relative Dividend Prediction Error 150

List of Tables

1.1 Monthly Stock Level Summary Statistics 60

1.2 Summary Statistics: High-Frequency Stock Pair Level Correlations 60

1.3 Stock-Day Level Summary Statistics: Trading and Quoting by Institution 61

1.4 Comparison of Treated and Control Stocks? 62

1.5 The Effects of Designated Market Makers on Liquidity, Volume, and Asset Prices . . 63

1.6 The Effects of Designated Market Makers on the Volatilities of Liquidity, Volume, and Asset Prices 64

1.7 The Separate Effects of Liquidity Provision Obligations and Incentives on Liquidity, Volume, and Asset Prices 65

1.8 How Do Market Makers Set Quotes? 66

1.9 The Effects of Common Intermediation on Co-Movement 67

1.10 Details of the 2018 Market Maker Program 68

1.10 Details of the 2018 Market Maker Program (continued) 69

1.10 Details of the 2018 Market Maker Program (continued) 70

1.11 Details of the 2019 Market Maker Program 71

1.11 Details of the 2019 Market Maker Program (continued) 72

1.11 Details of the 2019 Market Maker Program (continued) 73

1.11 Details of the 2019 Market Maker Program (continued) 74

1.11 Details of the 2019 Market Maker Program (continued) 75

1.12 The Strategic Complementarity of Liquidity Provision: The Effects of Designated Market Makers on Other Traders' Liquidity Provision 76

2.1 Summary Statistics 109
2.2 Do Mutual Funds Hold Stocks with Correlated Characteristics? 110
2.3 Regressions of Quarterly Returns on (Surprise) Flow-Induced Trading 110
2.4 Regressions of Quarterly Returns on (Surprise) Flow-Induced Trading 111
2.5 Predicting (Surprise) Flow-Induced Trading 112
2.6 Predicting Returns with (Surprise) Flow-Induced Trading 113

3.1 Summary Statistics for Financial Institutions by Type 151
3.2 The Market Feedback Effect on Investment 152
3.3 The Market Feedback Effect on Investment – Instrument with *Expected* DIT 152
3.4 Capital Structure Rebalancing Effect 153
3.5 Firms Subject to Dividend-Induced Trading are Similar to Dividend-Paying Firms . 153
3.6 The Market Feedback Effect on Investment – Robustness Tests 154
3.7 The Market Feedback Effect on Investment – Alternative Clustering 154
3.8 Capital Structure Rebalancing Effect – Time x Industry Fixed Effects 155
3.9 Capital Structure Rebalancing Effect – Instrument with *Expected* DIT 155

Chapter 1

In Good Times and in Bad: High-Frequency Market Making Design, Liquidity, and Asset Prices

1.1 Introduction

A key purpose of financial markets is liquidity provision. The intermediaries that supply liquidity are market makers (MMs), which are specialized firms that stand ready to both buy and sell. Today, market making is dominated by high-frequency traders (HFTs). Exchanges and regulators are concerned that these voluntary liquidity providers can simply exit during crises, causing liquidity to evaporate when traders need it the most, as for example during the 2010 Flash Crash (e.g. Kirilenko et al., 2017). As a solution, major exchanges including the NYSE, Euronext, London Stock Exchange, Deutsche Börse, and Brazil's B3 enroll designated market makers (DMMs) that accept liquidity provision obligations in return for incentives such as preferential trading fees. Since the 2018 legislative framework MiFID II, EU regulators even require exchanges to enter DMM agreements with voluntary market makers.

However, market making design is challenging: obligations guarantee a minimum of liquidity provision but also constrain DMMs; incentives encourage voluntary liquidity provision but also harmful strategies; and despite their potentially large consequences for the functioning of financial markets, the separate effects of the elements of these policies and consequently the optimal policy design are largely unknown.

This paper tackles this challenge by empirically investigating how exchanges and regulators can improve the liquidity and stability of modern financial markets. In particular, should exchanges and regulators use liquidity provision obligations or incentives? Does the answer differ in good times and crises? How can they prevent liquidity dry-ups? And how should they assign MMs to assets?

We address these questions using unique full-depth message-level trade and quote data from the Brazilian exchange B3, one of the largest exchanges in the world with a total stock market capitalization of 1 trillion USD. The key advantage of these data is that they reveal market participants' identities. The data cover the universe of activity on B3: all trades and quotes in all assets from 2014 to 2020 timestamped to the millisecond. In total, the raw message data include 50 billion observations. While our analysis is set in Brazil, the MMs include the most sophisticated algorithmic traders in the world, such as Citadel Securities and Virtu Financial.

For identification, we use two market maker programs that provide abrupt exogenous variation in MM activity as well as liquidity provision obligations and incentives. In 2018 and 2019, the exchange implemented centralized programs assigning 2 to 5 out of 13 global high-frequency MMs to each included stock. The programs impose stock-specific liquidity provision obligations: DMMs must maintain continuous, two-sided quotes within maximum bid-ask spreads and at minimum lot sizes. In return, DMMs receive incentives to voluntarily provide liquidity: B3 grants DMMs stock-specific trading fee discounts on *all* trades in their designated stocks. B3 decided program stock selection, obligations, and incentives two months in advance and quasi-randomly conditional on few observable variables.

We exploit the various features of the MM programs using several difference-in-differences designs which identify the effects of the overall programs, and separately estimate the effects of liquidity provision obligations and incentives, in normal times and in crisis. We examine effects on

the behavior of market makers as well as the level, volatility, and co-movement of liquidity, volume, and returns. While the programs provide random variation in many program elements, the identifying assumption underlying our research design is not random assignment; it is that outcomes would have trended in parallel absent the programs. Consistent with this assumption, we find the relevant outcomes for treated and control groups evolve in parallel before the shocks.

We begin by examining how DMMs behave during normal times. We find each DMM simultaneously runs two strategies for each asset. The first strategy only serves to fulfill their obligations. We call this strategy the "mandatory activity." We can isolate the mandatory activity because the programs require DMMs to post very large and uncommon quantities, which is a special advantage of our setting. The strategy is very active; e.g., Citadel Securities sends about one program-quantity message per second per stock. But as the quantity requirement exposes DMMs to substantial execution risk, this strategy quotes wide spreads and consequently generates little volume: total mandatory passive trades from all DMMs combined amount to only 1% of trading volume. Simultaneously, DMMs run a second strategy that we call the "voluntary activity." As a reward for fulfilling their obligations, DMMs receive a fee discount for all trading in their assigned stocks. This gives them a competitive advantage over all other liquidity providers, an advantage they exploit with a strategy that quotes smaller quantities but at aggressive prices which attracts large volume: the voluntary activity intermediates 30% of total volume. The voluntary activity also crosses the spread, but more often provides liquidity.

Next, we study how the programs affect market outcomes. We start by investigating the programs' overall effects using a difference-in-differences methodology comparing stocks that were newly included in the MM programs to those that were not. The combination of obligations and incentives is overall beneficial for stocks. Liquidity improves: quoted and effective spreads decline by up to 17%, and depth increases by 50%. In particular, the fee discount fully passes through into lower spreads. We find this large liquidity improvement is partly driven by strategic complementarity in liquidity provision. In addition to the DMMs, other traders that do not receive the fee discount also tighten spreads and provide more depth after the programs start. The liquidity improvement attracts volume, which rises by 15%, and lifts asset prices by 4%. In addition to the

programs' effects on levels, we examine their effects on stability. We find DMMs drastically lower the volatilities of liquidity and volume, but there is no substantial effect on return volatility. DMMs stabilize liquidity, which also stabilizes volume because trading demand depends on liquidity. Overall, the programs are win-win-win-win-lose situations: investors and stocks benefit from improved liquidity, the increase in volume overcompensates the exchange for the fee discounts, DMMs profit from the fee discount, and MMs that were not selected as DMMs lose business.

Our next step is to disentangle the effects of liquidity provision obligations and incentives. To this end, we exploit abrupt changes in stock-specific obligations and incentives from the 2018 to the 2019 program. We find tightening liquidity provision requirements, intended to improve market quality, has the opposite effect. Tight obligations decrease liquidity, volume, and asset prices, likely by exposing MMs to higher execution risk which lowers inventory capacity and inhibits MMs' ability to voluntarily provide liquidity. Incentives to voluntarily provide liquidity, by contrast, are beneficial. Together, these results show that in normal times, the positive effects are driven by the incentives, not the obligations.

However, these results flip in crises. Specifically, we examine the COVID-19 market crash during which stock prices fall by over 40% and liquidity dries up: within days and at the 95th percentile, program stocks' quoted spreads surge seven-fold from 0.15% to 1%. We find that during the crash, the voluntary activity withdraws and the mandatory activity becomes the liquidity provider of last resort. As market conditions worsen, liquidity provision by the voluntary strategy exhibits an abrupt absolute drop of 8% of total volume, a change that fully reverses as market conditions normalize. At the exact same time, the share of volume intermediated by the mandatory strategy quintuples from 1% to 5% and drops back to 1% as market conditions normalize.

We investigate whether this MM behavior affects stocks' resilience in crises by exploiting variation generated by the 2019 market maker program. First, we document that tighter program requirements mitigate the liquidity dry-up during the crash. MM obligations act as insurance, costly in normal times but beneficial in crises. Second, we find that larger incentives, i.e. more reliance on voluntary HFT liquidity provision, exacerbate the liquidity dry-up because the voluntary activity withdraws. Finally, we find the program overall mitigates the liquidity dry-up. In

sum, the programs are beneficial during both normal times and crises, but the effects of obligations and incentives are inverted. This suggests that exchanges and regulators should combine incentives with countercyclical liquidity provision obligations.

Another important choice exchanges make is how to assign MMs. Here, we document a dark side of the programs: MMs hedge across assets causing excess co-movement of liquidity, volume, and returns. Hence, by choosing which MMs to assign to which stocks, exchanges create co-movement clusters. We begin by examining how MMs behave. In practice, MMs act as intermediaries in several assets, making joint portfolio and liquidity supply decisions (Easley et al., 2020). Indeed, theoretical models of market making with multiple assets predict that MMs' demand for one stock depends negatively on inventory of the stock itself and, crucially, negatively on inventory of other stocks as well (Andrade, Chang, and Seasholes, 2008). We test whether MMs' quoting behavior is consistent with this prediction. In particular, we test how MMs revise price quotes in response to inventory shocks in the stock itself and in other stocks. Concretely, we examine MMs' quoting behavior in event time which preserves the temporal ordering of market events and MMs' quote revisions and only consider MMs' reactions in excess of any market reaction by controlling for stock returns.

We find MMs revise quotes aggressively after their passive orders are hit. Consider a MM who does not currently hold inventory; in response to accommodating a market sell order, the MM decreases quotes and widens spreads for that asset and for other assets. This cross-asset quoting behavior gives rise to a spillover effect through the portfolios of MMs; and as the shock moves all assets intermediated by the MM in the same direction, this mechanism can cause nonfundamental co-movement.

We estimate the causal effect of common intermediation on co-movement by exploiting exogenous variation in market making activity created by the two MM programs. Our difference-in-differences estimator compares stock pairs that are assigned a common MM to stock pairs that are assigned disjoint sets of MMs. We find common intermediation raises co-movement across all three dimensions: liquidity, volume, and returns. The increases are economically meaningful, up to 20% of the mean. Overall, there is a trade-off: DMMs boost and stabilize liquidity but cause excess

co-movement.

Our results are robust to a battery of additional tests. Most importantly, we estimate all difference-in-differences models separately for the 2018 and the 2019 programs. For all 15 variables we consider, levels, volatilities, and correlations of liquidity (measured as quoted spreads, effective spreads, and depth), volume, and returns, the results are consistent across the two natural experiments, which gives us high confidence in the robustness of the findings.

Related Literature. This paper relates to the literature studying the effects of DMMs. The literature finds DMMs improve market quality and benefit asset prices (Nimalendran and Petrella, 2003; Venkataraman and Waisburd, 2007; Anand, Tanggaard, and Weaver, 2009; Comerton-Forde et al., 2010; Perotti and Rindi, 2010; Menkveld and Wang, 2013; Skjeltorp and Odegaard, 2015; Clark-Joseph, Ye, and Zi, 2017; Bellia et al., 2020; Bessembinder, Hao, and Zheng, 2020). In contrast to the literature, we observe DMM behavior at the message level, as well as the exact DMM contract parameters. Our setting allows us to separate DMMs' mandatory and voluntary activity and disentangle their effects, in normal times and crises. These differences are policy relevant because exchanges and regulators can choose obligations and incentives.

We also contribute to the extensive literature exploring the behavior and effects of HFTs (e.g. Brogaard, Hendershott, and Riordan, 2014; Menkveld, 2016; Gider, Schmickler, and Westheide, 2019; Kervel and Menkveld, 2019; Korajczyk and Murphy, 2019). Specifically, recent work documents that HFTs often withdraw during crises (Anand and Venkataraman, 2016; Kirilenko et al., 2017; Brogaard et al., 2018; Raman, Robe, and Yadav, 2021). We contribute to this literature by showing that this behavior and its adverse effects can be mitigated by imposing affirmative obligations on HFTs. Another contribution of this paper is that HFTs hedge across assets which generates excess co-movement.

Next, this paper is related to the vast literature on the economics of market making (e.g. Stoll, 1978; Amihud and Mendelson, 1980; Hasbrouck and Sofianos, 1993; Madhavan and Smidt, 1993; Comerton-Forde et al., 2010; Hendershott and Menkveld, 2014). While existing findings are typically based on trade data from NYSE specialists, we observe algorithmic MMs' behavior at the

message level. Moreover, research in market microstructure predominantly examines markets one asset at a time, yet market participants make decisions across assets, especially with the rise of algorithmic trading (Easley et al., 2020). Our analysis documents that, consistent with theoretical models (e.g. Andrade, Chang, and Seasholes, 2008), MMs hedge across assets and cause excess co-movement.

This paper also relates to the extensive literature that examines how financial institutions impact asset prices and market quality (e.g. Brunnermeier and Pedersen, 2009; He and Krishnamurthy, 2013; Adrian, Etula, and Muir, 2014; He, Kelly, and Manela, 2017; Koijen and Yogo, 2019). In particular, this paper contributes to the literature that connects intermediaries to commonality in asset returns and liquidity. Greenwood and Thesmar (2011) and Anton and Polk (2014) find common ownership by mutual funds is associated with higher return co-movement. Barberis, Shleifer, and Wurgler (2005) and Claessens and Yafeh (2013) document index inclusion increases stocks' correlation with said index, though Chen, Singal, and Whitelaw (2016) cast doubt on the result. While this literature shows how shocks to demand-side investors create low-frequency co-movement, our paper studies how sell-side intermediaries cause high-frequency spillover effects. Also related are papers that investigate cross-stock predictability (Hasbrouck and Seppi, 2001; Harford and Kaul, 2005; Andrade, Chang, and Seasholes, 2008; Tookes, 2008; Pasquariello and Vega, 2015).

Relatedly, Chordia, Roll, and Subrahmanyam (2000) document important commonality in liquidity. Building on their work, Coughenour and Saad (2004) and Anand and Venkataraman (2016) show commonality in liquidity between stocks handled by the same liquidity provider. Similarly, Malceniece, Malcenieks, and Putniņš (2019) find HFTs increase return and liquidity market betas. Complementing these supply-side results, Karolyi, Lee, and van Dijk (2012) and Koch, Ruenzi, and Starks (2016) provide evidence that commonality in liquidity is driven by correlated liquidity demand. We contribute to this literature by showing that cross-asset hedging by MMs causes excess co-movement in liquidity, volume, and returns.

1.2 Data and Empirical Strategy

We use message data for all listed assets traded on the Brazilian Stock Exchange (B3). Raw messages include every order submitted and track lifetime order status in a market-by-order fashion. B3 marks every order with a buy/sell indicator, order type, quantity, bid or offer price, and aggressor flag and then records all changes (e.g., partial fulfillment, cancellation, modification, rejection, etc.) sequentially. As our audit trail data record orders at all price levels, the raw messages allow us to reconstruct the full depth limit order book of every market at every moment. Messages are timestamped to the millisecond. Our sample spans 2014 to 2020. In total, the raw message data include about 50 billion observations.

A special feature of our data is that they contain unique entering firm IDs. These identifiers are assigned by B3 to exchange participants allowed to operate in the market and remain constant across time, assets, and firm lifespan. If a market participant stops trading, e.g. because of an acquisition, B3 freezes the identifier and never uses the same code again. The exchange provides the identities of all registered firms since 1980, which enables us to observe the actual firm behind the masked ID and hence, all orders and trades done by the firm in stocks, options, and futures.

1.2.1 Institutional Details

B3 – Brasil, Bolsa, Balcão, formerly BM&FBovespa – is the only exchange in Brazil and resulted from of a merger between the country's last exchanges in 2008. In 2017, then BM&FBovespa took over CETIP, Brazil's largest depository, clearinghouse, and over-the-counter (OTC) market. This acquisition made B3 the largest OTC market in the country. Today, B3's services include trading, clearing, and settlement of equities and derivatives as well as registration, clearing, settlement, and custody systems for OTC fixed-income securities and OTC derivatives. According to the CME, this wide range of product offerings makes B3 the world's third largest exchange and the largest in Latin America by total market value across all its products (stocks, options, futures, OTC fixed income, OTC derivatives). In terms of its total stock market capitalization of 1 trillion USD, B3 ranks about 20th in the world, similar to the main stock exchanges in Germany, Switzerland, Australia,

and Singapore. B3 is a public company, trading in its own market with a market cap of about 15 billion USD, putting it in line with the world's top stock exchanges such as Nasdaq and Deutsche Börse.

B3 launched electronic limit order book platforms in the early 1990s, around the same time CME introduced its electronic trading platform Globex. Trading at B3 became fully electronic in 2009, when the last derivatives pits were shut down. B3's electronic trading platform is very similar to interfaces used at US exchanges, with market data feeds following standard FIX message protocols optimized for low latency. This allows B3 to route orders directly to CME, for example, which cross-lists several products with B3. Order priority allocation follows price, time, and quantity. For the same price level, orders entered first are entirely filled first (FIFO matching). The regular trading session starts at 10:00am and closes at 4:55pm. Stocks are traded at one-centavo increments and in 100-share minimum lot sizes. Odd lots are submitted and traded in a separate, parallel market – named "fractional" stocks by B3 – where investors may submit orders for fewer than 100 shares. We observe fractional stock trading but do not use it in this paper.

1.2.2 The Natural Experiments

We exploit B3's 2018 and 2019 MM programs as sources of exogenous variation in MM activity, obligations, and incentives. We obtain all program details directly from B3.

Timeline. In 2018 and 2019, B3 implemented centralized programs with standardized contract parameters to enroll MMs in stocks starting on one set date. MMs commit to maintaining minimum quantities at maximum spreads in exchange for fee discounts. Figure 1.1 provides a timeline of the two programs. The first program was announced in August 2018. Financial institutions could apply to the program until September 10. On September 12, B3 made and publicly released all program decisions. The program began on October 15, 2018 and ended on November 1, 2019. In September 2019, B3 announced a second program. Financial institutions had to apply by September 20 and B3 made and announced all program decisions on September 27. The program started on November 2, 2019, immediately following the end of the first program, and ended on April 30, 2021. For the

2018 and the 2019 programs, B3 chose program stocks and parameters two months and DMMs one month in advance. This means that B3's decisions are not based on changes occurring after the program start.

Program Inclusion. The 2018 program included 54 out of about 400 stocks while the 2019 program included 89. For the 2018 program, B3 selected program stocks quasi-randomly conditional on three observable variables. First, B3 selected stocks from the exchange's IBrX 100 index, which tracks the 100 top stocks based on average daily volume. Second, B3 excluded a small number of top stocks because of an internal notion that these stocks were already sufficiently liquid. Third, B3 gave preference to stocks without pre-existing bilateral DMM contracts (B3 has allowed issuing companies to enter bilateral agreements with trading firms for designated market making since 2003. About half of IBrX 100 stocks had entered these bilateral agreements which were unaffected by the market maker programs). While B3 did not use a random number generator to decide program inclusion, B3 explains the selection process as manual conditional stratified randomization; decision makers at B3 aimed to maximize variety across included stocks. While we confirm this empirically by showing that conditional on the three variables other stock characteristics are unrelated to treatment status, our identification strategy does not rely on random assignment. For the 2019 program, B3 increased the program's scope to all except the top IBrX 100 stocks.

DMM Assignment. The 2018 (2019) program divided included stocks into 2 (3) groups based on a liquidity cutoff. Group A included the 9 (13) stocks with historical average daily volume above R$80 (R$100) million. Group B included stocks below the cutoff, with the exception that the 2019 program further assigned stocks below R$30 million to group C. B3 offered two DMM slots for each stock in group A, three slots for group B, and five slots for group C. The 2018 volume cutoff reused an old value from a 2015 policy. B3 allowed trading firms to apply for stocks selectively. However, trading firms had to apply for five stocks in group B or C in order to submit one bid for a stock in group A. B3 had full discretion in choosing MMs and aimed to maximize variety across stocks; i.e. when given the choice between two trading firms, the exchange selected the firm with the lowest total number of slots. The 2018 program filled all DMM slots for each stock in the program. The

2019 program filled 96% of slots. The 2018 (2019) program selected 9 (13) trading firms as DMMs. The DMMs are mostly HFTs and include some of the most sophisticated algorithmic traders in the world, such as Citadel Securities, Virtu Financial, and Jane Street.

Activity Parameters. The programs set stock-specific minimum activity parameters. Specifically, the programs require DMMs to maintain continuous, two-sided quotes within maximum bid-ask spreads and at minimum lot sizes during at least 80% of each trading day. Moreover, DMMs must participate in at least 5% of daily trading volume. B3 set stock-specific parameters by applying mechanical rules to data covering the 12-month period ending three months before the programs. The rules aim to balance liquidity provision and execution risk. B3 set minimum quantities to about 0.3% of past average daily shares traded rounded to 1000s and maximum spreads to the past average round trip cost of each stock at the minimum quantity rounded up to the next 0.25%. Firms could propose stricter parameters, but final contract parameters remained unchanged from the values initially proposed by B3. Appendix Tables 1.10 and 1.11 detail the designated MMs and program parameters for each stock. Appendix Figure 1.13 shows histograms of program parameters.

Benefits and Enforcement. B3 fully waived trading fees for MMs in their designated stocks. The only exceptions are group A stocks in the 2019 program, for which MMs received a 75% instead of a 100% discount. B3 is a monopoly and consequently charges high fees. The fees range from 1.1 to 2.3 basis points. The more a market participant trades, the lower the fee. On average, B3 collects 1.85 basis points from each side of the trade. As market makers trade large volumes, they paid close to 1.1 basis points per transactions before the programs. For comparison, major stock exchanges typically charge less than 0.5 basis points. This means the fee discount is a very large competitive advantage and consequently, a very large shock. B3 evaluated DMMs after the first six months of each program, at which point B3 could sanction or replace firms that failed to fulfill their obligations. However, all MMs complied and by the end of the programs, no MMs had been substituted.

1.2.3 Observing Market Makers' Trades and Quotes

Figure 1.2 illustrates how traders access the Brazilian financial market. B3 gives direct market access to about 100 executing firms. End investors submit orders through executing firms. Importantly, executing firms merely transmit orders to the exchange; they do not trade with clients as in a dealer market. For example, Goldman Sachs executes orders for institutional clients and XP for retail investors. While we are working with B3 to obtain account IDs, our current data reveal the executing firm that sends the message to B3. However, while some MMs (highlighted in red) are executing firms (e.g. Credit Suisse), MMs and executing firms are not synonymous. Notably, foreign MMs are required to trade via domestic execution brokers. B3 reveals to us which execution brokers MMs use. MMs only use one execution broker for each stock. MMs may use different execution brokers for different stocks, but in practice they use one or only a few executing firms. E.g., Virtu Financial submits orders via UBS and QE Trading submits orders via Genial Institutional.

Execution brokers have multiple clients. Hence, while most activity originating from an executing firm that executes for a MM does stem from the MM because of MMs' high activity level, messages from other clients contaminate this inference. However, the programs' unique structure helps us identify messages from MMs. The programs require MMs to maintain high and unusual minimum quantities. While 100 is the most common (and required minimum) order quantity for most stocks, the programs set minimum quantities from 2,000 to 60,000 shares. MMs would not voluntarily provide such depth, which is why B3 incentivizes them. Therefore, MMs do not quote above the minimum quantity, meaning they quote *exactly* at the minimum quantity. As the minimum quantities are very uncommon lot sizes, messages at the minimum quantity from execution brokers serving MMs are quotes from MMs. Specifically, they are messages sent by MMs to fulfill their program obligations. We therefore call these messages their "mandatory activity".

While B3 confirms this logic, we also verify it empirically. Figure 1.3 shows the distribution of order quantities in 1000s for the first stock alphabetically, the large brewing company Ambev, on October 1, before the 2018 program begins, and on October 15th, the program's first day. The minimum quantity for Ambev is 25,000 shares. Before the shock, there are virtually no quotes

at 25,000 shares. In fact, most quotes are for 1,000 shares or fewer. However, after the shock, a substantial fraction of messages enter at exactly 25,000 shares, while the rest of the distribution remains almost unchanged. This bunching at exactly 25,000 shares confirms that MMs quote exactly at minimum quantities. This pattern is not unique to Ambev, as shown in section 1.3.1, where we compare MMs' mandatory and voluntary activity.

We measure MMs' voluntary activity as quotes by MMs' execution brokers excluding messages at the minimum quantity. This measure contains MMs' voluntary activity, though it also contains some messages from other clients. However, as there are about 100 execution brokers and MMs are very dominant in the stocks they intermediate (e.g., about half of all messages come from MMs), the contamination is likely small. We are also working with B3 to obtain account instead of executing firm IDs, meaning our future inference will be exact.

1.2.4 Comparison to Alternative Data Sources

Trade and quote data from order-driven markets almost never reveal traders' identities. The fact that researchers have proxied trading by the same entity as trades that occur almost simultaneously (e.g. within 20 microseconds) in different products (e.g. Ernst, 2021) shows the difficulty of obtaining identified message data. This makes our ability to observe MM behavior at the message level an important feature of this paper.

Previously, researchers have gained insights based on other data sources. While we cannot list all papers that have used trader identities, several stand out as important. First, Madhavan and Smidt (1991) and Madhavan and Smidt (1993) obtain transactions directly from one NYSE specialist firm. Hasbrouck and Sofianos (1993) use the NYSE specialist equity trade (SPET) file which covers NYSE specialist trades. More commonly, researchers use the NYSE Specialist Equity Trade Summary (SPETS) file, which provides daily data on specialists' trades (e.g. Madhavan and Sofianos, 1998; Hendershott and Seasholes, 2007; Comerton-Forde et al., 2010; Hendershott and Menkveld, 2014). Further, NYSE Consolidated Equity Audit Trail Data (CAUD) are transaction-level data that contain a flag indicating whether the buyer or seller are institutional investors (e.g. Boulatov, Hendershott, and Livdan, 2013; Hendershott, Livdan, and Schürhoff, 2015). Similarly,

NASDAQ, as well as the Trade Capture Report from the CME to the CFTC, supply transaction-level data that identify whether the buyer or seller are high-frequency traders (HFTs) (e.g. Brogaard, Hendershott, and Riordan, 2014; Kirilenko et al., 2017; Brogaard et al., 2018).

Next, the Investment Industry Regulatory Organization of Canada (IIROC) provides message-level data containing anonymous market participant IDs. Brogaard, Hendershott, and Riordan (2019) and Korajczyk and Murphy (2019) use the IDs to classify whether market participants are HFTs. Likewise, Aquilina, Budish, and O'Neill (2021) rely on message-level data with anonymous market participant IDs through a request of the Financial Conduct Authority (FCA) to the London stock exchange. Kervel and Menkveld (2019) and Meling (2021) obtain transaction-level data from Sweden and Norway, respectively, that reveal the executing firm that sends the order.

While our data share similarities with each of these data sources, our data have several advantages: first, they are at the message level, which is the most granular; second, they reveal identities; third, at least for the mandatory activity, they identify the market markers themselves, not just the executing firms which often execute trades for many clients; fourth, they cover algorithmic MMs instead of specialists (of course, each advantage only applies to a subset of the data sources listed above).

1.2.5 Variable Construction

In addition to the raw message data, we use other datasets. All MM program details are from B3, which also provides a mapping between anonymous market participant IDs and their identities as well as information on which market participant each MM uses as an execution broker in each stock. Finally, we obtain additional stock and accounting data from Compustat Global Fundamentals Annual and Compustat Global Security Daily.

We use the raw message data to reconstruct the complete limit order book at each instant, thereby allowing us to recover best bids and offers at that precise time. We use the data in this form to investigate MMs' high-frequency behavior at the message level. Further, we retain complete limit order book snapshots at the 5-minute frequency, which is the most common intraday frequency (e.g. Marshall, Nguyen, and Visaltanachoti, 2012). We also compute daily and monthly frequency

data by aggregating from the 5-minute frequency. This ensures that our low-frequency measures capture intraday instead of end-of-day information.

We construct two key input datasets: stock x time data and trader x stock x time data. For the former, we construct several key variables. Brunnermeier and Pedersen (2009) define liquidity as the difference between transaction prices and fundamental values. While fundamental values are unobserved, Comerton-Forde et al. (2010) argue that effective spreads best capture the spirit of their model. Therefore, we measure liquidity as effective spreads and additionally examine quoted spreads and depth, all constructed following Marshall, Nguyen, and Visaltanachoti (2012). Effective spreads are volume-weighted mean 2*abs((trade price - midpoint) / midpoint) during the respective time period. Quoted bid-ask spreads are end-of-period (best offer - best bid) / midpoint. We construct three measures of depth: depth on the first five price levels, total depth, and depth within 0.5% of the midpoint. We choose the latter as our main measure of depth because we believe it best captures liquidity, but the results are qualitatively the same across the different measures of depth. We compute R$ depth by multiplying depth in number of shares with the open price. This prevents intraday volatilities and correlations of depth being driven by stock returns. Depth and volume are in log R$.

Returns are relative changes in end-of-period bid-ask midpoints. We winsorize returns cross-sectionally at the 1 and 99% levels. Next, we compute standard deviations and stock pair level correlation coefficients of log quoted spreads, log effective spreads, log depth, log volume, and returns by month. Quoted spreads, effective spreads, depth, volume, and returns in levels, standard deviations, and correlation coefficients form the main input data into our empirical exercises at the stock x time level. In all exercises, we exclude microcap stocks, which are defined as stocks with a market capitalization less than R$100M.

Next, we construct trader x stock x time data. We compute the number of trades, total number of shares traded, and total R$ value traded separately for buys vs. sells and for liquidity-taking vs. liquidity-providing trades. We also retain prices and quantities of end-of-period best active bids and offers. Finally, we count the number of messages independently for each message type. Importantly, we split MMs' mandatory and voluntary activity. To this end, we use the information

provided by B3 to link market participant IDs to their identities. Then, we map these identities to the MMs for which they transmit orders. Finally, we separate the mandatory and voluntary activities by whether a message is sent at the program quantity, as described in section 1.2.3.

1.2.6 Summary Statistics

Table 1.3 gives an overview of the 13 MMs in our sample. The table provides stock-day level summary statistics by MM, sorted alphabetically. Panel A summarizes their mandatory activity, panel B summarizes their voluntary activity. At the time, 1 USD buys about 4 R$s. We begin by looking at MMs' mandatory activity. Average stock-day trading volume ranges from R$300k for Credit Suisse to R$1.5M for Jane Street. The average trade size is large, around R$60k, meaning each MM's mandatory activity only trades about a dozen times a day in each stock. However, MMs send a large amount of messages. Citadel Securities sends the most messages, at 19,000 messages per stock per day, i.e. about 1 message per stock per second, and 1,700 messages per trade. Finally, by definition, the mandatory activity trades only passively.

We also examine MMs quoting styles by reporting the fraction of messages that are new orders, fills, cancels, and replace orders. Two distinct quoting styles emerge. Citadel Securities is a good example of the first style. Their share of new and cancel orders is 50% and their share of fill and replace messages is 0% (rounded to the first decimal place). Credit Suisse is a good example of the second style. Their share of new, fill, and cancel messages is close to 0 and their share of replace orders is close to 100%.

Panel B reports analogous statistics for the voluntary activity. Here, average stock-day trading volume is much larger, ranging from 12M for Green Post Trading to 52M for Headlands Technologies. At the same time, average trade size is much lower, typically less than R$10k. This means the voluntary activity trades much more often, thousands of times per stock per day. While the voluntary activity also sends an enormous amount of messages, from 8k to 95k messages per stock-day, the message to trade ratio is much lower, ranging from 3 to 25. Further, the voluntary activity often takes liquidity. The most aggressive institutions are Citadel Securities and Tucana Bay, which are the aggressor in almost 60% of their volume. Finally, new, fill, cancel, and replace orders account

for about 40, 15, 35, and 10% of messages, respectively.

Table 1.1 reports summary statistics of monthly stock characteristics for the sample spanning the nine-month event windows around the MM programs. There are about 3600 observations, about 200 stocks during 18 months. The stocks in the sample have a mean market capitalization of R$12B. From the 5th to the 95th percentiles, stock prices range from R$2 to R$69. The natural experiments fall into a good period for stock prices. The average monthly return is 3%. The Brazilian stock market is very liquid: at the 5th percentile, quoted bid-ask spreads are only 5 basis points. Even at the median, bid-ask spreads are only 12 basis points. Only the very right side of the distribution is illiquid, with a 1.5% quoted spread at the 95th percentile. Effective spreads are similar to quoted spreads. Average depth is R$700k on the first five price levels, R$1.3M within 0.5% of the midpoint, and R$5M in the full-depth book. Average daily volume is R$50M, which increases to R$180M at the 95th percentile and to over R$2B for the most traded stock, the oil firm Petrobras. The average daily number of trades is 5400 and the average trade size is R$8400.

Next, we summarize the distributions of intraday (5-minute) standard deviations of log quoted spreads, log effective spreads, log depth, log volume, and returns. This differs from the standard deviations of the respective variables above because it is computed within instead of across stocks. Liquidity is volatile: quoted and effective spreads have a mean standard deviation of about 50% and depth of 80%. Further, volume is volatile with a mean standard deviation of 300%. Finally, mean intraday return volatility is 20 basis points.

Finally, we summarize correlations between program stocks. In contrast to the stock-level variables discussed above, correlations are *stock-pair-level* variables. This is why there are many more, 170k, observations. Table 1.2 reports the corresponding summary statistics. The variables are percent correlation coefficients of intraday (5-minute) log quoted spreads, log effective spreads, log depth, log volume, and returns. First, liquidity is correlated across stocks, with means of 17, 11, and 51% for quoted spreads, effective spreads, and depth. So is volume, as the mean correlation is 34%. Finally, returns have a mean correlation of 7%. Idiosyncratic volatility dominates at high-frequency.

1.3 Results

1.3.1 Liquidity Provision Obligations and Incentives in Normal Times

Market Maker Behavior

We begin by examining how DMMs behave. We find each DMM runs two strategies for the same asset at the same time. The "mandatory activity" fulfills the DMM obligations, and the "voluntary activity" exploits the fee discount to quote aggressive prices. Plots (a) and (b) of Figure 1.4 examine the voluntary and mandatory activity, respectively, in event time during a ± four-month event window around the programs. The red line in plot (a) shows the number of messages from the voluntary activity as a percent fraction of all messages sent by other executing firms for stocks included in the program. The line starts at 40% because B3 selected trading firms that were already making markets to be DMMs. Then, with the start of the program, this fraction jumps to 70%. The gray and black lines report trading instead of quoting activity. With the start of the program, the share of trades in which the voluntary activity participates jumps from 45 to 65% and its share of volume jumps from 35 to 50%. That the share of trades is greater than the share of volume means the average trade size is low, as reported by the summary statistics in Tables 1.3 and 1.1. Overall, plot (a) demonstrates that fee discounts incentivize MMs to increase their voluntary activity.

Plot (b) displays the analogous results for the mandatory activity. Quoting activity explodes. Two months before the program begins, the fraction is visually indistinguishable from 0. Only about 0.01% of messages are sent at the minimum quantity. This fraction increases to about 1% in the pre-treatment period, when MMs run live tests. When the program starts, this fraction jumps to 40% and then climbs to 60% over three months. In addition to demonstrating how active the mandatory strategy is, this finding also further demonstrates the validity of inferring the mandatory activity from minimum quantities (in combination with execution brokers) as described in section 1.2.3. In contrast, the share of trades and volume in which the mandatory activity participates stays close to zero. The mandatory activity participates in 1% of volume and 0.2% of trades. The fact that the share of volume is greater than the share of trades means the average trade size is high, as reported by the summary statistics in Tables 1.3 and 1.1. This is because the minimum

quantities are large.

Finally, plot (c) separates the voluntary trading volume into active and passive trades. Passive volume increases from 20 to 30%, active volume rises from 15 to 20%. This shows that DMMs exploit their competitive advantage to intermediate large volume. However, they also trade actively, either to hedge or to take directional bets. Note that there is no analogue of plot (c) for the mandatory activity, which, by definition, only trades passively.

Next, Figure 1.5 digs deeper into DMMs' behavior by examining prices instead of quantities. It displays the fractions of time the voluntary and mandatory activity spend on different price levels. The red line in plot (a) shows that upon the programs' start, the share of time that the voluntary activity provides the best price in the market increases from 20 to 35%. This demonstrates that DMMs exploit their competitive advantage coming from the fee discount by quoting aggressive prices. The gray line denotes that the share of time during which the voluntary activity quotes on price levels 2 to 5 only rises slightly, from 15 to 17%. Plot (b) reports the analogous results for the mandatory activity, which usually quotes unattractive prices, almost never quoting at the top of the book.

In sum, Figure 1.6 illustrates DMMs' typical quoting behavior using an example snapshot of the bid-side of a limit order book. Price levels decrease from the top to the bottom and queue priority declines from left to right. We highlight orders from the mandatory and voluntary activity in red and blue, respectively. The figure illustrates that the mandatory activity quotes large quantities deep in the book while the voluntary activity quotes small quantities at the top of the book. Appendix Figure 1.14 contains a dynamic instead of a static example.

Designated Market Makers' Overall Effects

We begin by investigating the effects of combining liquidity provision obligations and incentives. To this end, we exploit exogenous variation in market making activity created by the MM programs using a difference-in-differences identification strategy to estimate DMMs' casual effects on the levels and volatilities of liquidity, volume, and asset prices. For the first program, we compare stocks included in the MM program to stocks that are not part of the program. We use the

2019 program to compare stocks newly included in the program to stocks included since the 2018 program. Hence, the treated stocks in the 2018 experiment are the control stocks in the 2019 experiment. We estimate DMMs' causal effects on stock-level outcomes separately for the 2018 and 2019 program using

$$y_{i,t} = \alpha_i + \phi_t + \beta MM_i \times Post_t + \delta' X_{i,t} + \epsilon_{i,t}, \tag{1.1}$$

where MM indicates that stock i is newly included in the MM program. X denotes a vector of control variables. When estimating this regression, we cluster standard errors by stock. The 2019 program starts almost 4 months before the COVID crisis begins. As this section examines the effects during normal times, we set the event windows to $\pm$ 4 months around the events and cut off the sample at the start of the COVID-19 stock market crash on February 21, 2020.

The validity of our difference-in-differences identification strategy relies on the assumption that outcomes for the treatment and control groups would have trended in parallel without the treatment. This assumption likely holds because B3 selected program stocks quasi-randomly conditional on being IBrX 100 index constituents, not being one of a few top stocks, and not having a pre-existing bilaterally contracted DMM. Table 1.4 confirms this selection procedure using regressions of the treatment indicator on an indicator for inclusion in the IBrX 100 index, an indicator for a pre-existing bilateral DMM relationship, and key stock characteristics. The characteristics include key outcome variables of our analysis; the Fama and French (2015) characteristics size, value, profitability, and investment; and major additional firm characteristics. As expected, the coefficient on the IBrX 100 indicator is positive and statistically significant for both programs; the coefficient on the indicator for the bilateral DMM relationship is negative and significant for the 2018 experiment and positive and significant for the 2019 experiment; the top stock indicator is statistically significant and negative for the 2018 experiment and missing for the 2019 experiment because the top stocks are not in the sample; and importantly, the coefficients on all other characteristics are small and statistically insignificant. Motivated by this empirical fact, we choose the three variables B3

used to determine treatment as the set of control variables. The fact that treated and control firms have similar stock characteristics, at lest conditional on the control variables, also suggests that the parallel trends assumption is likely fulfilled. We further document the absence of differential pre-trends when examining dynamic treatment effects.

DMMs' Effects on Liquidity, Volume, and Asset Prices. Table 1.5 reports estimates of equation 1.1 for the two natural experiments separately. The rows report results for different dependent variables: log quoted spreads, log effective spreads, log depth, log volume, and cumulative abnormal returns. We begin by examining the 2018 experiment in panel A. First, stocks in the program experience a statistically significant 10% decline in quoted bid-ask spreads. Similarly, effective spreads drop by 10%. Next, row 3 finds a 50% increase in depth. These three effects demonstrate that DMMs increase liquidity. Row 4 shows the liquidity improvement attracts volume, which increases by 15%. To interpret the size of this effect, recall that Figure 1.4 documents that the programs also increase the share of volume in which MMs participate by 15%. As MMs only take one side of each trade, this means that the programs increase trading volume even after stripping out the trading volume of MMs.

Finally, we examine pricing effects. Stock prices react when the relevant information is revealed; therefore, in all exercises involving price reactions, we examine returns around the announcement, not the start, of the program. We find program stocks experience positive cumulative abnormal returns of 3%. Panel B shows that the results are very similar for the 2019 experiment. In fact, the confidence intervals overlap in most cases. This further corroborates the robustness of the findings.

We also estimate dynamic treatment effects using

$$y_{i,t} = \alpha_i + \phi_t + \sum_{s \neq 0} \beta_s MM_{i,t} \times \mathbb{1}\{s = t - E + 1\} + \delta' X_{i,t} + \epsilon_{i,t}, \tag{1.2}$$

where E is the time of the event and the β_s are the dynamic treatment effects. Figure 1.7 reports estimates of equation 1.2, pooling the two program samples. Plots (a) and (b) trace out the effects

on quoted and effective spreads, respectively. The pattern is the same in both plots. Treated and control stocks trend in parallel before the shock. When the shock begins, program stock spreads fall sharply relative to control stocks. After a few months, there seems to be a partial reversal, though the confidence intervals overlap for all post-treatment effects. Plot (c) reports effects on depth. Treated and control stocks trend in parallel from period -4 to -1. In period 0, depth increases for treated stocks. This is expected because MMs run live tests before the program starts. Next, when the program starts, depth rises sharply and remains elevated. Plots (d) and (e) show the effect on volume and asset prices. Again, there are no differential pre-event trends, but volume and asset prices increase after the shock. Overall, DMMs are beneficial for stocks. They boost liquidity, which attracts volume and lifts asset prices.

The fee discount fully passes through into lower spreads: the spread decrease is about the same as the fee discount for a roundtrip transaction; 2 basis points (a 10% decline of a pre-treatment mean of the treated group of 20 basis points) in comparison to saving the 1 basis point trading fee in two transactions. We find this large liquidity improvement is partly driven by strategic complementarity in liquidity provision. In particular, we examine how other liquidity providers respond to the MM program by constructing hypothetical limit order books that ignore DMMs. We use the hypothetical books to compute quoted spreads and depth provided by non-DMMs. Appendix Table 1.12 reports the results. We find non-DMMs tighten spreads and provide more depth after the programs start even though they do not receive the fee discount.

Designated Market Makers' Effects on Volatility. In addition to the programs' effects on levels, we examine their effects on stability. In particular, we examine the volatilities of liquidity, volume, and returns. Table 1.6 reports the results for the two natural experiments separately. The rows report results for volatilities of different dependent variables: intraday (5-minute) quoted spreads, effective spreads, depth, volume, and returns. The columns differ by including different sets of fixed effects. We begin by examining the 2018 experiment in panel A.

First, rows 1 to 3 present the results for percent volatilities of liquidity measures. For the volatility of quoted spreads, we find a negative and statistically significant coefficient of -6 across all

specifications, a large effect amounting to a third of a standard deviation. Similarly, we find that the volatilities of effective spreads and depth also decline. DMMs stabilize liquidity. Next, row 4 shows a negative and statistically significant coefficient of -50 for the volatility of volume, a large effect amounting to 70% of a standard deviation. This is consistent with the decreased volatility of liquidity. The demand for trading depends on its cost. And as the cost of trading stabilizes, so does volume. Finally, the effect on return volatility is positive but small and statistically insignificant. Panel B demonstrates the 2019 experiment yields very similar results. In fact, the confidence intervals overlap in most cases, which further supports the robustness of these findings.

Next, we examine the effects' dynamics by estimating equation 1.2. Figure 1.8 reports the results. Plots (a) to (d) display the results for the intraday volatilities of quoted spreads, effective spreads, depth, and volume. The patterns are very similar across the four graphs. There are no differential pre-treatment trends, followed by a sharp, concave drop when the event begins. The effect plateaus quickly and does not reverse. Plot (e) shows that there is no statistically significant response of intraday return volatility. Overall, DMMs stabilize liquidity and volume.

Effects of Liquidity Provision Obligations and Incentives

The MM programs combine the two major elements exchanges use to improve market quality: liquidity provision obligations and incentives. Which of these two elements is responsible for the improvement in market quality? We answer this question by exploiting abrupt variation in obligations and incentives generated by the MM programs. First, B3 changes stock-specific maximum spreads and minimum quantities from the 2018 to the 2019 program. For some stocks, maximum spreads decrease (increase) and/or minimum quantities increase (decrease), i.e. requirements tighten (loosen). For other stocks, the program parameters remain unchanged. Second, in an attempt to boost fee revenue, B3 lowers the fee discount from 100% to 75% for the 13 most traded program stocks, a negative treatment reducing the incentive for voluntary liquidity provision. We exploit each change using a difference-in-differences identification strategy based on equation 1.1.

Table 1.7 reports the results. Panel A displays the effects of tightening DMM requirements. We estimate equation 1.1 for the following shock: stocks for which B3 tightens requirements are

positively treated; stocks for which B3 loosens requirements are negatively treated; stocks for which B3 does not change the requirements or changes the maximum spread and minimum quantity in the same direction (i.e. tightens one and loosens the other) form the control group. We find increases in quoted and effective spreads, decreases in depth and volume, and negative cumulative abnormal returns. The effects are statistically significant for spreads and volume, but not for returns and depth. Tight obligations, intended to improve market quality, have the opposite effect, likely because they expose MMs to higher execution risk which lowers inventory capacity and inhibits their ability to voluntarily provide liquidity.

Panel B displays the effects of the incentive to voluntarily provide liquidity. Again, we estimate equation 1.1, now for the following shock: stocks for which B3 reduces the fee discount are *negatively* treated, and the remaining program stocks form the control group. We find the opposite effects from panel A, i.e. incentives are beneficial: fee discounts lower quoted and effective spreads and increase depth, volume, and asset prices. Overall, this shows that in normal times the positive effects are driven by the incentives, not the obligations.

1.3.2 Liquidity Provision Obligations and Incentives During Crises

Thus far we have examined effects during normal times. However, crises are particularly important to exchanges and regulators because liquidity often evaporates during crises which can be detrimental to welfare because crises are when investors have the largest and most immediate trading needs. Therefore, we now explore MM behavior as well as market outcomes during the COVID-19 stock market crash. Figure 1.9 demonstrates the severity of the crash in our setting. Panel (a) shows cumulative returns on the equal-weighted program stock portfolio. The downturn begins on February 21, 2020. Stock prices fall by 10% until March 4, when asset prices crash, falling by 40% until they reach the bottom on March 23. After that, they partially recover.

Panel (c) shows how the crisis affected liquidity. We plot the time series of the median and 95th percentile quoted spreads of program stocks. We also plot the fraction of time for which the program maximum spread is binding. Before the crisis, program stocks are very liquid. Median spreads are 0.1%, 95th percentile spreads are 0.15%, and the maximum spread never binds. Then,

at the beginning of the crisis, the market looks resilient, only displaying modestly increased spreads. However, on March 12, the day of the global "Black Thursday" crash, liquidity suddenly dries up: median quoted spreads of program stocks quadruple from 0.1 to 0.4% and at the 95th percentile, quoted spreads surge seven-fold from 0.15% to 1%. In normal times, the maximum spreads imposed by the MM programs never bind, but on March 12, maximum spreads are binding 20% of the time. After that, liquidity remains scarce but improves slowly. Finally, panel (c) examines trading volume. Despite the liquidity dry-up, average daily volume rises from R$60M to R$80M, highlighting how desperate investors are to trade.

Market Maker Behavior

As before, we first investigate MMs' behavior. We find that voluntary liquidity providers withdraw and mandatory intermediaries become the liquidity providers of last resort. Figure 1.10 plots the time series of the liquidity provision shares of the mandatory and voluntary activities, i.e. passive trading volume of the respective activity scaled by total trading volume. Before the shock, the voluntary activity is responsible for a large share, 40%, of liquidity provision. However, during the crash this share abruptly drops to 32%, reaches its minimum almost exactly at the height of the crisis, and fully reverses when market conditions normalize. The mandatory activity follows the opposite pattern. It starts very low, at less than 1% of liquidity provision. Then, at the exact same time as the voluntary activity withdraws, the average share of liquidity provision by the mandatory activity quintuples to 5%, peaking very close to the crisis' worst point, after which it drops back to pre-crises levels as market conditions normalize.

Effects on Stocks' Resilience in Crises

In parallel to the preceding analysis of normal times, we now examine the effects of the overall program, obligations, and incentives during the crisis. We exploit variation across stocks generated by the 2019 MM program to investigate whether MMs' crisis behavior affects market outcomes. Concretely, we estimate dynamic treatment effects using equation 1.2, where the time of the shock is the start of the crash, the 21st February of 2020, and the treatment indicator isolates different

stock subgroups depending on the effect of interest. Causal inference is more difficult in this setting than in the preceding analysis of normal times because the MM programs start during normal times and are already in place when the crisis begins. One particular concern is that the subgroups we compare differ in additional dimensions that may be correlated with their sensitivity to crashes. We address this concern by comparing subgroups that are as similar as possible and by controlling for time-specific effects of known differences.

Figure 1.11 reports the results. Most importantly, we study the effects of affirmative liquidity provision obligations during crises by comparing stocks with tight to stocks with loose obligations. While the exchange sets maximum spreads based on historic liquidity levels, we exploit that B3 rounds maximum spreads to the next 0.25%, which generates quasi-random variation in liquidity provision obligations. Picture two nearly identical stocks that would have been assigned maximum spreads of 0.5% and 0.51%. The former would have a relatively tight maximum spread at 0.5%, while the latter would have been rounded up to a relatively loose maximum spread of 0.75%. Specifically, we compare stocks in the top and bottom halves of the rounding error and control for historic pre-rounding spread $\times$ time heterogeneous slopes. Plots (a), (b), and (c) report the estimated effects of tighter obligations. We find tight obligations mitigate the liquidity dry-up: quoted spreads of stocks with tight obligations spike by 0.1% less during the crash, a large effect of 100% of the pre-crisis median. The effects on volume and returns are small.

Next, the 2019 program generates variation in voluntary liquidity provision incentives by granting DMMs 100% fee discounts in most program stocks and 75% fee discounts in the 13 program stocks for which historic average daily trading volume exceeded R$100M. While this cutoff is arbitrary, it does mean treated (100% fee discount) and control (75% fee discount) stocks differ according to historic trading volume. We address this by controlling for historic trading volume $\times$ time heterogeneous slopes. Plots (d), (e), and (f) report the results. We find that stocks with larger incentives experience worse liquidity dry-ups: quoted spreads of stocks with higher incentives spike by 0.1% more during the crash, a large effect of 100% of the pre-crisis median. Consistently, volume also declines. Again, there are no effects on returns. These results indicate that strong reliance on voluntary HFT liquidity provision is harmful during crises because these liquidity providers

withdraw.

Finally, we examine the program's overall effects by comparing program stocks to the most similar non-program stocks, i.e. stocks right outside the IBrX 100 cutoff, ranked 101 to 150. The only reason the control stocks are not part of the program is that they are not part of the index, i.e. because their historic trading volume fell just short of an arbitrary cutoff. This means treated and control stocks differ according to the metric used for index inclusion. We address this by controlling for index metric $\times$ time heterogeneous slopes. Further, the fact that treated and control stocks differ by IBrX 100 index inclusion raises the concern that they differ along additional dimensions such as their investor base. However, attenuating this concern, the IBrX 100 is not a top index. The most important index in the Brazilian stock market is the Ibovespa, and the largest Brazil ETF (EWZ) tracks the MSCI Brazil 25/50 Index, both of which track substantially fewer stocks than the IBrX 100. Panels (g), (h), and (i) report the program's estimated effects on liquidity (quoted spreads), volume, and returns. Before the crash, treated and control stocks trend in parallel, but when the crisis begins, program stocks experience a significantly less severe liquidity dry-up. The effects on volume and returns are small.

While the program is overall beneficial in both normal times and crises, the effects of obligations and incentives invert. In normal times, exchanges can increase market quality by encouraging voluntary liquidity provision. By contrast, imposing tight liquidity provision requirements constrains MMs and can decrease market quality. During crises, however, tight obligations are beneficial while strong reliance on HFT liquidity is harmful. This suggests that exchanges and regulators should combine incentives with countercyclical liquidity provision obligations. Of course, it may be difficult to dynamically adjust obligations. If obligations have to be static, our results suggest that obligations should be set such that they do not bind during normal times but become binding during crises, as is the case for the program we examine in this paper.

1.3.3 How to Assign Market Makers to Assets?

Another important choice exchanges have to make when designing liquidity provision programs is how to assign MMs. Here, we document a dark side of these programs: MMs hedge across assets

which causes excess co-movement of liquidity, volume, and returns. Hence, by choosing which MMs to assign to which stocks, exchanges create co-movement clusters.

Market Maker Behavior

We begin by examining how MMs behave. In practice, MMs act as intermediaries in several assets, making joint portfolio and liquidity supply decisions (Easley et al., 2020). Indeed, theoretical models of market making with multiple assets predict that MMs' demand for one stock depends negatively on inventory of the stock itself and, crucially, negatively on inventory of other stocks as well (Andrade, Chang, and Seasholes, 2008). We test whether MMs' quoting behavior is consistent with this prediction. Specifically, we test how MMs revise price quotes in response to inventory shocks in the stock itself and in other stocks.

Our approach addresses two challenges. First, our specification allows us to focus on MMs' reactions in excess of market reactions by controlling for stock returns. Second, we address a reverse causality problem: algorithmic MMs respond to market events in fractions of a second. Therefore, market events and quote revisions happen simultaneously in data synchronized to regular time periods. But the causality may run both ways: MMs respond to market events and MM quotes attract volume. Our high-frequency, message-level data allow us to address this reverse causality problem by examining MMs' quoting behavior in event time, which preserves the temporal ordering of market events and quote revisions.

Concretely, we conduct this analysis by MM $\times$ stock $\times$ trading day for computational reasons and treat the bid and ask side of the market separately. We denote the liquidity provider by i, the stock by n, and the day by d. There are two event time clocks. The first clock is only needed for variable definitions; every trade is an event and event time is denoted by $\tilde{t}$. The second clock is the frequency at which we estimate MM behavior; each active change in a MM's best standing quote starts a new period and event time is denoted by t. We only start a new event period after *active* changes to avoid the mechanical effects of passive execution on standing quotes. We denote the set of stocks in which i is a MM as $\mathcal{I}_i$. $\mathcal{T}$ is the set of all trades. A trade is described by the tuple $(i, n, \tilde{t})$.

The MM may react to any market event that occurred since the last quote update, including market events in other stocks. Therefore, we construct explanatory variables that summarize all trades that occur between event time $t-1$ and event time t in all stocks intermediated by MM i. We denote the set of these trades by

$$\mathcal{M}_{i,t} = \{(j,m,s) \in \mathcal{T} : t-1 < s < t, m \in \mathcal{I}_i\}. \tag{1.3}$$

Depending on the stock under consideration, we divide this set into four disjoint subsets

$$\begin{aligned}
\mathcal{M}^{OF}_{i,n,t} &= \{(j,m,s) \in \mathcal{M}_{i,t}, j \neq i, m = n\} \\
\mathcal{M}^{OFabsorbed}_{i,n,t} &= \{(j,m,s) \in \mathcal{M}_{i,t}, j = i, m = n\} \\
\mathcal{M}^{OFother}_{i,n,t} &= \{(j,m,s) \in \mathcal{M}_{i,t}, j \neq i, m \neq n\} \\
\mathcal{M}^{OFabsorbedOther}_{i,n,t} &= \{(j,m,s) \in \mathcal{M}_{i,t}, j = i, m \neq n\}.
\end{aligned} \tag{1.4}$$

The OF and $OFabsorbed$ subsets only contain trades in the stock itself, the $OFother$ and $OFabsorbedOther$ subsets only contain trades in other stocks. The $OFabsorbed$ and $OFabsorbedOther$ subsets only contain trades where the MM was the liquidity provider, the OF and $OFother$ subsets only contain trades where the MM was not the liquidity provider. There are four main dependent variables, one summarizing each subset. These four variables are all defined in the same way by

$$x_{i,n,t} = \frac{1}{\sum_{\{m \in \mathcal{M}^x_{i,n,t}\}} R\$\overline{Vol}_m} \sum_{(j,m,s) \in \mathcal{M}^x_{i,n,t}} OrderSide_{j,m,s} R\$Vol_{j,m,s}, \tag{1.5}$$

where $x \in \{OF, OFabsorbed, OFother, OFabsorbedOther\}$. We normalize by dividing by past average daily R$ volume, $R\$\overline{Vol}$, summed over all stocks in the respective set. $OrderSide$ is 1 if the buyer is the aggressor, -1 if the seller is the aggressor. $R\$Vol$ is the R$ value of the trade. The

only difference across the four variables is which trades are included in the summation, the $\mathcal{M}^x$. OF is order flow intermediated by other liquidity providers. $OFabsorbed$ is order flow absorbed by the MM (i.e. changes in inventory). $OFother$ is order flow in other stocks intermediated by other liquidity providers. $OFabsorbedOther$ is order flow absorbed by the MM in other stocks (i.e. changes in inventory of other stocks). We use these variables to estimate

$$\begin{aligned}\Delta p_{i,n,t} = \alpha_{i,n,d} + \beta_{1,i,n,d} OF_{i,n,t} + \beta_{2,i,n,d} OFabsorbed_{i,n,t} \\ + \beta_{3,i,n,d} OFother_{i,n,t} + \beta_{4,i,n,d} OFabsorbedOther_{i,n,t} + \beta_{5,i,n,d} r_{n,t} + \epsilon_{i,n,t},\end{aligned} \tag{1.6}$$

where p denotes log best price quotes of the MM. We treat the two sides of the market separately: p is the best bid of the MM in one dataset and the best ask in another dataset. For estimation, we then append the two datasets which imposes symmetry. We control for log stock returns, r, the analogue of the dependent variable at the stock level instead of the MM level. This makes this specification particularly stringent because we only attribute quote revisions unexplained by market movements to the independent variables of interest.

The coefficient estimates are subscripted because we estimate equation 1.6 separately for each MM x stock x day combination. This is conceptually equivalent to including MM x stock x day fixed effects. We report mean coefficients and compute Fama-MacBeth standard errors robust to within MM x day correlation from the distribution of the coefficient estimates (Fama and MacBeth, 1973). This accounts for correlations across stocks as well as for autocorrelation within stocks during the trading session. Our approach is broadly similar to Brogaard, Hendershott, and Riordan (2019).

β_1 measures the relationship between MM quotes and order flow. The similar relationship between stock returns and order flow is well-studied and known to be positive (e.g. Hasbrouck and Sofianos, 1993; Madhavan and Smidt, 1993). Here, we control for stock returns and examine how MMs respond to order flow in excess of market movements. β_2 measures how MMs react to order flow when they were the liquidity providers. Researchers have examined the relationship between quote revisions and specialist trades (e.g. Hasbrouck and Sofianos, 1993; Madhavan and Smidt,

1993) but not at this level of granularity because market data almost never reveal trader identities. Existing findings are typically based on daily trade data (e.g. Comerton-Forde et al., 2010). In contrast, our data allow us to examine MM behavior at high frequency. Finally, β_3 and β_4 measure the analogous effects across stocks. Research in market microstructure usually examines markets one asset at a time. Yet, market participants make trading decisions across assets, even more so with the rise of algorithmic trading (Easley et al., 2020). Our two cross-asset variables take this into account and allow us to study how MMs respond to information and inventory shocks in other stocks.

Table 1.8 reports estimates of equation 1.6. Quote revisions and stock returns are in percent. The sample spans the MM programs' duration. Columns 1 and 2 report results for the mandatory activity, columns 3 and 4 for the voluntary activity. Columns 2 and 4 control for stock returns, columns 1 and 3 do not. The estimated coefficients are qualitatively similar, just smaller once we control for stock returns. All relevant coefficient estimates are statistically significant.

In the main specification for the mandatory activity, column 2, the coefficient on order flow is 0.7, indicating MMs are more sensitive to order flow than the market. Second, the coefficient on absorbed order flow is much larger, 8. This means MMs revise quotes much more aggressively when they are on the other side of the order flow, likely to revert inventory. In response to providing liquidity amounting to 1% of daily volume, MMs revise quotes by 8 basis points in excess of the market response. Third, the coefficient on order flow in other stocks is very close to zero. However, MMs revise quotes when their limit orders in other assets are hit. The coefficient on order flow absorbed in other stocks is much larger, 4. In response to a 1% shock in other stocks, MMs revise quotes by 4 basis points, again in excess of any market response. Finally, the coefficient on returns is 0.9. This simply captures that MMs move their quotes with market prices. The coefficient is close to but less than one because MMs do not always react to market events when quoting deeper in the book. Column 4 reports the analogous results for the voluntary activity. The results are qualitatively the same, except that the coefficients are larger, meaning that the voluntary activity revises quotes more aggressively. This makes sense, because the orders of the voluntary activity are more exposed.

Effects of Common Intermediation on Co-Movement

Market events spill over to other assets through MMs' cross-asset hedging behavior and as the shocks move all assets intermediated by the MM in the same direction, this mechanism can cause nonfundamental co-movement. Using a difference-in-differences identification strategy, we exploit exogenous variation in market making activity created by the MM programs to estimate the causal effect of common intermediation on the co-movement of liquidity, volume, and asset prices. Figure 1.12 illustrates the programs as shocks to common intermediation. The nodes in the graph are the ticker symbols of 10 randomly selected program stocks. Nodes are connected if they have a common MM. The color of the connection indicates when the two stocks were connected. Before the program, there are no connections. The 2018 program introduces the gray and red connections. Then, from the 2018 to the 2019 program, the red connections are severed and the blue connections are added. Hence, there are positive and negative shocks to common intermediation.

We use this shock to estimate the causal effect of common intermediation on co-movement. This analysis is analogous to that described in section 1.3.1, except it is at the stock pair level instead of the stock level. Similar to equation 1.1, we estimate the difference-in-differences regression,

$$\rho_{i,j,t} = \alpha_{i,j} + \phi_t + \beta Connected_{i,j} \times Post_t + \delta' X_{i,j,t} + \epsilon_{i,j,t}, \tag{1.7}$$

separately for each MM program. Importantly, the panel variable is the stock pair, i.e. the combination of stock i and stock j. The dependent variable is the correlation between stock i and stock j. The main explanatory variable is the interaction between a binary variable that indicates whether stocks i and j have a common MM and a post-shock indicator. β is the average treatment effect. In the main specification, we include stock pair and time fixed effects and cluster standard errors by stock pair.

We run this analysis on the stock pair sample in which stock i and stock j are both part of the MM program. Hence, the treatment group consists of stock pairs that were assigned a common MM and the control group consists of stock pairs for which each stock was assigned a DMM, but

their DMM sets are disjoint. This means the results cannot be driven by the direct effects of having a MM.

Table 1.9 reports estimates of equation 1.7 for the two natural experiments separately. The rows report results for different dependent variables: the correlation of intraday (5-minute) quoted spreads, effective spreads, depth, volume, and returns. The columns differ by the inclusion of different sets of fixed effects. We begin by examining the 2018 experiment in panel A. First, for liquidity co-movement, the estimated average treatment effects are positive, and statistically significant across all specifications, for all three measures of liquidity. The treatment effect estimates imply increases between 5 and 20% of the mean. For trading volume, we find a statistically significant effect of 10% of the mean. Finally, for return co-movement, the estimated average treatment effect is positive, statistically significant, and implies that common intermediation raises the correlation coefficient of high-frequency returns by 10% of the mean, a significant increase. Panel B demonstrates the 2019 experiment yields very similar results. In fact, the confidence intervals overlap in most cases, further confirming the robustness of the findings. Overall, common intermediation raises co-movement across all three dimensions: liquidity, volume, and returns.

1.4 Conclusion

We study how exchanges and regulators can improve the liquidity and stability of modern electronic financial markets, in good times and in bad. We address this question using unique message-level trade and quote data from the Brazilian exchange B3 that reveal market participants' identities. Further, we exploit two natural experiments that provide rich and abrupt exogenous variation in market maker (MM) activity as well as liquidity provision obligations and incentives: two MM programs that prescribe stock-specific maximum bid-ask spreads and minimum lot sizes and grant trading fee discounts in return. While the experiments occur in Brazil, the MMs include the most sophisticated algorithmic traders in the world, such as Citadel Securities, Virtu Financial, and Jane Street.

During normal times, each designated market maker (DMM) simultaneously runs two strategies

for each asset. The first strategy, the "mandatory activity," only serves to fulfill their obligations. This strategy quotes wide spreads and consequently generates little volume. Simultaneously, DMMs run a second strategy, the "voluntary activity," which exploits their competitive advantage due to the trading fee discount by quoting aggressive prices and intermediating large volumes.

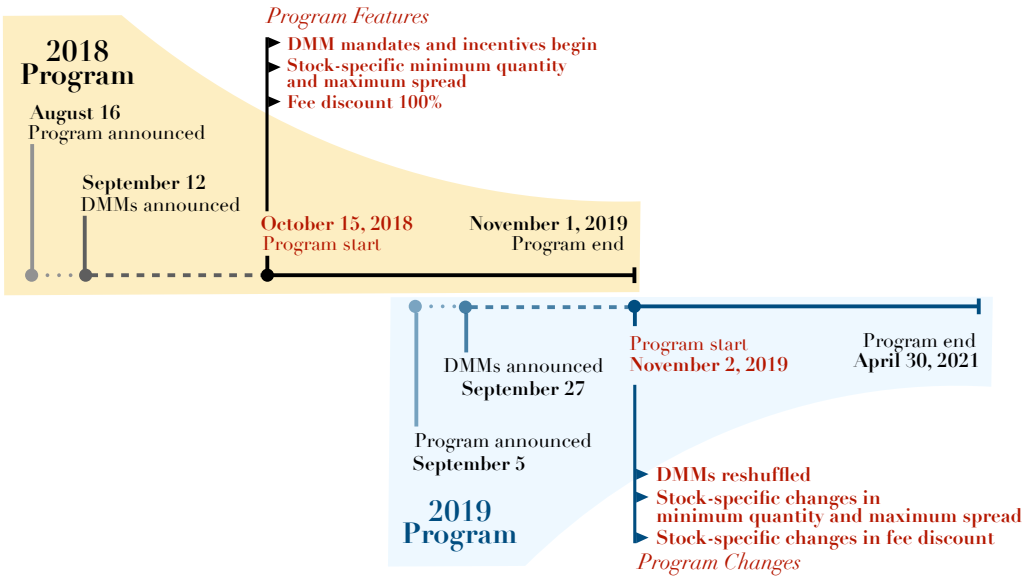
At the stock level, the combination of obligations and incentives improves and stabilizes liquidity which attracts volume and lifts asset prices. In normal times, these positive effects are driven by the program incentives, while tight obligations, intended to improve market quality, have the opposite effect, likely because they expose MMs to higher execution risk which lowers inventory capacity and inhibits MMs' ability to voluntarily provide liquidity.

In crises, however, the results flip: stocks with larger incentives experience worse liquidity dry-ups because the voluntary activity withdraws; in contrast, tight obligations mitigate liquidity dry-ups because the mandatory activity steps in as the liquidity provider of last resort. These results suggest that exchanges and regulators should combine incentives with countercyclical liquidity provision obligations. Of course, it may be difficult to dynamically adjust obligations. If obligations have to be static, our results suggest that obligations should be set such that they do not bind during normal times but become binding during crises.

Finally, which market makers are assigned to which stocks is consequential: consistent with theoretical models of market making, we document empirically that inventory shocks in one asset shift MMs' demand for the asset itself and other stocks they intermediate. This cross-asset hedging behavior of MMs gives rise to a spillover effect through MMs' portfolios that causes nonfundamental co-movement. We confirm this mechanism by documenting that stock pairs that are assigned a common DMM experience increases in the correlations of liquidity, volume, and returns. This shows a trade-off: on the one hand, DMMs increase liquidity, on the other hand, they cause excess co-movement.

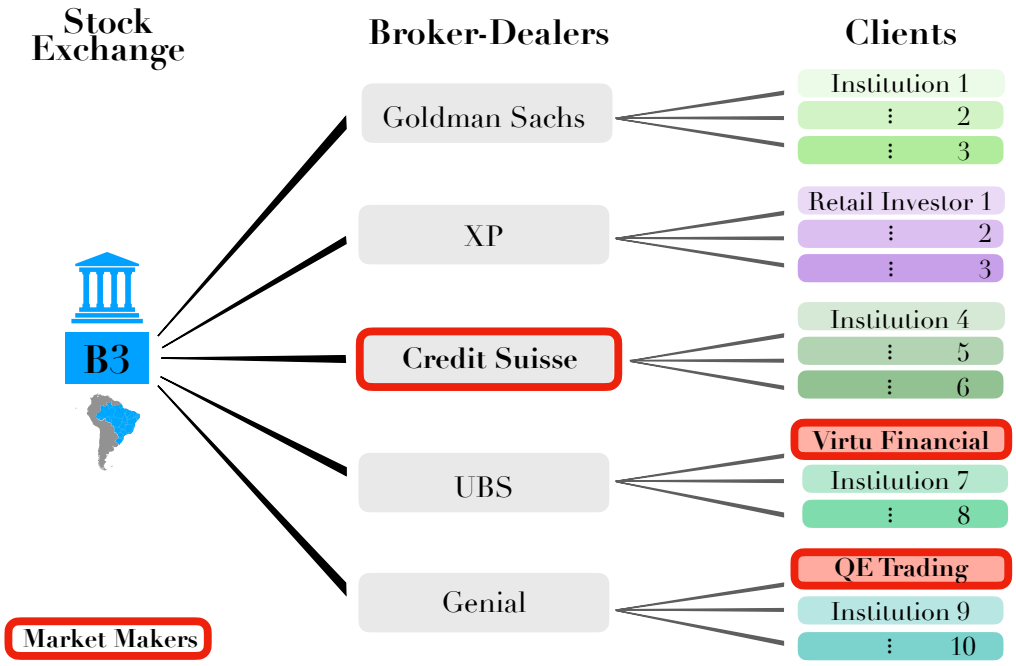
Figures

Figure 1.1: The Timeline of B3's Market Maker Programs



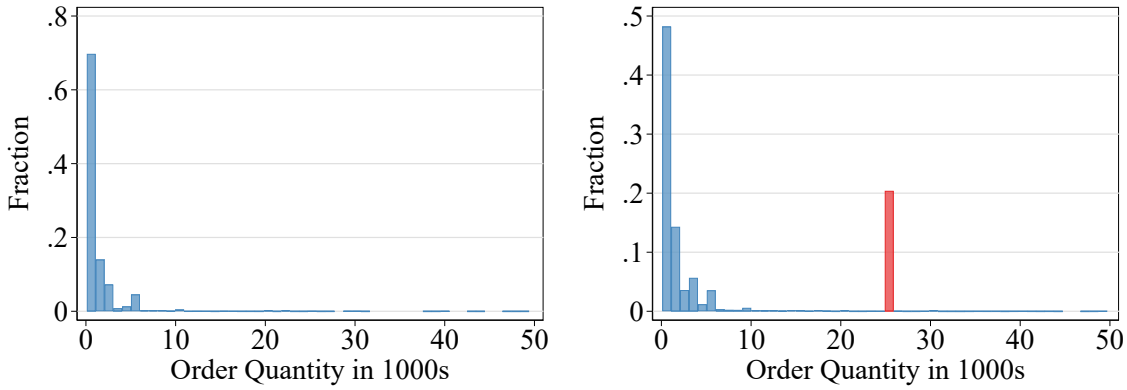
This figure illustrates the timeline of B3's 2018 and 2019 market maker program. B3 publicly announced the 2018 (2019) program on August 16, 2018 (September 5, 2019). Market making firms had until September 10, 2018 (September 20, 2019) to apply for each stock they were interested in. B3 made and publicly released all program decisions on September 12, 2018 (September 27, 2019). The program began on October 15, 2018 (November 2, 2019). It ended on November 1, 2019 (April 30, 2021).

Figure 1.2: Institutional Background



This figure illustrates how traders access Brazilian financial markets. B3 gives direct market access to about 100 broker-dealers. We observe broker-dealer identities in the message data. End investors submit orders through broker-dealers. For example, Goldman Sachs executes orders for financial institutions and XP does so for retail investors. Some broker-dealers are themselves market makers, e.g. Credit Suisse. In addition, Credit Suisse executes orders for institutional clients. Some market makers are not broker-dealers and submit orders via broker-dealers. E.g., Virtu Financial submits orders via UBS and QE Trading submits orders via Genial Institutional. Additionally, UBS and Genial Institutional both also execute orders for other institutional clients.

Figure 1.3: Example of Bunching at the Minimum Quantity (at 25)

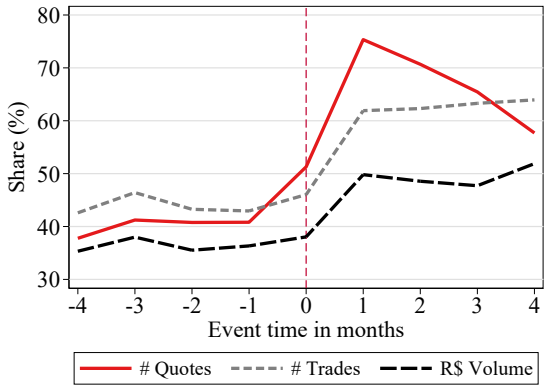


(a) Before the Program

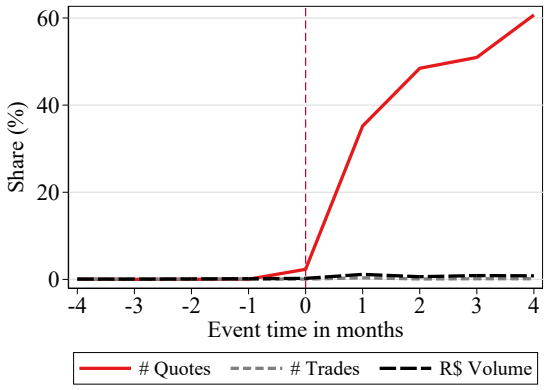
(b) After the Program

This figure illustrates that DMMs quote exactly at the minimum quantity the market maker programs require. The example shows the distribution of order quantities in 1000s for the first stock alphabetically, ABEV3, on October 1st, 2018, before the shock and on October 15th, the first day after the shock. The minimum quantity for ABEV3 is 25,000 shares.

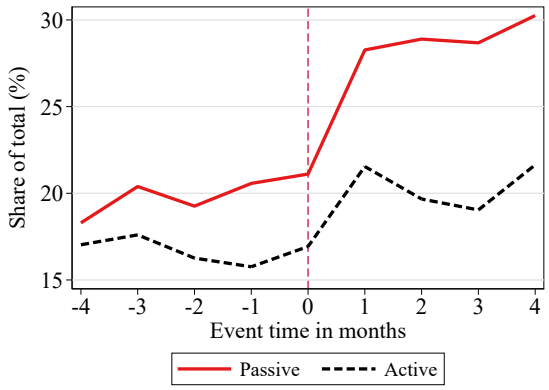
Figure 1.4: Quoting and Trading: Mandatory vs. Voluntary Activity



(a) Voluntary activity



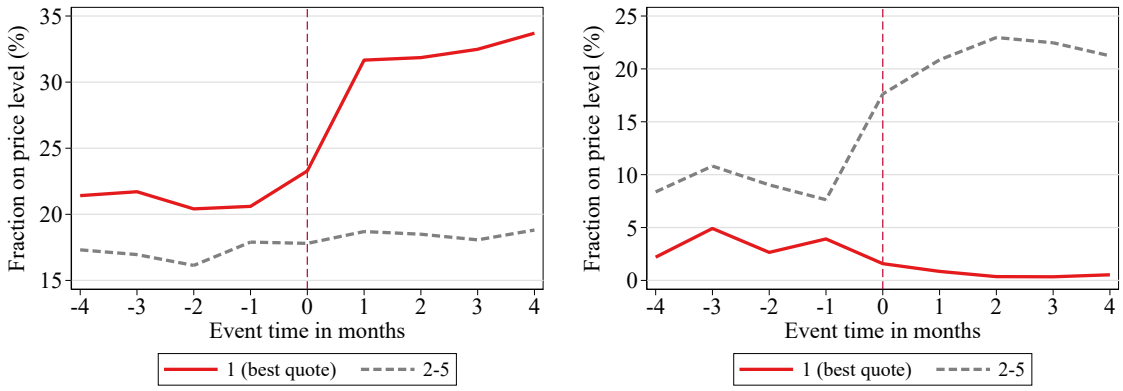
(b) Mandatory activity



(c) Voluntary activity: active vs. passive

This figure compares the quoting and trading of the mandatory and voluntary activity. Panels (a) and (b) plot the number of quotes as the percent fraction of all quotes from other market participants, trading volume as the percent fraction of total volume, and the number of trades as the percent fraction of the total number of trades, all for the voluntary and mandatory activity, respectively. Panel (c) shows total passive and total active trades separately as the percent fraction of total volume. We only show this split for the voluntary activity because, by definition, the mandatory activity only trades passively. Event time is in months around the market maker programs.

Figure 1.5: Fraction of Time On Each Price Level: Mandatory vs. Voluntary Activity

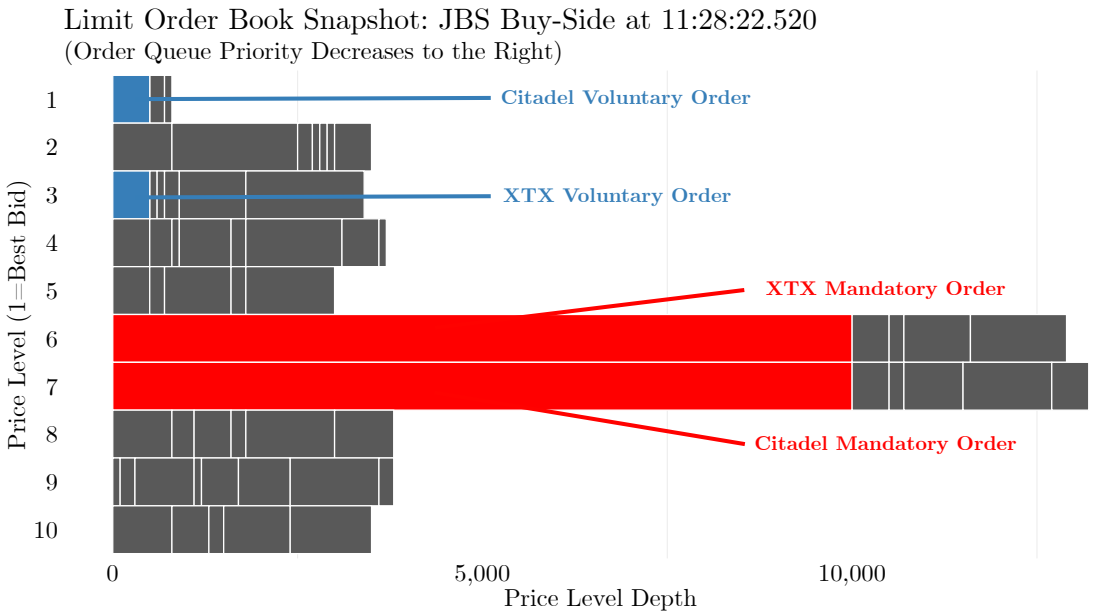


(a) Voluntary activity

(b) Mandatory activity

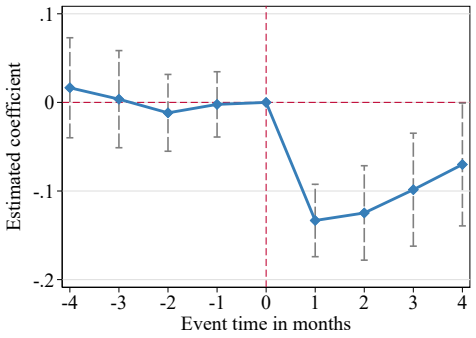
This figure shows the percent fraction of time market makers quote on different price levels. Panel (a) shows results for the voluntary activity, panel (b) for the mandatory activity. The red line corresponds to the best price. The gray line corresponds to price levels 2, 3, 4, or 5. We average across the two sides of the market. Event time is in months around the market maker programs.

Figure 1.6: Illustration of the Mandatory vs. Voluntary Activity

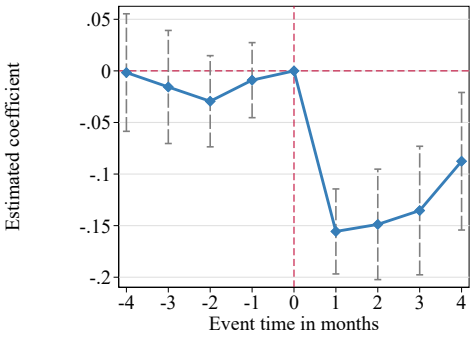


This figure illustrates the typical quoting behavior of the mandatory vs. voluntary activity using an example limit order book snapshot, JBS at 11:28:22.520 on the first day of the 2019 market maker program, 11/04/2019. We only show the first 10 price levels of the bid side. Each block corresponds to one order, with its width indicating its lot size and queue priority decreasing to the right. The market makers in this stock are Citadel Securities and XTX Markets. We highlight their mandatory orders in red and their voluntary orders in blue. The figure shows that, typically, the mandatory orders provide large quantities deep in the book, while the voluntary orders provide small quantities at the top of the book.

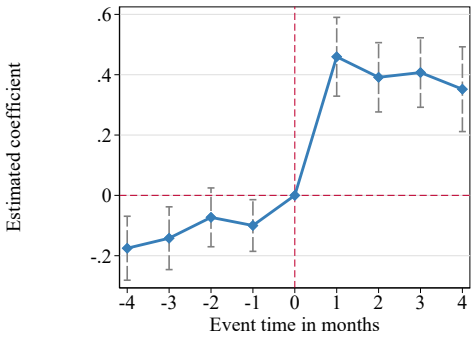
Figure 1.7: Dynamic Treatment Effects of DMMs on Liquidity, Volume, and Asset Prices



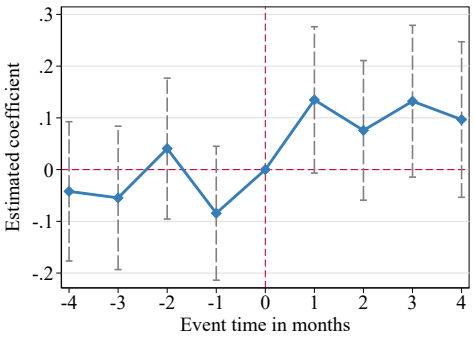
(a) Log Quoted Spread



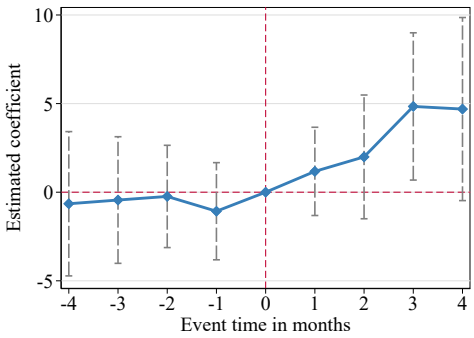
(b) Log Effective Spread



(c) Log Depth



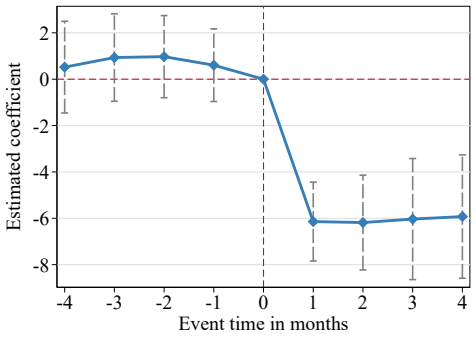
(d) Log Volume



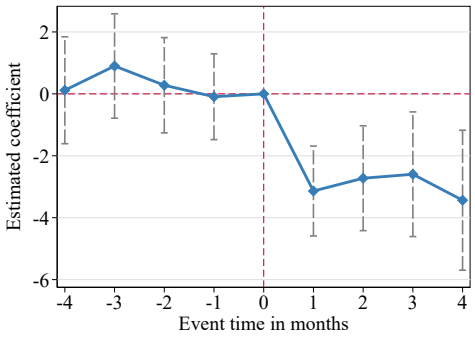
(e) Cumulative Abnormal Returns

This figure reports estimates of equation 1.2 and shows dynamic treatment effects of the market maker assignment on log quoted bid-ask spread, log depth, log R$ volume, and percent cumulative abnormal returns. Standard errors are clustered by stock. The gray lines span 95% confidence intervals. Event time is in months around the market maker programs.

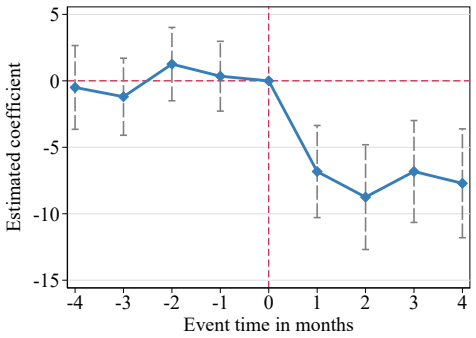
Figure 1.8: Dynamic Treatment Effects of Designated Market Makers on Volatilities



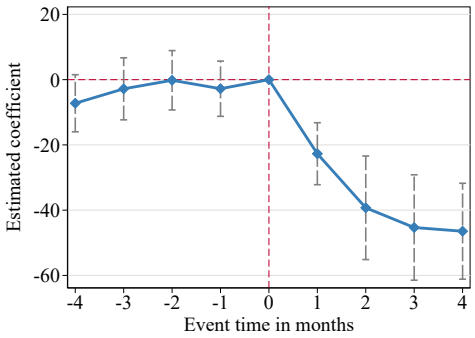
(a) Sd(Log Quoted Spread)



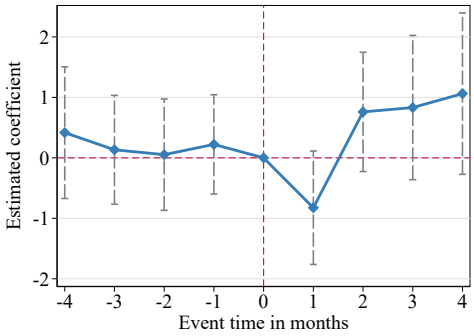
(b) Sd(Log Effective Spread)



(c) Sd(Log Depth)



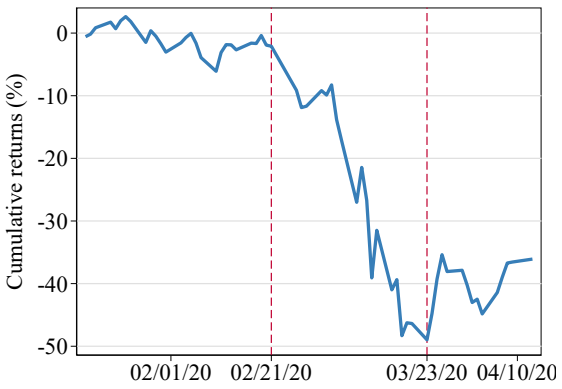
(d) Sd(Log Volume)



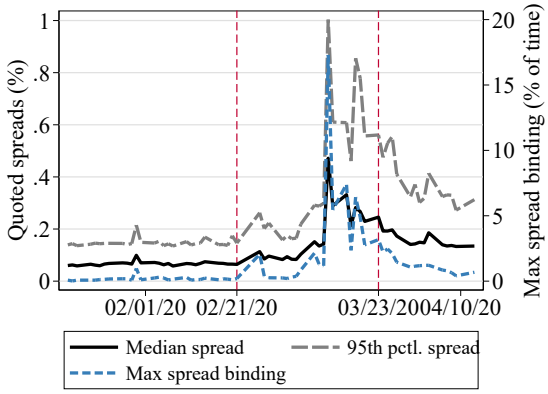
(e) Sd(Returns)

This figure reports estimates of equation 1.2 and shows dynamic treatment effects of the market maker assignment on the percent intraday (5-minute) volatilities of log quoted bid-ask spreads, log depth, log R$ volume, and returns. Standard errors are clustered by stock. The gray lines span 95% confidence intervals. Event time is in months around the market maker programs.

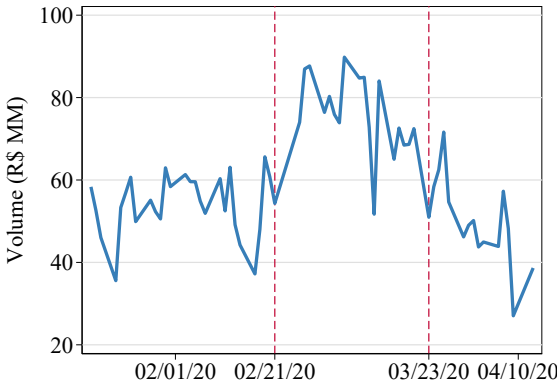
Figure 1.9: The Anatomy of the COVID-19 Crash



(a) Cumulative Returns



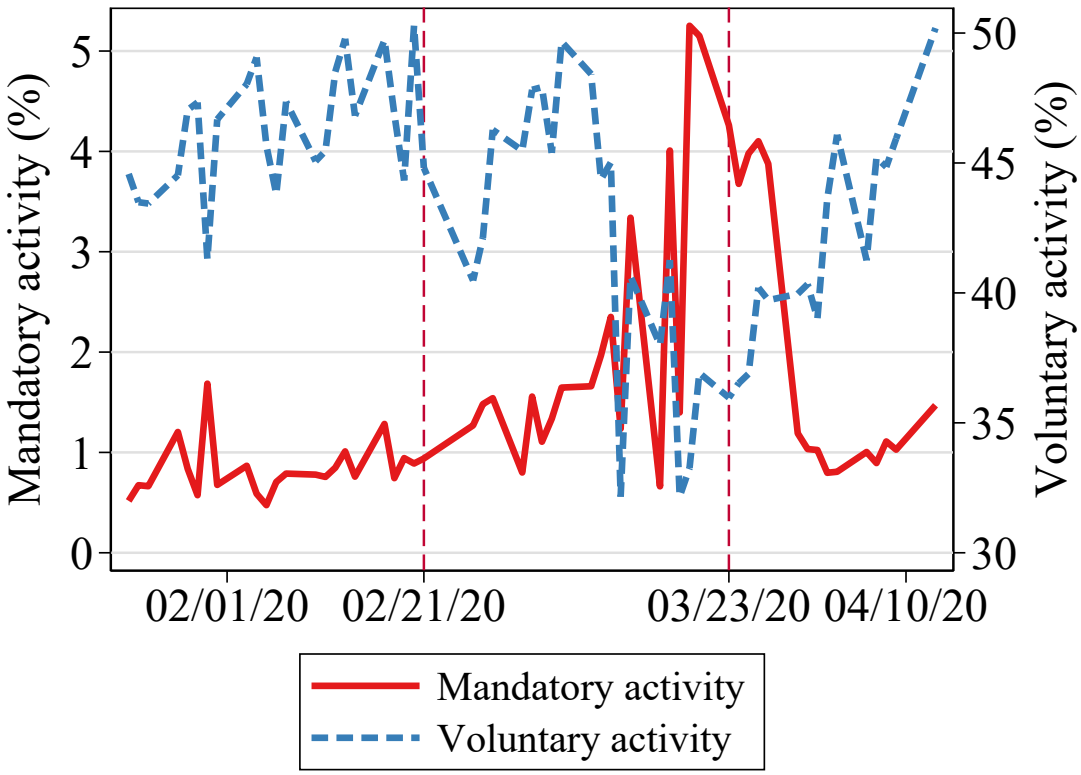
(b) Liquidity



(c) Volume

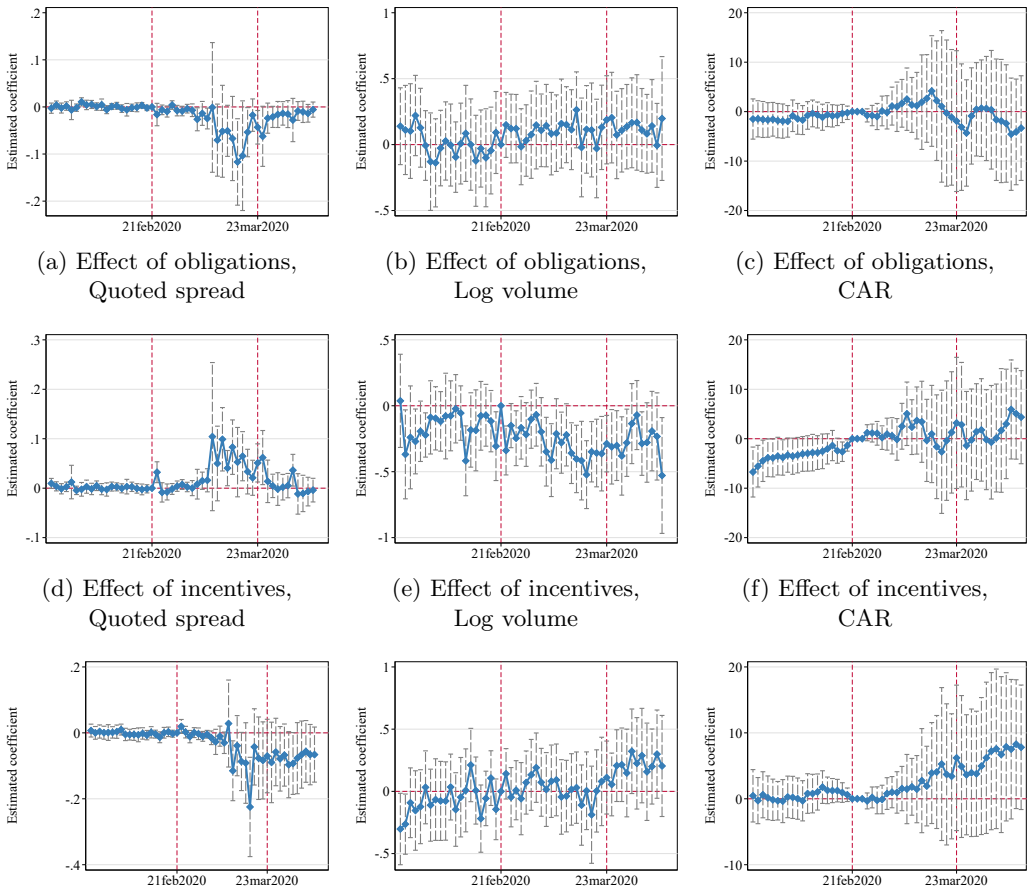
This figure shows time series of summary statistics for stocks in the 2019 market maker program during the 2020 stock market crash. Panel (a) plots cumulative returns on the equal-weighted portfolio of program stocks. Panel (b) plots median and 95th percentile quoted spreads as well as the fraction of time the program spread is binding. Panel (c) plots average daily trading volume. The variables in panels (b) and (c) are smoothed using a 5-day moving average. The red vertical lines indicate the start and the end of the COVID-19 stock market crash.

Figure 1.10: Share of Liquidity Provision During the COVID-19 Crash



This figure reports statistics describing the behavior of the mandatory and voluntary activities during the 2020 COVID-19 crash. The figure plots the percent share of liquidity provision, i.e. passive trading volume of the respective activity scaled by total trading volume. The red vertical lines indicate the start and the end of the COVID-19 stock market crash.

Figure 1.11: Effects of Obligations, Effects of Incentives, and Overall Program Effects on Liquidity, Volume, and Asset Prices During the COVID-19 Crash



(a) Effect of obligations, Quoted spread

(b) Effect of obligations, Log volume

(c) Effect of obligations, CAR

(d) Effect of incentives, Quoted spread

(e) Effect of incentives, Log volume

(f) Effect of incentives, CAR

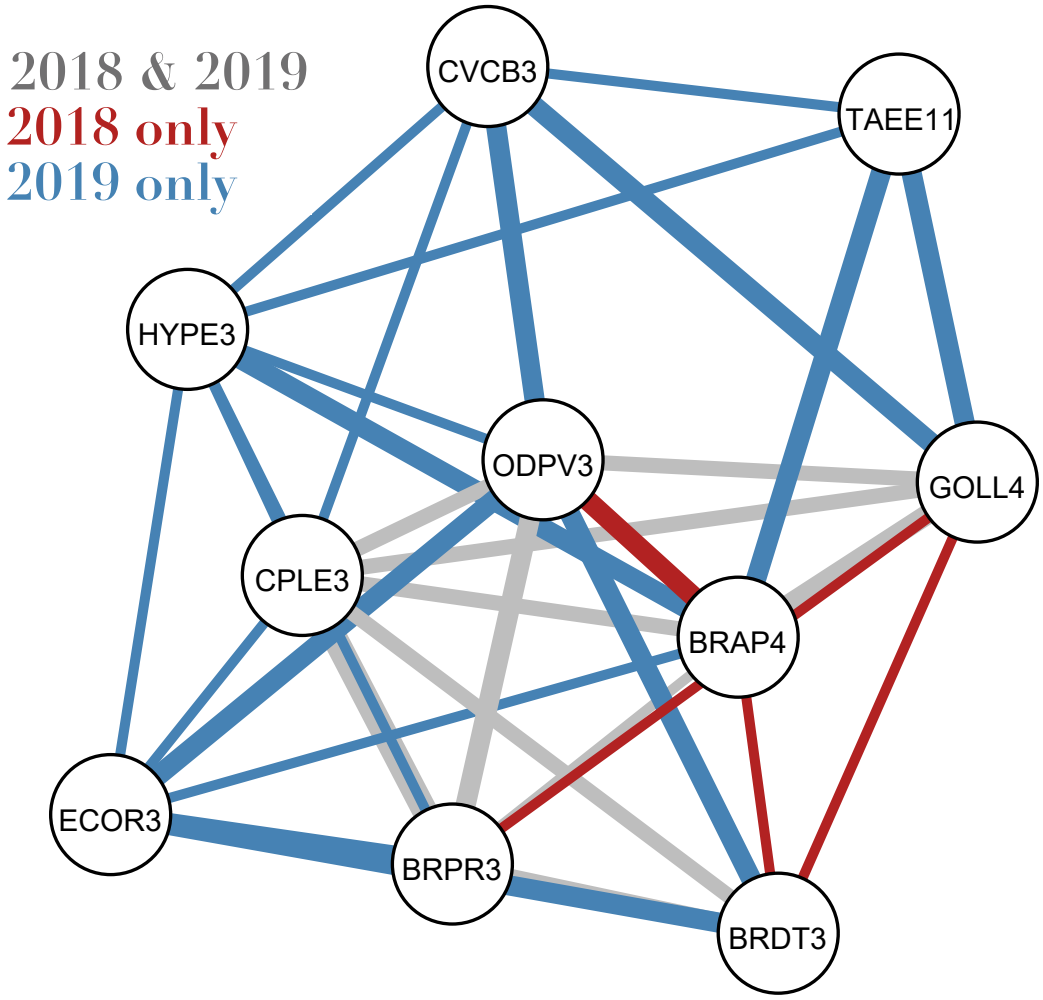
(g) Overall effect, Quoted spread

(h) Overall effect, Log volume

(i) Overall effect, CAR

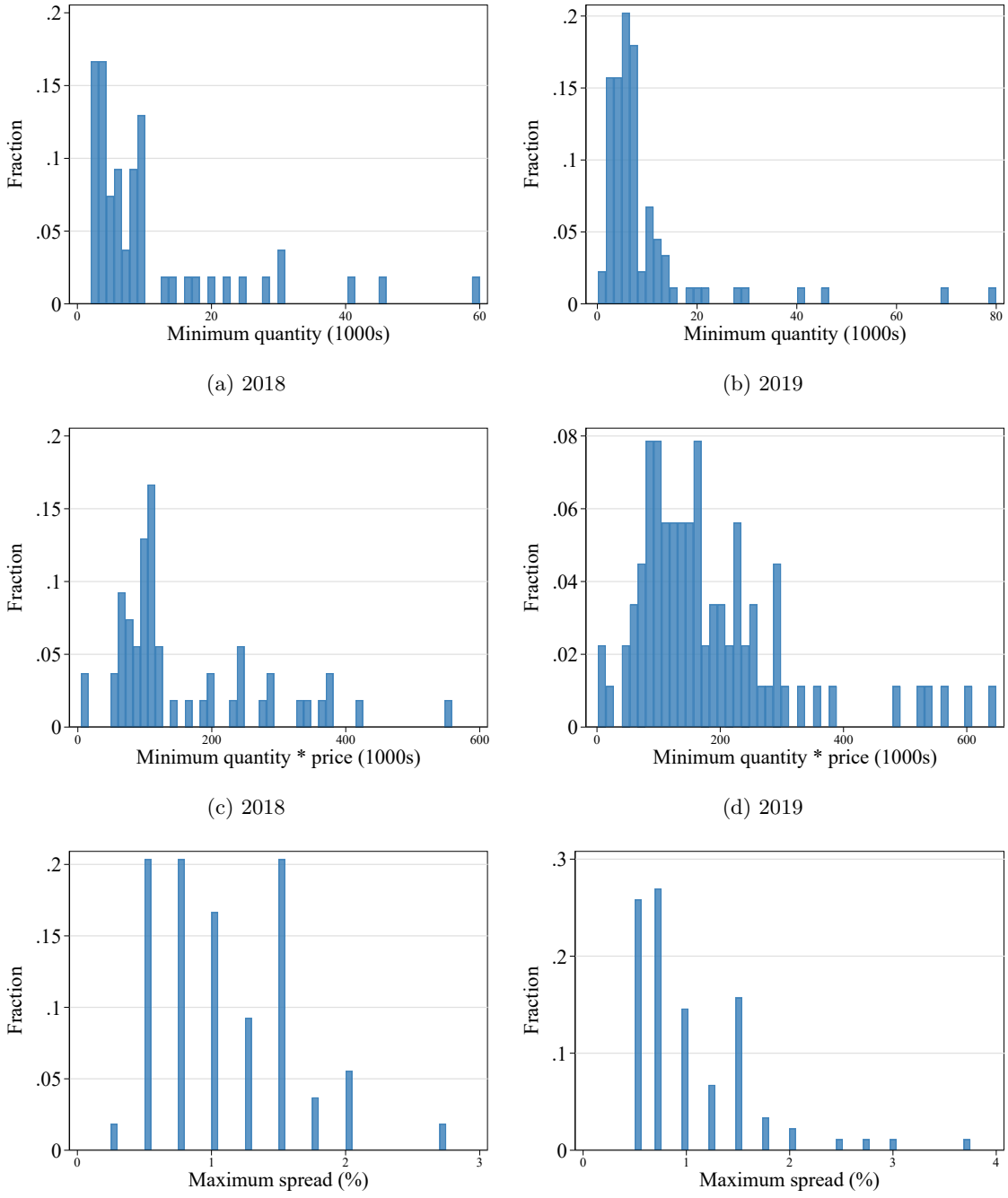
This figure shows overall program effects (a, b, c), effects of tight program obligations (d, e, f), and effects of incentives (g, h, i) during the 2020 stock market crash. We report effects on quoted spreads (a, d, g), log trading volume (b, e, h), and cumulative log returns (c, f, i). The red vertical lines indicate the start and the end of the COVID-19 stock market crash.

Figure 1.12: The Evolution of the Common Intermediation Network (10 Stock Subsample)



This figure illustrates the natural experiments we exploit. In 2018, the Brazilian stock exchange assigns 2 or 3 (out of 9) market makers to each of 54 (out of around 400) stocks. In 2019, the Brazilian stock exchange assigns 2 to 5 (out of 13) market makers to each of 89 stocks. The nodes in the graph are the ticker symbols of 10 randomly selected program stocks. Nodes are connected if they have a common market maker. Gray edges indicate that the two stocks are connected during both market maker programs. Red edges indicate that stocks were connected during the 2018 program, but got disconnected. Blue edges indicate new connections created by the 2019 program.

Figure 1.13: Distribution of Program Parameters



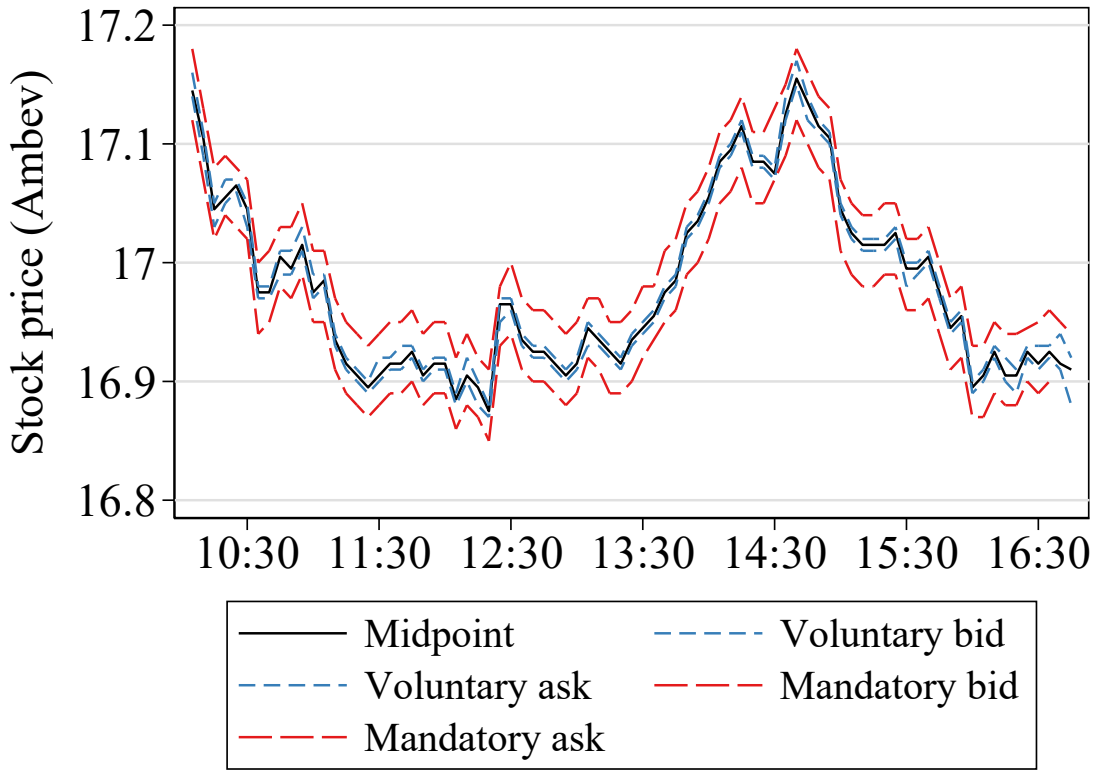
(a) 2018 (b) 2019

(c) 2018 (d) 2019

(e) 2018 (f) 2019

This figure shows histograms of program parameters. The panels on the left correspond to the 2018 program, the panels on the right to the 2019 program. The first row summarizes minimum quantities in 1000 shares. The second row shows the corresponding R$ execution risk by multiplying the minimum quantities in 1000 shares with the respective stock prices on the first day of the respective program. The last row shows the distribution of percent maximum spreads.

Figure 1.14: Illustration of the Mandatory vs. Voluntary Activity: Quoting of Tucana Bay in Ambev on 1st Day of 2018 Program



This figure illustrates the typical quoting behavior of the mandatory vs. voluntary activity using an example. We plot Tucana Bay's quotes for Ambev on the first day of the 2018 program, 10/15/2018. We plot the midpoint in black, Tucana Bay's voluntary bids and asks in blue, and Tucana Bay's mandatory bids and asks in red. The figure shows that, typically, the mandatory activity quotes deep in the book, while the voluntary activity quotes tightly around midpoints.

Tables

Table 1.1: Monthly Stock Level Summary Statistics

	mean	sd	p5	p25	p50	p75	p95
Market value	12756.6	29829.4	140.0	728.4	3663.7	11603.0	48710.7
Price	26.7	33.4	2.4	8.5	18.8	33.5	68.8
Return	3.1	13.7	-16.1	-3.5	2.4	9.2	22.7
Quoted spread	33.6	53.0	4.9	7.5	11.8	31.4	145.2
Effective spread	33.8	53.1	5.0	7.5	11.7	32.0	143.9
Depth (5 levels)	707.2	1533.1	64.1	146.8	281.3	581.0	2931.5
Depth (0.5%)	1379.2	3904.1	26.3	153.9	485.5	1535.0	4638.7
Depth (total)	5214.0	10357.6	138.0	690.1	2832.6	6369.0	17303.6
Volume	49.9	129.8	0.0	0.2	8.9	55.8	181.4
Number of trades	5401.2	8019.3	2.0	38.2	2126.8	8212.4	19812.9
Average trade size	8422.9	33401.6	1329.2	3161.0	5341.8	8592.9	17415.9
Sd(log quoted spread)	48.5	18.4	16.3	35.4	49.9	62.1	75.5
Sd(log effective spread)	37.6	14.5	13.7	26.7	38.6	48.1	60.0
Sd(log depth)	80.0	25.2	45.6	61.2	75.1	98.0	123.2
Sd(log volume)	310.9	68.2	205.6	244.6	322.4	364.6	410.1
Sd(return)	23.5	8.2	14.5	18.3	21.7	26.8	38.0
Observations	3601						

This table reports stock-level summary statistics. We report means, standard deviations, 5th, 25th, 50th, 75th, and 95th percentiles. The variables are stock market value in million R$s, stock price, returns in %, quoted and effective bid-ask spreads in basis points, depth in 1000 R$s, daily volume in million R$s, daily number of trades, average trade size, order flow, order imbalance, percent intraday (5-minute) standard deviations of log quoted and effective spreads, log depth, log volume, and returns. The sample contains the eight-month event window around the respective market maker program at the monthly frequency.

Table 1.2: Summary Statistics: High-Frequency Stock Pair Level Correlations

	mean	sd	p5	p25	p50	p75	p95
ρ(log quoted spread)	17.3	17.9	-5.0	5.3	13.3	26.4	52.9
ρ(log effective spread)	10.9	9.9	-3.9	5.1	10.6	17.0	26.7
ρ(log depth)	51.3	21.3	12.6	39.1	54.1	66.7	80.6
ρ(log volume)	33.8	30.9	-3.9	7.6	23.1	65.3	83.1
ρ(returns)	7.3	10.7	-6.3	0.5	5.5	12.7	26.7
Observations	169344						

This table reports stock pair level summary statistics. We report means, standard deviations, 5th, 25th, 50th, 75th, and 95th percentiles. The variables are percent intraday (5-minute) correlation coefficients of returns, log quoted and effective spreads, log depth, and log volume. The sample contains the eight-month event window around the respective market maker program at the monthly frequency.

Table 1.3: Stock-Day Level Summary Statistics: Trading and Quoting by Institution

Market Maker	Volume	Trade size	Trades	Quotes	Ratio	% Ag.	% New	% Fill	% Cancel	% Repl.
Panel A: Mandatory Activity										
Bradesco	725	41.1	18	9040	513.0	0.0	39.4	0.1	39.4	21.2
Citadel Securities	823	73.1	11	19082	1694.9	0.0	50.0	0.0	50.0	0.0
Credit Suisse	296	81.5	4	5102	1404.6	0.0	0.4	0.0	0.3	99.3
Green Post Trading	303	57.6	5	7053	1338.2	0.0	0.1	0.0	0.1	99.7
Headlands Technologies	944	50.9	19	15105	814.6	0.0	50.0	0.1	50.0	0.0
Jane Street Capital	1521	55.1	28	8836	320.3	0.0	49.9	0.3	49.8	0.0
Optiver	471	68.1	7	2260	327.1	0.0	20.9	0.1	20.8	58.2
QE Trading	373	61.3	6	5703	938.7	0.0	39.7	0.0	39.7	20.5
Spire X Trading	306	45.9	7	969	145.6	0.0	45.2	0.3	45.0	9.4
Tarpika	620	64.8	10	11997	1253.5	0.0	49.2	0.0	49.2	1.5
Tucana Bay	874	39.5	22	2064	93.3	0.0	49.9	0.4	49.8	0.0
Virtu Financial	570	55.1	10	12453	1202.0	0.0	43.1	0.0	43.1	13.7
XTX Markets	593	63.6	9	2881	309.0	0.0	48.6	0.1	48.5	2.8
Panel B: Voluntary Activity										
Bradesco	22798	5.8	3927	93871	23.9	46.8	47.4	3.5	45.4	3.8
Citadel Securities	30960	7.6	4073	14365	3.5	58.0	43.9	18.1	34.9	3.1
Credit Suisse	20200	4.7	4266	21593	5.1	29.1	44.9	14.4	32.2	8.4
Green Post Trading	11976	7.3	1630	9938	6.1	31.5	38.3	11.7	29.1	20.9
Headlands Technologies	48132	9.8	4916	16814	3.4	21.1	40.5	21.6	33.6	4.4
Jane Street Capital	30657	12.5	2454	8030	3.3	37.5	38.9	27.1	24.6	9.4
Optiver	22679	6.7	3400	25076	7.4	51.9	41.3	8.2	36.5	14.0
QE Trading	13347	7.9	1689	22596	13.4	37.6	28.5	5.2	25.5	40.8
Spire X Trading	12754	7.0	1827	12139	6.6	44.1	42.1	12.2	33.8	12.0
Tarpika	24724	4.6	5388	26065	4.8	33.2	43.4	11.7	33.9	11.0
Tucana Bay	44771	7.8	5734	25575	4.5	59.9	45.8	14.3	38.2	1.7
Virtu Financial	24487	6.4	3837	27034	7.0	44.2	45.7	11.3	38.5	4.5
XTX Markets	26722	6.2	4320	10563	2.4	58.3	42.5	27.0	28.7	1.8

This table provides stock-day level summary statistics by market maker. We report average stock-day trading volume and average trade size in 1000 R$s, the average number of trades and messages, the ratio of quotes to trades, and the % share of volume in which the market maker is the aggressor. We also report the fraction of messages that are new orders, fills, cancels, and replace orders. Panel A reports summary statistics for the mandatory activity, panel B for the voluntary activity. The sample comprises program stocks during the market maker programs.

Table 1.4: Comparison of Treated and Control Stocks?

	(1) 2018	(2) 2019
IBrX 100	0.503***	0.819***
	(0.0967)	(0.192)
1(bilateral DMM)	-0.249***	0.465***
	(0.0591)	(0.106)
1(top stock)	-0.565***	
	(0.102)	
Log market value	0.00711	-0.00674
	(0.0286)	(0.0951)
Log book equity	-0.00873	0.00702
	(0.0194)	(0.0232)
Log PP&E	-0.00680	0.0211
	(0.0137)	(0.0245)
Profitability	-1.155	0.432
	(1.055)	(2.112)
Investment	-3.409	1.842
	(3.214)	(1.697)
Asset growth	-0.353	-0.599
	(0.270)	(0.398)
Log debt	0.0217	-0.0284
	(0.0233)	(0.0469)
Log quoted spread	0.0761	0.00563
	(0.0681)	(0.217)
Log volume	0.0338	-0.0618
	(0.0297)	(0.118)
Sd(log quoted spread)	-0.000571	0.00550
	(0.00152)	(0.00347)
Sd(log volume)	-0.000159	-0.000675
	(0.000409)	(0.00154)
Sd(return)	0.000409	0.0152
	(0.00437)	(0.0171)
R-Squared	0.428	0.412
N	196	87

This table reports results of regressions of a treatment indicator on stock characteristics. Columns 1 and 2 report results for the 2018 and 2019 market maker program, respectively. The sample is the cross-section of treated and control stocks at the beginning of the event window, i.e. 3 months before the treatment. The first regression contains about 200 stocks, the universe of 400 stocks excluding small stocks with a market capitalization below R$100M. The second regression includes all stocks included in the 2019 MM program. ***, **, and * denote significance at the 10%, 5%, and 1% levels.

Table 1.5: The Effects of Designated Market Makers on Liquidity, Volume, and Asset Prices

Dep. Var	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: The 2018 Market Maker Program						
Log quoted spread	-0.14**	-0.13***	-0.15**	-0.08**	-0.09**	-0.09**
	(0.06)	(0.05)	(0.06)	(0.04)	(0.04)	(0.04)
Log effective spread	-0.16**	-0.15***	-0.16**	-0.09**	-0.10**	-0.10**
	(0.06)	(0.05)	(0.06)	(0.04)	(0.04)	(0.04)
Log depth	0.52***	0.52***	0.53***	0.49***	0.49***	0.49***
	(0.09)	(0.07)	(0.09)	(0.06)	(0.06)	(0.06)
Log volume	0.16*	0.15*	0.16*	0.14**	0.14**	0.14**
	(0.09)	(0.08)	(0.09)	(0.07)	(0.07)	(0.07)
CAR	7.79*	7.32	7.78*	3.81	3.81	3.81
	(4.56)	(4.52)	(4.57)	(3.19)	(3.19)	(3.19)
Panel B: The 2019 Market Maker Program						
Log quoted spread	-0.16***	-0.16***	-0.16***	-0.15***	-0.15***	-0.15***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Log effective spread	-0.18***	-0.18***	-0.18***	-0.17***	-0.17***	-0.17***
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Log depth	0.51***	0.50***	0.51***	0.52***	0.52***	0.52***
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Log volume	0.08	0.09	0.08	0.13	0.13	0.13
	(0.10)	(0.09)	(0.10)	(0.08)	(0.08)	(0.08)
CAR	2.92	2.87	2.87	3.33	3.27	3.27
	(3.49)	(3.50)	(3.51)	(3.51)	(3.52)	(3.52)
Controls		✓				✓
Time FE			✓		✓	✓
Stock FE				✓	✓	✓
Cluster by stock	✓	✓	✓	✓	✓	✓

This table reports estimates of equation 1.1. The rows report results for different dependent variables: log quoted spreads, log effective spreads, log depth, log volume, and percent cumulative abnormal returns. The columns include different sets of control variables and fixed effects. We report standard errors clustered by stock in parentheses. The sample contains the eight-month event window around the respective market maker program at the monthly frequency. ***, **, and * denote significance at the 10%, 5%, and 1% levels.

Table 1.6: The Effects of Designated Market Makers on the Volatilities of Liquidity, Volume, and Asset Prices

Dep. Var	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: The 2018 Market Maker Program						
Sd(log quoted spread)	-6.27***	-6.26***	-6.30***	-6.24***	-6.28***	-6.28***
	(1.30)	(1.31)	(1.30)	(1.24)	(1.24)	(1.24)
Sd(log effective spread)	-2.79***	-2.76***	-2.81***	-2.86***	-2.89***	-2.89***
	(1.02)	(1.01)	(1.02)	(0.98)	(0.98)	(0.98)
Sd(log depth)	-8.86***	-8.53***	-8.82***	-9.45***	-9.42***	-9.42***
	(2.24)	(2.14)	(2.25)	(2.05)	(2.05)	(2.05)
Sd(log volume)	-49.00***	-48.52***	-49.00***	-47.79***	-47.81***	-47.81***
	(6.50)	(6.54)	(6.51)	(6.55)	(6.57)	(6.57)
Sd(ret)	-0.80	-0.77	-0.85	0.29	0.22	0.22
	(0.63)	(0.61)	(0.62)	(0.55)	(0.55)	(0.55)
Panel B: The 2019 Market Maker Program						
Sd(log quoted spread)	-6.99***	-7.00***	-7.00***	-7.43***	-7.42***	-7.42***
	(1.69)	(1.69)	(1.68)	(1.55)	(1.54)	(1.54)
Sd(log effective spread)	-3.24**	-3.24**	-3.26**	-3.81***	-3.80***	-3.80***
	(1.31)	(1.31)	(1.30)	(1.20)	(1.19)	(1.19)
Sd(log depth)	-4.19	-4.16	-4.20	-4.01	-4.03	-4.03
	(2.56)	(2.54)	(2.57)	(2.54)	(2.56)	(2.56)
Sd(log volume)	-11.01	-10.97	-11.07	-11.51	-11.56	-11.56
	(8.04)	(8.05)	(8.07)	(8.22)	(8.25)	(8.25)
Sd(ret)	0.37	0.38	0.37	0.37	0.37	0.37
	(0.47)	(0.46)	(0.47)	(0.44)	(0.44)	(0.44)
Controls		✓				✓
Time FE			✓		✓	✓
Stock FE				✓	✓	✓
Cluster by stock	✓	✓	✓	✓	✓	✓

This table reports estimates of equation 1.1. The rows report results for percent intraday (5-minute) volatilities of different dependent variables: log quoted spreads, log effective spreads, log depth, log volume, and returns. The columns include different sets of control variables and fixed effects. We report standard errors clustered by stock in parentheses. The sample contains the eight-month event window around the respective market maker program at the monthly frequency. ***, **, and * denote significance at the 10%, 5%, and 1% levels.

Table 1.7: The Separate Effects of Liquidity Provision Obligations and Incentives on Liquidity, Volume, and Asset Prices

Dep. Var	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Effect of Obligations						
Log quoted spread	0.16	0.14	0.16	0.10***	0.10***	0.10***
	(0.13)	(0.11)	(0.13)	(0.03)	(0.03)	(0.03)
Log effective spread	0.16	0.15	0.17	0.11***	0.11***	0.11***
	(0.14)	(0.12)	(0.14)	(0.03)	(0.03)	(0.03)
Log depth	-0.15	-0.10	-0.15	-0.06	-0.06	-0.06
	(0.12)	(0.09)	(0.12)	(0.08)	(0.08)	(0.08)
Log volume	-0.24	-0.21	-0.25	-0.15	-0.15	-0.15
	(0.17)	(0.14)	(0.18)	(0.09)	(0.09)	(0.09)
CAR	-5.95	-5.95	-5.95	-5.95	-5.95	-5.95
	(4.73)	(4.74)	(4.76)	(4.71)	(4.75)	(4.75)
Panel B: Effect of Incentives						
Log quoted spread	-0.17***	-0.17***	-0.17***	-0.16***	-0.16***	-0.16***
	(0.06)	(0.05)	(0.06)	(0.04)	(0.04)	(0.04)
Log effective spread	-0.18**	-0.18***	-0.18***	-0.16***	-0.17***	-0.17***
	(0.07)	(0.05)	(0.07)	(0.04)	(0.04)	(0.04)
Log depth	0.25**	0.23**	0.25**	0.21**	0.22**	0.22**
	(0.11)	(0.10)	(0.11)	(0.09)	(0.09)	(0.09)
Log volume	0.26**	0.26**	0.27**	0.23**	0.24**	0.24**
	(0.13)	(0.11)	(0.13)	(0.10)	(0.10)	(0.10)
CAR	12.75***	12.75***	12.75***	12.75***	12.75***	12.75***
	(3.65)	(3.66)	(3.67)	(3.64)	(3.67)	(3.67)
Controls		✓				✓
Time FE			✓		✓	✓
Stock FE				✓	✓	✓
Cluster by stock	✓	✓	✓	✓	✓	✓

This table reports estimates of equation 1.1. The rows report results for different dependent variables: log quoted spreads, log effective spreads, log depth, log volume, and cumulative returns. The columns include different sets of control variables and fixed effects. We report standard errors clustered by stock in parentheses. The sample contains the eight-month event window around the respective market maker program at the monthly frequency. ***, **, and * denote significance at the 10%, 5%, and 1% levels.

Table 1.8: How Do Market Makers Set Quotes?

	Mandatory Activity		Voluntary Activity	
	(1)	(2)	(3)	(4)
OF	14.31***	0.74***	16.21***	1.74***
	(0.17)	(0.06)	(0.14)	(0.05)
OF absorbed	65.53***	8.08***	78.59***	18.74***
	(1.33)	(0.33)	(1.11)	(0.38)
OF (other stocks)	5.97***	-0.20***	7.76***	0.73***
	(0.23)	(0.07)	(0.15)	(0.03)
OF absorbed (other stocks)	19.77***	4.52***	26.51***	3.88***
	(1.57)	(0.46)	(1.06)	(0.22)
Return		0.90***		0.91***
		(0.00)		(0.00)
N events	4,551,149,609	4,551,149,609	6,866,048,833	6,866,048,833
N reactions	85,524,480	85,524,480	46,823,032	46,823,032
MM-stock-days	22,037	22,037	30,736	30,736
MM-days	1,610	1,610	1,629	1,629

This table reports estimates of equation 1.6 and shows how market makers revise price quotes in response to order flow, order flow where they provide liquidity, and both variables constructed using all other stocks the respective market maker intermediates. We also control for stock returns. Quote revisions and stock returns are in percent. We estimate equation 1.6 separately for each market maker x stock x day combination. This is conceptually equivalent to including market maker x stock x day fixed effects. We then compute Fama-MacBeth standard errors robust to within market maker x day correlation from the distribution of the coefficient estimates (Fama and MacBeth, 1973). We report standard errors in parentheses. The sample spans the duration of the market maker programs. ***, **, and * denote significance at the 10%, 5%, and 1% levels.

Table 1.9: The Effects of Common Intermediation on Co-Movement

Dep. Var	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: The 2018 Market Maker Program						
ρ(log quoted spread)	1.03**	1.01**	1.04**	1.03**	1.02**	1.01**
	(0.44)	(0.44)	(0.44)	(0.44)	(0.44)	(0.44)
ρ(log effective spread)	1.70***	1.71***	1.83***	1.85***	1.84***	1.86***
	(0.65)	(0.65)	(0.62)	(0.63)	(0.62)	(0.63)
ρ(log depth)	1.63***	1.61***	1.88***	1.81***	1.85***	1.79***
	(0.37)	(0.37)	(0.36)	(0.36)	(0.36)	(0.36)
ρ(log volume)	0.43	0.43	0.54*	0.56*	0.54*	0.56*
	(0.34)	(0.34)	(0.31)	(0.31)	(0.31)	(0.31)
ρ(return)	0.62**	0.60**	0.68***	0.79***	0.66***	0.77***
	(0.26)	(0.25)	(0.23)	(0.23)	(0.23)	(0.23)
Panel B: The 2019 Market Maker Program						
ρ(log quoted spread)	1.51***	1.50***	1.54***	1.56***	1.53***	1.55***
	(0.20)	(0.20)	(0.19)	(0.19)	(0.19)	(0.19)
ρ(log effective spread)	2.77***	2.76***	2.80***	2.82***	2.78***	2.80***
	(0.19)	(0.19)	(0.18)	(0.18)	(0.17)	(0.18)
ρ(log depth)	0.18	0.18	0.24*	0.26*	0.24	0.26*
	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)
ρ(log volume)	0.55***	0.53***	0.63***	0.65***	0.61***	0.64***
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
ρ(return)	0.64***	0.61***	0.69***	0.72***	0.66***	0.69***
	(0.10)	(0.11)	(0.10)	(0.10)	(0.10)	(0.10)
Time FE		✓			✓	✓
Stock FE			✓		✓	
Pair FE				✓		✓

This table reports estimates of equation 1.7. The rows report results for different dependent variables: the percent correlation of intraday (5-minute) log quoted spreads, log effective spreads, log depth, log volume, and returns. The columns include different sets of fixed effects. We report standard errors clustered by stock pair in parentheses. The sample contains the eight-month event window around the respective market maker program at the monthly frequency. ***, **, and * denote significance at the 10%, 5%, and 1% levels.

Table 1.10: Details of the 2018 Market Maker Program

Ticker	MM 1	MM 2	MM 3	% Spread	Q	% Presence
ABEV3	Headlands Technologies	Tucana Bay	-	.25	25000	80
ALPA4	Headlands Technologies	Tucana Bay	QE Trading	1.5	8000	80
AZUL4	Jane Street Capital	QE Trading	Virtu Financial	1.25	4000	80
BBDC3	Headlands Technologies	Credit Suisse Brasil	QE Trading	.5	4000	80
BBSE3	Credit Suisse Brasil	Tucana Bay	-	.5	9000	80
BPAC11	Headlands Technologies	XTX Markets	QE Trading	1.5	5000	80
BRAP4	XTX Markets	Tucana Bay	Spire X Trading	1	6000	80
BRDT3	Headlands Technologies	Credit Suisse Brasil	XTX Markets	1	10000	80
BRFS3	XTX Markets	Virtu Financial	-	.75	18000	80
BRML3	Credit Suisse Brasil	Tucana Bay	Bradesco	.75	20000	80
BRPR3	Headlands Technologies	Spire X Trading	QE Trading	1.5	10000	80
BRSR6	Headlands Technologies	Credit Suisse Brasil	XTX Markets	1	6000	80
BTOW3	XTX Markets	Spire X Trading	QE Trading	.75	4000	80
CESP6	Spire X Trading	Tucana Bay	QE Trading	1	4000	80
CIEL3	Credit Suisse Brasil	QE Trading	-	1	30000	80
CMIG4	Headlands Technologies	Credit Suisse Brasil	Spire X Trading	.75	28000	80
CPLE3	Headlands Technologies	Tucana Bay	XTX Markets	2	3000	80
CPRE3	Headlands Technologies	Tucana Bay	Spire X Trading	2	3000	80
CSAN3	Headlands Technologies	XTX Markets	QE Trading	.5	3000	80
CSNA3	Jane Street Capital	Spire X Trading	QE Trading	.75	41000	80

Table 1.10: Details of the 2018 Market Maker Program (continued)

Ticker	MM 1	MM 2	MM 3	% Spread	Q	% Presence
ELET3	Headlands Technologies	Credit Suisse Brasil	Virtu Financial	1	14000	80
ELET6	Credit Suisse Brasil	Tucana Bay	XTX Markets	1	6000	80
EMBR3	Jane Street Capital	QE Trading	Virtu Financial	.5	9000	80
ENGI11	Spire X Trading	Tucana Bay	QE Trading	1.5	3000	80
FIBR3	XTX Markets	Tucana Bay	-	.5	4000	80
GFSA3	Credit Suisse Brasil	Tucana Bay	Spire X Trading	1.5	8000	80
GNDI3	Credit Suisse Brasil	Tucana Bay	XTX Markets	1.5	4000	80
GOAU4	Spire X Trading	QE Trading	-	.75	45000	80
GOLL4	XTX Markets	Tucana Bay	Spire X Trading	.75	7000	80
HAPV3	Credit Suisse Brasil	XTX Markets	QE Trading	1.25	3000	80
HGTX3	Headlands Technologies	Credit Suisse Brasil	XTX Markets	1	6000	80
KLBN11	Headlands Technologies	XTX Markets	QE Trading	.5	5000	80
LAME3	Headlands Technologies	Tucana Bay	Credit Suisse Brasil	1.5	5000	80
MRFG3	Spire X Trading	Tucana Bay	QE Trading	1.25	13000	80
MRVE3	XTX Markets	Tucana Bay	Spire X Trading	.5	8000	80
ODPV3	Spire X Trading	Tucana Bay	QE Trading	1.5	7000	80
PARD3	Spire X Trading	Tucana Bay	QE Trading	1.5	4000	80
PCAR4	Jane Street Capital	Spire X Trading	Virtu Financial	.5	3000	80
POMO4	Headlands Technologies	Credit Suisse Brasil	Spire X Trading	1.75	30000	80
RAIL3	Credit Suisse Brasil	XTX Markets	-	.75	22000	80

Table 1.10: Details of the 2018 Market Maker Program (continued)

Ticker	MM 1	MM 2	MM 3	% Spread	Q	% Presence
RAPT4	Credit Suisse Brasil	Tucana Bay	QE Trading	1.5	16000	80
RENT3	XTX Markets	Tucana Bay	Spire X Trading	.75	8000	80
RNEW11	Headlands Technologies	XTX Markets	Spire X Trading	2	2000	80
SAPR11	Headlands Technologies	Credit Suisse Brasil	XTX Markets	.75	2000	80
SAPR4	Headlands Technologies	XTX Markets	Spire X Trading	1.25	8000	80
SBSP3	Jane Street Capital	Tucana Bay	Virtu Financial	.5	4000	80
SEER3	Headlands Technologies	XTX Markets	Credit Suisse Brasil	1.5	4000	80
SLCE3	Headlands Technologies	Credit Suisse Brasil	XTX Markets	1.25	2000	80
TPIS3	Headlands Technologies	Tucana Bay	Spire X Trading	2.75	10000	80
UGPA3	Headlands Technologies	Spire X Trading	-	.5	5000	80
USIM5	Headlands Technologies	Spire X Trading	-	.75	60000	80
VVAR11	XTX Markets	Tucana Bay	Spire X Trading	1	9000	80
WEGE3	Credit Suisse Brasil	Spire X Trading	QE Trading	.5	6000	80
WIZS3	Headlands Technologies	Tucana Bay	XTX Markets	1.75	10000	80

This table reports the 2018 market maker program details as provided by B3. The columns report the stock ticker symbol, the designated market makers, the maximum spread in percent, the minimum quantity in number of shares, and the minimum presence during each trading session in percent.

Table 1.11: Details of the 2019 Market Maker Program

Ticker	MM 1	MM 2	MM 3	MM 4	MM 5	% Spread	Q	% Presence
ALPA4	QE Trading	Credit Suisse	Tucana Bay	Tarpika	Citadel Sec.	1.5	8000	80
AZUL4	QE Trading	Optiver	Virtu Fin.	-	-	.75	4000	80
BBDC3	Headlands Tech.	Credit Suisse	Optiver	-	-	.5	4000	80
BBSE3	Credit Suisse	Tucana Bay	-	-	-	.5	9000	80
BEEF3	QE Trading	Spire X Trading	XTX Markets	Tarpika	Citadel Sec.	1.5	11500	80
BPAC11	QE Trading	XTX Markets	Citadel Sec.	Optiver	-	1.5	5000	80
BRAP4	Headlands Tech.	Optiver	Tucana Bay	-	-	.75	6000	80
BRDT3	Spire X Trading	Credit Suisse	Citadel Sec.	-	-	.75	10000	80
BRFS3	Tucana Bay	Optiver	-	-	-	.5	18000	80
BRML3	Bradesco	Credit Suisse	Tucana Bay	-	-	.75	20000	80
BRPR3	QE Trading	Green Post Tr.	Spire X Trading	Optiver	Citadel Sec.	1.5	10000	80
BRSR6	Spire X Trading	XTX Markets	Optiver	-	-	.75	6000	80
BTOW3	QE Trading	Spire X Trading	Virtu Fin.	-	-	.75	6000	80
CCRO3	Green Post Tr.	Tarpika	-	-	-	.5	10000	80
CESP6	QE Trading	Optiver	Citadel Sec.	-	-	1	4000	80
CIEL3	Spire X Trading	Tarpika	-	-	-	1.5	80000	80
CMIG3	QE Trading	Green Post Tr.	Spire X Trading	Tarpika	Credit Suisse	3.75	6000	80
CMIG4	Green Post Tr.	Optiver	-	-	-	.5	28000	80
CPLE3	QE Trading	Green Post Tr.	Spire X Trading	XTX Markets	Optiver	2	3000	80
CPLE6	XTX Markets	Virtu Fin.	Optiver	-	-	.75	3000	80

Table 1.11: Details of the 2019 Market Maker Program (continued)

Ticker	MM 1	MM 2	MM 3	MM 4	MM 5	% Spread	Q	% Presence
CPRE3	Optiver	-	-	-	-	3	1000	80
CRFB3	Citadel Sec.	Tarpika	Green Post Tr.	-	-	1	6500	80
CSAN3	QE Trading	Citadel Sec.	-	-	-	.5	3000	80
CSMG3	Spire X Trading	XTX Markets	Optiver	-	-	.75	3000	80
CSNA3	XTX Markets	Optiver	-	-	-	.5	41000	80
CVCB3	Virtu Fin.	Tarpika	XTX Markets	-	-	.75	5500	80
CYRE3	Spire X Trading	Tucana Bay	Tarpika	-	-	.75	6500	80
DTEX3	QE Trading	Spire X Trading	Credit Suisse	Tucana Bay	Citadel Sec.	1	12000	80
ECOR3	Credit Suisse	Optiver	Citadel Sec.	-	-	.5	4000	80
ELET3	QE Trading	Credit Suisse	-	-	-	.75	14000	80
ELET6	Tucana Bay	Tarpika	Citadel Sec.	-	-	.75	6000	80
EMBR3	Tucana Bay	Optiver	Virtu Fin.	-	-	.5	9500	80
ENAT3	QE Trading	XTX Markets	Tucana Bay	Tarpika	Citadel Sec.	1.5	5000	80
ENBR3	Green Post Tr.	XTX Markets	Tarpika	-	-	.75	5000	80
ENGI11	QE Trading	Green Post Tr.	XTX Markets	-	-	.75	3000	80
EZTC3	QE Trading	XTX Markets	Credit Suisse	Optiver	Tarpika	1	5000	80
GFSA3	QE Trading	Green Post Tr.	Spire X Trading	Optiver	Citadel Sec.	1.5	8000	80
GNDI3	Credit Suisse	Tucana Bay	Virtu Fin.	-	-	1	4000	80
GOAU4	Green Post Tr.	Headlands Tech.	Tarpika	-	-	.75	45000	80
GOLL4	XTX Markets	Tucana Bay	Virtu Fin.	-	-	.75	7000	80

Table 1.11: Details of the 2019 Market Maker Program (continued)

Ticker	MM 1	MM 2	MM 3	MM 4	MM 5	% Spread	Q	% Presence
HAPV3	QE Trading	Green Post Tr.	Credit Suisse	-	-	1	3000	80
HGTX3	Spire X Trading	XTX Markets	Credit Suisse	-	-	.75	6000	80
HYPE3	Headlands Tech.	Optiver	Tarpika	-	-	1	5000	80
IRBR3	Spire X Trading	XTX Markets	Tarpika	-	-	.75	2000	80
JBSS3	XTX Markets	Citadel Sec.	-	-	-	.5	10000	80
KLBN11	QE Trading	Headlands Tech.	Credit Suisse	-	-	.5	7500	80
LAME3	Green Post Tr.	Spire X Trading	XTX Markets	Credit Suisse	Citadel Sec.	1.5	5000	80
LAME4	Headlands Tech.	Optiver	Tarpika	-	-	.5	7000	80
LIGT3	Green Post Tr.	XTX Markets	Tucana Bay	Tarpika	Citadel Sec.	1.25	5000	80
LINX3	Green Post Tr.	Spire X Trading	Citadel Sec.	-	-	1.25	3000	80
LOGG3	QE Trading	Green Post Tr.	Spire X Trading	Optiver	Headlands Tech.	2	6500	80
MDIA3	Spire X Trading	XTX Markets	Credit Suisse	Virtu Fin.	Citadel Sec.	.75	2500	80
MRFG3	QE Trading	Tucana Bay	Virtu Fin.	-	-	1	13000	80
MRVE3	QE Trading	Spire X Trading	Virtu Fin.	-	-	.5	8000	80
MULT3	Headlands Tech.	Tarpika	Optiver	-	-	.75	8000	80
MYPK3	XTX Markets	Credit Suisse	Tucana Bay	Tarpika	Citadel Sec.	1	5000	80
NATU3	Green Post Tr.	Spire X Trading	XTX Markets	-	-	.75	8000	80
ODPV3	QE Trading	XTX Markets	Credit Suisse	Tarpika	Citadel Sec.	1.5	7000	80
PARD3	QE Trading	Green Post Tr.	Spire X Trading	XTX Markets	Optiver	1.5	4000	80
PCAR4	QE Trading	Citadel Sec.	-	-	-	.5	3000	80

Table 1.11: Details of the 2019 Market Maker Program (continued)

Ticker	MM 1	MM 2	MM 3	MM 4	MM 5	% Spread	Q	% Presence
POMO4	Green Post Tr.	Tarpika	Headlands Tech.	Tucana Bay	Citadel Sec.	1.75	30000	80
PSSA3	Green Post Tr.	XTX Markets	Tarpika	-	-	.5	4000	80
QUAL3	Green Post Tr.	Tucana Bay	Citadel Sec.	-	-	1.25	8000	80
RAIL3	XTX Markets	Credit Suisse	-	-	-	.5	22000	80
RAPT4	QE Trading	Green Post Tr.	Spire X Trading	Tucana Bay	Tarpika	1.5	16000	80
RENT3	Green Post Tr.	Spire X Trading	-	-	-	.5	8000	80
RNEW11	Green Post Tr.	Optiver	-	-	-	2.5	100	80
SAPR11	Headlands Tech.	Credit Suisse	Optiver	-	-	.75	2000	80
SAPR4	QE Trading	Green Post Tr.	Spire X Trading	XTX Markets	Tucana Bay	1.25	8000	80
SBSP3	Spire X Trading	Credit Suisse	Optiver	-	-	.5	4500	80
SEER3	Spire X Trading	XTX Markets	Credit Suisse	Optiver	Citadel Sec.	1.5	4000	80
SLCE3	Green Post Tr.	Tarpika	Citadel Sec.	-	-	1	3000	80
SMLS3	Green Post Tr.	Headlands Tech.	Citadel Sec.	-	-	.75	2000	80
SMTO3	QE Trading	Green Post Tr.	Spire X Trading	Credit Suisse	Citadel Sec.	1	3000	80
SULA11	XTX Markets	Credit Suisse	Citadel Sec.	-	-	1	4000	80
TAEE11	Headlands Tech.	Tucana Bay	Virtu Fin.	-	-	.5	3500	80
TEND3	QE Trading	Green Post Tr.	Spire X Trading	XTX Markets	Optiver	1.5	4000	80
TIET11	Green Post Tr.	Tucana Bay	Optiver	Tarpika	Citadel Sec.	1	8000	80
TIMP3	Headlands Tech.	Tucana Bay	Optiver	-	-	.5	12000	80
TOTS3	Green Post Tr.	XTX Markets	Credit Suisse	Tucana Bay	Citadel Sec.	1.25	4000	80

Table 1.11: Details of the 2019 Market Maker Program (continued)

Ticker	MM 1	MM 2	MM 3	MM 4	MM 5	% Spread	Q	% Presence
TPIS3	Green Post Tr.	Optiver	-	-	-	2.75	10000	80
TRPL4	XTX Markets	Tarpika	Optiver	-	-	.5	4000	80
TUPY3	QE Trading	Spire X Trading	Tucana Bay	Tarpika	Citadel Sec.	1.75	5000	80
UGPA3	Spire X Trading	Virtu Fin.	-	-	-	.5	12000	80
USIM5	Green Post Tr.	Headlands Tech.	-	-	-	1.25	70000	80
VLID3	QE Trading	Spire X Trading	Tucana Bay	Optiver	Citadel Sec.	1.5	5000	80
VVAR3	QE Trading	Tucana Bay	Tarpika	-	-	.75	13500	80
WEGE3	QE Trading	Credit Suisse	Optiver	-	-	.5	6000	80
WIZS3	QE Trading	Green Post Tr.	Spire X Trading	XTX Markets	Optiver	1.75	10000	80

This table reports the 2019 market maker program details as provided by B3. The columns report the stock ticker symbol, the designated market makers, the maximum spread in percent, the minimum quantity in number of shares, and the minimum presence during each trading session in percent. The unabbreviated names of the respective designated market makers are Citadel Securities, Green Post Trading, Credit Suisse Brasil, Virtu Financial, and Headlands Technologies.

Table 1.12: The Strategic Complementarity of Liquidity Provision: The Effects of Designated Market Makers on Other Traders' Liquidity Provision

Dep. Var	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: The 2018 Market Maker Program						
Log quoted spread (non-DMMs)	-0.11*	-0.09**	-0.11*	-0.05	-0.05	-0.05
	(0.06)	(0.04)	(0.06)	(0.04)	(0.04)	(0.04)
Log depth (non-DMMs)	0.20*	0.18**	0.20*	0.08	0.08	0.08
	(0.10)	(0.08)	(0.10)	(0.07)	(0.07)	(0.07)
Panel B: The 2019 Market Maker Program						
Log quoted spread (non-DMMs)	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Log depth (non-DMMs)	0.16**	0.15*	0.16**	0.14*	0.14*	0.14*
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Controls		✓				✓
Time FE			✓		✓	✓
Stock FE				✓	✓	✓
Cluster by stock	✓	✓	✓	✓	✓	✓

This table reports estimates of equation 1.1. The rows report results for different dependent variables: log quoted spreads ignoring DMM quotes and log depth ignoring DMM quotes. The columns include different sets of control variables and fixed effects. We report standard errors clustered by stock in parentheses. The sample contains the eight-month event window around the respective market maker program at the monthly frequency. ***, **, and * denote significance at the 10%, 5%, and 1% levels.

Chapter 2

Identifying the Price Impact of Fire Sales Using High-Frequency Surprise Mutual Fund Flows

2.1 Introduction

What is the price impact of mutual fund flow-induced fire sales? From a policy maker's perspective, this question is important for financial stability regulation; from a researcher's perspective, it is important because fund flow-induced trading generates the state-of-the-art natural experiment for changes in asset prices. Since Edmans, Goldstein, and Jiang (2012), researchers use fire sales triggered by contemporaneous fund flows as an instrument for stock returns and find that market misvaluation distorts corporate decisions (e.g. Khan, Kogan, and Serafeim, 2012; Derrien, Kecskes, and Thesmar, 2013; Phillips and Zhdanov, 2013; Norli, Ostergaard, and Schindele, 2014; Lee and So, 2017; Bonaime, Gulen, and Ion, 2018; Dessaint et al., 2018; Eckbo, Makaew, and Thorburn, 2018; Lou and Wang, 2018). Moreover, using flow-induced trading, Coval and Stafford (2007), Duffie (2010), Greenwood and Thesmar (2011), Ben-Rephael, Kandel, and Wohl (2012), Jotikasthira,

Lundblad, and Ramadorai (2012), Lou (2012), Anton and Polk (2014) and Li (2019) show that price pressure effects are large, arbitrage capital is slow-moving and fund flows explain returns of stocks, factors and asset classes, as well as return volatility and comovement.

While mutual fund flow-induced trading provides the best available natural experiment for changes in asset prices, the literature acknowledges that fund flows may not be exogenous to firm fundamentals because retail investors chase past performance (e.g. Edmans, Goldstein, and Jiang, 2012). In particular, the contemporaneous correlation between quarterly fund flows and the returns of the stocks in their portfolios could be driven by reverse causality: instead of outflows inducing fire sales, which drive down prices, poor stock returns reduce mutual fund returns, which in turn trigger outflows. As only low-frequency fund flows are easily observable and as fund flows and stock returns occur simultaneously from a low-frequency perspective, this issue is largely unexamined and unsolved. Therefore, this paper obtains high-frequency fund flows, investigates the relevance of the reverse causality problem, proposes a new method to isolate a component of flows that is plausibly exogenous to stock returns and uses that method to revisit classic empirical finance questions.

To begin, I show that the standard regression of stock returns on mutual fund flow-induced fire sales suffers from reverse causality if retail investors chase within-quarter mutual fund returns and mutual funds hold few stocks and/or stocks with correlated factor loadings. Empirically, I show that these conditions are met. In particular, while it is well known that retail investors chase past performance at low frequency (e.g. Chevalier and Ellison, 1997), I trace out retail investors' high-frequency purchase/redemption responses to mutual fund returns and find strong evidence for within-quarter return chasing.

In addition, I demonstrate the importance of the reverse causality issue at the stock instead of the mutual fund level. The literature typically examines the price impact of mutual fund flow-induced fire sales using the quarterly event study depicted in Figure 2.1, on the left. In the event quarter, when mutual funds that hold a stock receive extreme outflows, stock prices crash, followed by a slow recovery. But increasing the event study frequency to daily, as in the right panel, stock prices crash before the event; returns precede flows. Consistent with this interpretation, I also show that daily stock returns strongly Granger cause daily flow-induced trading in calendar time.

In addition to high-frequency return chasing, low-frequency return chasing can also be problematic. Fundamental shocks may drive future corporate decisions and fund flows. However, this endogeneity issue is easier to mitigate because researchers can control for past stock returns. I therefore emphasize the reverse causality issue.

I propose a solution inspired by the literature on the real effects of monetary policy, which also faces a reverse causality issue. At low frequency, instead of monetary policy affecting economic activity, the Fed may respond to changes in economic conditions. A popular solution is to extract monetary policy surprises using high-frequency data and then aggregating the high-frequency surprises to low frequency (Cochrane and Piazzesi, 2002; Gorodnichenko and Weber, 2016; Nakamura and Steinsson, 2018). Analogously, I isolate high-frequency *surprise flows* which I aggregate to quarterly surprise mutual fund flows. I then construct surprise flow-induced trading as a new instrument for stock returns.

Fund flows come from two sources: retail investor trades in response to new information and trades in response to random liquidity shocks (e.g., retail investors may redeem mutual fund shares to make a down payment on a mortgage). The former gives rise to the identification concerns, the latter, surprise flows, are orthogonal to firm fundamentals. If researchers can isolate surprise flows, they can construct a valid instrument for returns. In this paper, I extract surprise flows by orthogonalizing high-frequency mutual fund flows with respect to contemporaneous fund returns and with respect to a long history of mutual fund returns and flows. This means, when instrumenting stock returns using surprise flow induced trading, I only attribute stock returns that occur after mutual fund flows to the demand shock. Hence, the only threats to identification are omitted variables that trigger flows today and returns tomorrow, a much weaker concern because it only applies if retail investors have skill, which is unlikely because fund flows are dumb money (Frazzini and Lamont, 2008).

I use surprise flows to reevaluate important findings in the literature. There are two competing forces when comparing the price impact of flow-induced trading to surprise flow-induced trading. First, the price impact of surprise flow-induced trading is not inflated by the reverse causality mechanism. But second, fund managers likely trade more drastically in response to surprise flows

than in response to expected flows because they have less time to execute trades. Therefore, which force dominates, is an empirical question. I find that using the new instrument, the estimated price impact of fire sales shrinks by around one third compared to existing estimates but is still economically and statistically significant; mutual fund flow-induced fire sales have price impact, but equity markets are deeper than previously suggested and the corresponding instrument is weak.

Many papers drop modest fund flows because they may not trigger a fast trading response that moves prices (e.g. Coval and Stafford, 2007; Edmans, Goldstein, and Jiang, 2012). Because the choice of cutoff value is arbitrary, I estimate the price impact of surprise flow-induced fire sales for all combinations of negative and positive cutoffs. I find more extreme fund flows trigger more drastic fire sales that result in larger price impacts. Using all mutual fund flows, a 1% demand shock triggers a 0.2% return. But a 1% demand shock induced by the most extreme mutual fund flows triggers a 1% return. Further, in the spirit of the causal inference using machine learning literature (Mullainathan and Spiess, 2017), choosing optimal flow cutoffs that maximize the strength of the instrument generates a strong instrument. The optimal cutoffs focus on extreme flows in the bottom 7.5% and the top 5% of the flow distribution. This is similar to, but more extreme than in Coval and Stafford (2007), who use flows in the 1st and 10th decile. The new, strong instrument could be used to reevaluate the multitude of findings referenced above that are based on evidence from fund flows.

The reverse causality issue highlighted in this paper does not necessarily invalidate existing research findings. Many papers provide additional evidence unaffected by this critique. Hence, only the particular evidence involving fund flows should be reevaluated. A good example is Anton and Polk (2014), who also show unaffected evidence from the 2003 mutual fund flow scandal. The literature also highlights the stock price reversal that follows fire sale events as evidence for temporary price pressure (e.g. Coval and Stafford, 2007). While stock price reversal does suggest that the abnormal returns generated by flow-induced trading are at least partially the result of price pressure, it cannot show the absence of the reverse causality mechanism. In addition, the stock price reversal pattern may be explained by risk factor loadings (Wardlaw, 2020).

Further, the market feedback effect literature uses flows very uniformly, closely following Ed-

mans, Goldstein, and Jiang (2012). Outside this literature, papers use flows more heterogeneously, often making the reverse causality problem more subtle. For example, Greenwood and Thesmar (2011) relate stock return volatility to flow volatility and Anton and Polk (2014) relate stock return comovement to the absolute value of flows to mutual funds that hold both stocks. But the reverse causality issue still applies in higher moments. Volatile stock returns imply volatile mutual fund returns which cause volatile flows. And when two stocks crash, funds that hold both stocks suffer bad returns and, consequently, outflows. In addition, both papers lag the explanatory flow variable, but this does not fully eliminate the concern because higher moments of returns are persistent.

This paper is related to Berger (2019), who finds the instrument as constructed in the market feedback effect literature implies unbalanced treatment and control groups, indicative of selection bias. To solve this issue she proposes testing for market feedback effects in a homogeneous subsample. While this approach improves the test, it does not address the reverse causality problem at the center of this paper. Further, eliminating the reverse causality issue by taking surprise flows addresses the Berger (2019) concern: past returns and firm characteristics predict flow-induced trading, but not surprise flow-induced trading.

Also, Wardlaw (2020) documents a mechanical issue in the standard construction of the instrument in the market feedback effect literature. The standard instrument contains the lagged stock price in the numerator (dollar portfolio position) and the stock price in the denominator (dollar trading volume). Therefore, the instrument for returns mechanically contains returns. Only the market feedback effect literature is subject to this critique, because papers outside this literature do not typically scale the demand shock by dollar trading volume. I demonstrate that even the Wardlaw (2020) corrected instrument suffers from an economic problem that applies more widely. I then propose a solution and use the improved instrument in applications.

This paper focuses on the price impact of mutual fund flow-induced fire sales, but also contributes to the literature examining high-frequency fund flows (Edelen and Warner, 2001; Greene, Hodges, and Rakowski, 2007; Rakowski and Wang, 2009; Rakowski, 2010; Ben-Rephael, Kandel, and Wohl, 2011) by relating high-frequency fund flows and returns to fund portfolios, by expanding the prediction horizon and by using a new, larger panel dataset.

2.2 Empirical Strategy

I illustrate a simplified version of the environment in Figure 2.2. Stock returns form mutual fund returns which become public information after markets close and potentially trigger mutual fund flows the next day. Flows are also determined by random liquidity needs of retail investors. Redemptions and purchases settle after markets close. Mutual funds absorb these flows with their cash position and trade over the next days to rebalance it. This forced trading may cause stock returns. Note that this description is stylized and events may fall closer together for sophisticated agents. Now, cumulating daily flows and returns over a quarter hides their temporal ordering so that quarterly flows and returns seem to happen contemporaneously, no matter which happens first at the daily frequency. Hence, the impact of flow-induced trading on stock prices cannot be inferred from the contemporaneous quarterly correlation.

2.2.1 Mutual Fund Flow-Induced Trading

There are N firms, indexed by n. Stocks are held by I investors, indexed by i. I denote the time in days by t and the time in quarters by q. I make this distinction because the literature works with quarterly data, while retail investors observe and could potentially chase daily mutual fund returns. The goal is to test whether mutual fund flow-induced trading has a causal effect on asset prices. This is the price impact of fire sales, as well as the first stage in an instrumental variable regression of a corporate outcome on stock returns that tests for market feedback effects.

Mutual funds receive net dollar flows, $Flow_{i,q}$. They hold $A_{i,q}$ in assets under management (AuM), so relative flows are $flow_{i,q} = Flow_{i,q}/A_{i,q-1}$. Flows are shocks to the liability side and force funds to adjust the asset side. The ensuing trades are the demand shocks. However, starting with Edmans, Goldstein, and Jiang (2012), the literature recognizes that actual trades are endogenous. Hence, they construct hypothetical flow-induced trading

$$FIT_q(n) = \frac{\sum_{i=1}^{I} Shares_{i,q-1}(n) g(flow_{i,q})}{SharesOutstanding_{q-1}(n)}. \tag{2.1}$$

The numerator is the net amount of shares mutual funds buy or sell to accommodate flows. $g()$ maps flows into relative changes in held shares that move prices. I call $g()$ the fire sale function. For example, if funds keep portfolio weights constant and the demand elasticity is constant, $g()$ is simply the identity; in response to a 10% outflow, funds sell 10% of each portfolio position, resulting in a price impact proportional to the fraction of outstanding shares held by the fund. The market feedback effect literature uses the identity, except they drop flows greater than -5%, arguing that flows>-5% do not force funds to trade quickly and likely do not generate large price impacts. Outside of this literature, researchers use different fire sale functions. In the empirical part, I start by using the identity as the simplest benchmark but then examine all relevant candidates. The measure is scaled by shares outstanding to make it a relative demand shock. The literature also uses different denominators. In addition to shares outstanding, common choices are total shares held by mutual funds (e.g. Lou, 2012) and trading volume (e.g. Edmans, Goldstein, and Jiang, 2012). However, since the Wardlaw (2020) critique, using trading volume is no longer state-of-the-art. Hence, I only examine the first two.

I now turn from quarterly to daily frequency, because retail investors can make trading decisions daily and can use daily mutual fund returns to inform those decisions. This is because mutual funds publish their net asset value once a day, after markets close, and retail orders settle at that price. Of course, most retail investors review their portfolios infrequently and also base their investment decisions on long-term performance metrics. But this behavior does not give rise to the reverse causality problem. Instead, I ask whether the fraction of retail investors who do review their portfolio on any given day take within-quarter fund returns into account when making purchase/redemption decisions, because this is the potential source of reverse causality.

2.2.2 Decomposition of the Mutual Fund Flow Instrument

Where do mutual fund flows come from? First, retail investors face liquidity shocks. They may redeem mutual fund shares to make a down payment on a mortgage. This is the original idea behind the fund flow instrument. Random retail redemptions are exogenous to firm fundamentals but change asset prices by triggering fire sales. But in addition, it is well known that retail investors

chase past mutual fund performance (e.g Chevalier and Ellison, 1997). In a Berk and Green (2004) model, retail investors chase past performance because they use realized mutual fund returns, $MFret_{i,t}$, to update their beliefs about the ability of the mutual fund manager. I write this as

$$flow_{i,t} = \eta_t + \sum_{l=0}^{L} \alpha_l MFret_{i,t-l} + u_{i,t}, \tag{2.2}$$

where I allow for macro shocks to flows. Retail investors may react to information beyond returns. For example, they may redeem shares in response to redemptions by other retail investors because of strategic complementarities (e.g. Chen, Goldstein, and Jiang, 2010; Goldstein, Jiang, and Ng, 2017). For ease of notation, I omit these terms here, but include them later in the empirical part.

Retail investors observe mutual fund returns after markets close and can therefore only react to the exact fund return the next day. Yet, sophisticated retail investors may be able to infer the contemporaneous return and react the same day. Hence, to be conservative, I also allow for contemporaneous return chasing, i.e. the sum starts at zero. Next, for a stylized mutual fund that only holds equities, neither incurs trading costs nor receives lending revenues and only trades at the end of each quarter, the mutual fund return is the portfolio weight-weighted mean of stock returns

$$MFret_{i,t} = \sum_{n=1}^{N} w_{i,t-1}(n) r_t(n). \tag{2.3}$$

I make this simplification for ease of notation. Stock returns follow a linear asset pricing model. They are driven by factor returns, f_t, and idiosyncratic shocks

$$r_t(n) = \beta(n)' f_t + \epsilon_t(n). \tag{2.4}$$

Substituting equations 2.3 and 2.4 into equation 2.2, I decompose mutual fund flows into the following components:

$$
\begin{aligned}
flow_{i,t} &= \eta_t \\
&+ \sum_{l=0}^{L} \alpha_l w_{i,t-1-l}(n) r_{t-l}(n) \\
&+ \sum_{l=0}^{L} \alpha_l \sum_{m \neq n} w_{i,t-1-l}(m) \beta(m)' f_{t-l} \\
&+ \sum_{l=0}^{L} \alpha_l \sum_{m \neq n} w_{i,t-1-l}(m) \epsilon_{i,t-l} \\
&+ u_{i,t}.
\end{aligned} \tag{2.5}
$$

This decomposition shows why quarterly flow-induced trading is not exogeneous to contemporaneous quarterly stock returns. $FIT_q(n)$ is a function of the daily flows that fall into quarter q, and the first term above contains daily stock returns that fall into quarter q. This means the instrument for returns depends on returns, a reverse causality problem.

In addition, the first term implies that $FIT_q(n)$ depends on stock returns that occurred before the current quarter. While this is not a problem in the regression of stock returns on flow-induced trading, it does create an endogeneity problem in the market feedback effect regression of a corporate outcome on stock returns instrumented with flow-induced trading. Longer-term stock returns may be the result of fundamental shocks that influence corporate decisions in the future. Hence, both, the corporate outcome and the instrument may be driven by the fundamental shock, an endogeneity problem. Yet, while this endogeneity concern is generated by the same economic mechanism as the reverse causality issue, it is easier to mitigate because researchers can control for past stock returns. I therefore emphasize the reverse causality issue.

Note that if retail investors do not chase returns, $\alpha_l = 0$, these critiques do not apply. However, this is strongly rejected empirically. It is well known that retail investors chase low-frequency returns (e.g. Chevalier and Ellison, 1997) and in this paper, I show that retail investors chase within-quarter returns. Also, if portfolio weights are small, the direct dependence of the instrument on the stock return itself is weak. But that does not mean that constructing the instrument excluding

concentrated mutual funds solves the problem, because the second term is also problematic. The instrument depends on the systematic mutual fund return. Hence, if factor loadings are correlated within a portfolio, then factor returns drive both: they drive stock returns directly and they drive flows via mutual fund returns. Most stocks in the portfolio of a value fund load on HML. Accordingly, high HML returns at the beginning of a quarter cause high returns for the stocks in the portfolio, meaning high mutual fund returns that trigger inflows later in the quarter. Together, the first two terms show the instrument is invalid if retail investors chase within-quarter mutual fund returns and portfolios are concentrated and/or contain stocks with correlated factor loadings.

Terms three and four are exogenous to stock returns and can be used to construct demand shocks. Term three contains flows triggered by mutual fund returns stemming from idiosyncratic returns of other stocks in the portfolio. Say a fund receives outflows because it had poor returns due to a fire at GM's factory in a foreign country. This flow could be used to instrument the return of Goldman Sachs. In practice, these shocks are difficult to identify, because they depend on the asset pricing model. Hence, I do not use them in this paper.

Term four contains surprise flows, flows generated by random liquidity shocks to retail investors. I use this term to construct the new instrument. Specifically, I aggregate daily surprise flows to quarterly surprise flows and substitute flows with surprise flows when constructing the instrument. This gives surprise flow-induced trading

$$SFIT_q(n) = \frac{\sum_{i=1}^{I} Shares_{i,q-1}(n) g(surpriseflow_{i,q})}{SharesOutstanding_{q-1}(n)}, \tag{2.6}$$

where

$$surpriseflow_{i,q} = \frac{\sum_{t \in q} u_{i,t} A_{i,t-1}}{A_{i,q-1}}. \tag{2.7}$$

I extract daily surprise flows as the residuals of a cross-sectional regression motivated by equation 2.2. Daily mutual fund flows are the dependent variable. The independent variables are, first, 65

lags of mutual fund returns, including the contemporaneous fund return. 65 is the the maximum number of business days in a quarter. This solves the reverse causality problem outlined above. However, when using the flow instrument to test for market feedback effects on corporate decisions, additional variables may raise endogeneity concerns, as discussed above. I address this by adding control variables to equation 2.2. In addition to the 65 daily fund return lags, I include 65 daily fund flow lags and quarterly fund return and flow lags from one quarter to three years ago. I also estimate the regression separately for active and index mutual funds because the main return chasing motive, using returns to update beliefs about the fund manager's ability as in Berk and Green (2004), does not apply to index funds. This accounts for the finding in Dannhauser and Pontiff (2019) that the difference between active and index mutual funds explains a large share of the heterogeneity in the flow-return relation. By construction, surprise flows are orthogonal to past and contemporaneous mutual fund returns. This implies that I only attribute stock returns that occur after mutual fund flows to the demand shock. Hence, the only threats to identification are omitted variables that trigger flows today and returns tomorrow, a much weaker concern.

In future work, potential additional concerns about the endogeneity of fund flows can be addressed in this step. Any measure of surprise flows that isolates random liquidity needs of retail investors can be used to construct instruments that are exogenous to firm fundamentals.

2.3 Data

2.3.1 Mutual Fund Data

All data I use are standard with one exception: daily mutual fund flows and returns from Morningstar. I use the Morningstar variables *daily return index*, *estimated fund level net flow (comprehensive) (daily)* and *fund size - comprehensive (daily)* to construct daily flows and returns. The data become reliably available starting in July 2008, which dictates the start of my sample. When an exercise requires high-frequency Morningstar data, the sample is July 2008 to December 2017. Whenever it does not, I use the common 1980 to 2017 sample. I restrict the sample to equity mutual funds that are in the MFLinks database. These are the funds that I can match with CRSP mutual

fund characteristics and Thomson Reuters portfolios. The availability of MFLinks data prescribes the end of my sample. I merge Morningstar and MFLinks using share class cusips. Then, I collapse the data from the share class to the portfolio level by taking size-weighted means where necessary. There are about 3500 mutual funds in the average cross-section. Out of these, I can match about 2900 to daily Morningstar data. With ten years of daily data, this yields over six million observations. For comparison with the literature, I also construct quarterly mutual fund flows, returns and characteristics using CRSP.

I make minor modifications to the mutual fund flow data. First, I require a minimum fund size of $10 million. Next, a small number of observations show extreme spikes in assets under management that only last one day. I assume these are data entry mistakes. Therefore, I replace fund size with its lagged value where it implies last and next period flows with an absolute value greater than 25% with opposite signs. This leaves a few daily flows with an absolute value greater than 25%, less than 0.1% of observations. I set these outliers to missing because many of these are likely data errors. For the same reason, I set quarterly flows less than -100% or greater than 200% to missing. Lastly, a small number of daily observations is missing. As I estimate regressions with 65 daily lags, one missing observation would eliminate 65 observations. To avoid this loss of information, I fill missing lags with their cross-sectional mean.

2.3.2 Portfolio Holdings Data

The Thomson Reuters Mutual Fund Holdings database provides quarterly fund-level portfolios starting in 1980. The sources for this database are SEC mandated disclosures in Forms N-30D, N-Q and N-CSR as well as voluntary disclosures. I merge the Thomson Reuters holdings data with CRSP mutual fund data using MFLinks. I merge to the Thomson Reuters file date.

2.3.3 Stock Level Data

Stock data are from CRSP. Accounting data are from the CRSP/Compustat Merged Fundamentals Quarterly and Annual databases. Specifically, I use the stock characteristics corresponding to a standard 6-factor asset pricing model (Fama and French, 2018), beta, log market equity, log Tobin's

Q, profitability, investment and momentum as the standard control variables because they are the most prominent sources of variation in the cross-section of expected returns. To construct market beta, I take the 1-month T-bill rate and the market return from Kenneth French's website. I winsorize characteristics at the 1 and 99% levels as in Green, Hand, and Zhang (2017). I restrict the sample to US ordinary common stocks that trade on the NYSE, AMEX or Nasdaq, have non-missing returns and have at least 1% mutual fund ownership.

2.3.4 Summary Statistics

While the literature typically uses low-frequency fund flows from CRSP, I obtain high-frequency fund flows from Morningstar because CRSP does not provide high-frequency flows.[1] Table 2.1 compares Morningstar to CRSP data. Panel A reports summary statistics for each year-end quarter where Morningstar data are available, from 2008 to 2017. I compare Morningstar and CRSP data in terms of number of funds, median fund total net assets, total net assets and total absolute fund flows. In an average cross-section, the CRSP sample contains about 3500 funds and the Morningstar sample about 2900. By construction, the Morningstar data are a strict subsample. The number of funds decreases by about 10% after the financial crisis. Median total net assets increase from \$110 to \$280 million. They are similar but lower for Morningstar data. Total TNA increase from \$3 to \$7 trillion for CRSP and from \$2 to \$4.5 trillion for Morningstar data. Flows are more stable over time, meaning that relative flows decrease over time. For CRSP, total quarterly absolute fund flows are about \$220 billion; for Morningstar, they are about \$150 billion. Overall, high-frequency Morningstar data cover about three quarters of the low-frequency CRSP data.

Next, Panel B shows fund level instead of aggregate summary statistics. It summarizes the distribution of mutual fund flows and returns at the daily and quarterly frequency, as indicated by (D) and (Q). Daily flows and returns are centered around zero with a standard deviation of about 1%. Quarterly flows are centered around zero with a standard deviation of 19%. Quarterly returns are centered around the quarterly equity premium with a standard deviation of 9%. Returns are

[1] CRSP provides daily fund returns and net asset values. However, CRSP does not provide daily total net assets which are necessary to compute fund flows

largely uncorrelated in the time series, so the ratio of the quarterly and the daily standard deviation is about the square root of the number of business days in a quarter. In contrast to that, flows are autocorrelated, so the ratio is larger.

The previous section shows that the standard regression of stock returns on flow-induced trading suffers from reverse causality if there is within-quarter mutual fund return chasing and portfolios are concentrated and/or contain stocks with correlated factor loadings. Panel C indicates that both mutual fund portfolio conditions hold. Panel C summarizes the distribution of mutual fund portfolio characteristics for the last cross-section in my sample, 2017 Q4. The number of stocks in a portfolio varies substantially. From the 5th percentile, to the median, to the 95th percentile, mutual funds hold 11, 64 and 497 stocks, respectively. Comparing this to the about 4000 stocks in the cross-section, this means most mutual funds hold concentrated portfolios. Below, I show the distribution of portfolio characteristics that correspond to a standard 6-factor asset pricing model (Fama and French, 2018). Portfolio characteristics are the portfolio weight-weighted mean of the characteristic across all stocks in a portfolio. Again, there is substantial dispersion. The median fund has a market beta of 1.1. This is as in Frazzini and Pedersen (2014), who argue mutual funds use high market exposure because they are leverage-constrained. From the 5th to the 95th percentile, mutual fund betas vary from 0.9 to 1.4. Mutual funds also choose differential exposure to firm size. From the 5th percentile, to the median, to the 95th percentile, mutual funds hold stocks with market values of $2, $88 and $220 billion, respectively.

The wide dispersion in mean portfolio characteristics across mutual funds suggests mutual funds hold stocks with correlated factor loadings. Of course, this is not a formal test. I report the results of a formal test in Table 2.2. Mutual fund portfolio position level regressions of stock characteristics on the portfolio weight-weighted mean of the same characteristic over all other stocks in the same portfolio strongly reject that factor loadings are uncorrelated within portfolios for all characteristics corresponding to a standard 6-factor asset pricing model (Fama and French, 2018).

2.4 Results

2.4.1 Mutual Fund Flows and Mutual Fund Returns

Equation 2.5 shows that the standard regression of stock returns on flow-induced trading suffers from reverse causality if there is within-quarter mutual fund return chasing (and portfolios are either concentrated or contain stocks with correlated factor loadings; both are empirical facts as shown in the previous section). While it is well known that retail investors chase past performance at low frequency (e.g. Chevalier and Ellison, 1997), I trace out retail investors' high-frequency purchase/redemption responses to mutual fund returns. Figure 2.3 shows that retail investors do indeed chase within-quarter mutual fund returns. I report results for the regression that extracts daily surprise flows, as described in section 2.2, a regression of daily mutual fund flows on lags of daily flows and returns. I include 65 lags, the maximum number of business days in a quarter. The regression also controls for quarterly fund return and flow lags from one quarter to three years ago and for time fixed effects. Standard errors are clustered by time because, similar to returns, flows are correlated in the cross-section. I estimate the regression separately for active and index mutual funds because the main return chasing motive, using returns to update beliefs about the fund manager's ability as in Berk and Green (2004), does not apply to index funds. I show the active fund results here and the index fund results in appendix Figure 2.7. I plot cumulative coefficients in blue and display the p-values of each coefficient as gray bars.

The left panel of Figure 2.3 shows the cumulative coefficients on daily past mutual fund returns. The cumulative coefficients trace out a concave impulse response and reach 0.2; a 10% mutual fund return triggers a 2% mutual fund flow within one quarter. For the first 20 business day lags all coefficients are highly statistically significant. After that, they remain positive but are less precisely estimated. The coefficient on the contemporaneous mutual fund return is insignificant. This is likely because retail investors only observe the contemporaneous return after markets close and only react to it the next day. On the right, I show the cumulative coefficients on past daily mutual fund flows. Flows are highly autocorrelated. A 10% flow today predicts a 6% flow within one quarter. Yet, curiously, flows feature a one-day reversal, so that a 10% inflow today predicts a

1% outflow tomorrow, but all other coefficients are positive.

Appendix Figure 2.7 shows the analogous results for index mutual funds. Puzzlingly, but consistent with Dannhauser and Pontiff (2019), I find that index fund flows are more elastic to past performance than active fund flows. A 10% index mutual fund return triggers a 3% in comparison to a 2% mutual fund flow within one quarter. However, the estimates are much less precise because less than 15% of observations correspond to index funds. Overall, this exercise shows that retail investors engage in within-quarter return chasing, meaning the reverse causality issue is relevant.

Of course, this empirical argument's limitation is that it could itself be subject to a reverse causality critique. I attribute the fact that mutual fund returns predict future mutual fund flows to retail investor return chasing. However, if mutual fund managers had a crystal ball, they could trade prior to flows, which could generate the same empirical pattern. Yet, to overturn the reverse causality critique, this anticipatory trading must be the only reason for the observed predictive pattern and fund managers would need to engage in anticipatory trading far in advance, up to one quarter. Further, the following subsection shows that most of the flow-induced cumulative abnormal returns realize before extreme flows. Therefore, anticipatory trading needs to be far greater than reactive trading. This is unlikely, though, since most flows are unpredictable, as indicated by the 4% R^2 of the flow predictive regression reported in Figure 2.3.

2.4.2 Flow-Induced Trading and Stock Returns

In Figure 2.4, I show the corresponding analysis at the stock level instead of the mutual fund level. I report results from a regression of daily flow-induced trading on daily lags of flow-induced trading and stock returns. Daily FIT is constructed as in equation 2.1, except at the daily instead of the quarterly frequency. I find the same pattern as in the previous figure. Looking at the graph on the left, the cumulative coefficients on stock returns are positive and highly statistically significant for the first 20 lags. For ease of interpretation, I normalize FIT. The cumulative coefficients reach 4; a 10% stock return predicts an additional 0.4 standard deviations of FIT within one quarter. Stock returns Granger cause the instrument, which is additional evidence for the relevance of the reverse causality issue. Looking at the graph on the right, FIT is autocorrelated because flows are

autocorrelated.

The literature typically illustrates the impact of flow-induced trading in event time instead of calendar time. One graph is ubiquitous: an event study around flow-induced fire sale events. Prominent examples that show a version of the graph include Coval and Stafford (2007), Duffie (2010), Edmans, Goldstein, and Jiang (2012), Khan, Kogan, and Serafeim (2012), Lou (2012), Jotikasthira, Lundblad, and Ramadorai (2012) and Dessaint et al. (2018). I replicate the graph from Edmans, Goldstein, and Jiang (2012) in the left panel of Figure 2.1. The definition of the event is that FIT falls into the bottom decile of the full sample distribution. They compute abnormal returns by deducting the CRSP equal-weighted return. I add Fama-MacBeth standard errors following Coval and Stafford (2007). In the literature, the typical interpretation of the graph is that there is no significant pre-event trend, prices crash by about 5% during the event quarter and prices recover from the fire sale over the following two years. Together, these observations are taken as evidence that flow-induced fire sales drive asset prices away from fundamentals.

Wardlaw (2020) already shows that correcting a mechanical mistake attenuates the empirical pattern in the left panel of Figure 2.1. This paper highlights a different issue. In the right panel of Figure 2.1, I show the result of the event study at the daily level, after making the Wardlaw (2020) correction. The event window is now one month instead of three years. The high-frequency event study shows that prices crash before the event. In fact, the price collapse largely stops at exactly the onset of the event. This is additional evidence that the reverse causality problem is relevant.[2]

The reverse causality problem can explain several nuances in the popular graph on the left of Figure 2.1. First, there is a pre-event trend. Stock returns over the last year form mutual fund returns over the last year which trigger flows. Second, the return during the event quarter is convex because stock returns on the first days of a quarter have more time to trigger flows than stock returns that occur at the end of the quarter. If the causality ran from flows to returns, we would expect the opposite. Stock returns on the first days of the quarter can only be driven down by flows from these first days, while stock returns during the end of the quarter follow many days of extreme

[2]In the daily event study the CAAR is an order of magnitude smaller than in the quarterly event study. This is because the daily shock is about an order of magnitude smaller than its quarterly analogue. Hence, the implied elasticities are similar.

flows. Finally, as pointed out by Wardlaw (2020), the price reversal does not stop at a CAAR of 0. The "recovery" is largely a relic of how abnormal returns are constructed. The sample of stocks held by mutual funds has higher average returns than the CRSP equal-weighted return and event stocks have size and value exposure.

2.4.3 The New Instrument - Surprise Flow-Induced Trading

As managers make corporate decisions at low frequency, market feedback effect regressions require a low-frequency instrument. Accordingly, I now turn from daily to quarterly frequency. The low-frequency instruments are constructed as described in equation 2.1 for the traditional instrument and as described in equations 2.6 and 2.7 for the new instrument. Table 2.3 estimates the price impact of fund flow-induced fire sales using the traditional and the new instrument. There are two competing forces when comparing the price impact of flow-induced trading to surprise flow-induced trading. First, the price impact of surprise flow-induced trading is not inflated by the reverse causality mechanism. But second, fund managers likely trade more drastically in response to surprise flows than in response to expected flows because they can liquidate positions in anticipation of expected flows but, by definition, not in anticipation of surprise flows. Therefore, which force dominates, is an empirical question. Table 2.3 reports results for regressions of quarterly stock returns on different versions of the instrument, controls and time fixed effects. As stock returns are highly correlated in the cross-section but not in the time series, I cluster standard errors by time. This regression is also the first stage of the common instrumental variable regression of a corporate outcome on returns.

In column 2, I construct FIT as in the literature. The estimated coefficient is 0.3 and highly statistically significant. It suggests a 1% demand shock triggers a 0.3% return. In column 4, I change my data source from CRSP to Morningstar but use the same object. I show this intermediate step to ensure the results are not driven by the change in data source and sample period. Now the coefficient is 0.37 and highly statistically significant, but less precisely estimated. This is because I lose over 70% of observations as there are only 10 years of Morningstar data in comparison to 37 years of CRSP data.

While the regressions in columns 1 to 4 are subject to the reverse causality critique, the following regressions employ the new, corrected instrument. In column 6, the coefficient drops to 0.23, but remains statistically significant. Demand is more elastic. A 1% demand shock triggers a 0.23% instead of a 0.37% return. The Kleibergen and Paap (2006) F-statistic of this regression is 5, meaning the instrument is weak. Columns 1, 3 and 5 show the conclusions do not change when not controlling for common risk factors. Overall, correcting the reverse causality problem indicates that markets are deeper than suggested. In column 7, I regress the old on the new instrument. By construction, the two variables are highly correlated. The estimated coefficient is 1.2 with a standard error of 0.07 and the R^2 is 52%.

When constructing flow-induced trading, I need to map mutual fund flows into trades that move prices. This is the purpose of $g()$ in equation 2.1. The literature proposes several candidate mappings which I illustrate in Figure 2.5. This figure also shows mutual funds' actual trading response to flows in red; I show position-weighted mean trading, the relative change in shares held, for 50 mutual fund flow bins. The mapping used in the analysis above corresponds to the gray 45 degree line, which is the simplest option. This mapping assumes mutual funds keep portfolio weights constant, scaling their entire portfolio up or down in response to flows. This is a rough approximation to mutual funds' actual trading behavior, but the actual trading response is stronger for outflows than for inflows. Naturally, mutual funds sell stocks they own, but they may initiate new positions. Lou (2012) accommodates this by scaling inflows and outflows with the empirically observed slopes. I show the resulting mapping in purple. By construction, it is a closer fit to realized trading.

However, as the goal is to construct a nonfundamental shock to asset prices, realized trading is not necessarily the relevant target. While modest flows trigger trades in the medium run, these trades can be smoothed using a fund's cash position to minimize price impact. Hence, only large flows that force funds to trade quickly trigger fire sales and thus have significant price impact. This is the reasoning in Coval and Stafford (2007) and Edmans, Goldstein, and Jiang (2012). I show the fire sale function from Coval and Stafford (2007) in green. They only use flows in the

extreme deciles of the mutual fund flow distribution.[3] In blue, I depict the mapping from Edmans, Goldstein, and Jiang (2012) which is standard in the market feedback effect literature and only uses flows less than -5%. Both mappings can be written as

$$\tilde{g}(flow_{i,q}) = \begin{cases} flow_{i,q} & \text{if } flow_{i,q} \leq \alpha_L \text{ or } flow_{i,q} \geq \alpha_H \\ 0 & \text{if } \alpha_L < flow_{i,q} < \alpha_H. \end{cases} \tag{2.8}$$

The construction of flow-induced trading contains many degrees of freedom. Hence, I report results of cross-sectional return predictive regressions for all combinations in Table 2.4. There are two choices, the fire sale function and the denominator. For each combination, I show the same quarterly, cross-sectional return regressions as in Table 2.3, for the traditional versus the new instrument and excluding and including controls. I report t-statistics in parentheses to allow the reader to quickly determine instrument strength and statistical significance. The results are similar to the results in table 2.3. Using the traditional instrument in the left two columns, the estimated coefficients are large and strongly statistically significant in all specifications. The two columns on the right report results for the new instrument. The estimates tend to be smaller and are consistently less statistically significant. Though when focusing on extreme flows using the Edmans, Goldstein, and Jiang (2012) or Coval and Stafford (2007) fire sale function and scaling FIT by shares observed in Panel A, the estimated price impact is similar.

In panel B, I scale FIT by shares outstanding instead of shares held by mutual funds. Therefore, many estimates are an order of magnitude larger than in panel A because the denominator is an order of magnitude larger. In Panel B, using the corrected instrument and except for the Coval and Stafford (2007) fire sale function, all estimated coefficients are much smaller than their uncorrected counterparts and statistically insignificant. Using the Coval and Stafford (2007) fire sale function, the estimated coefficients are 20% smaller and remain significant at the 10% level.

Overall, after correcting the reverse causality problem, I find that mutual fund flow-induced fire

[3] As Coval and Stafford (2007) use actual trades instead of hypothetical trades, their exact measure is not nested by equation 2.1. However, starting with Edmans, Goldstein, and Jiang (2012) the literature recognizes that actual trades are endogenous. Hence, I update the Coval and Stafford (2007) measure to reflect this.

sales have price impact but this price impact is smaller and generates weak instruments.

2.4.4 Which Mutual Fund Flows Trigger Fire Sales/Purchases?

The exact cutoffs in equation 2.8 from Coval and Stafford (2007) and Edmans, Goldstein, and Jiang (2012) are arbitrary. Therefore, I estimate the price impact of surprise flow-induced fire sales for all combinations of negative and positive cutoffs. This allows me to select the optimal $\tilde{g}()$. I pick the negative and positive flow cutoffs to maximize instrument strength. I pick them separately because there is no reason flows should map into forced trading symmetrically.

The heatmaps in Figure 2.6 report coefficient estimates and Kleibergen and Paap (2006) F-statistics for the grid search that picks the optimal cutoffs. Every grid cell corresponds to one quarterly cross-sectional regression of stock returns on surprise flow-induced trading and control variables. The positive percentile cutoff is on the x-axis and the negative percentile cutoff is on the y-axis. So the (0.9, 0.1) value corresponds to the Coval and Stafford (2007) fire sale function. The squares' color and size indicate the magnitude of the coefficient estimates or F-statistics corresponding to the respective grid cell. I use a grid with a step size of 2.5%.

I flip the y-axis to make the graph more intuitive. Now the regression corresponding to the bottom left does not drop any flows. It is the same as in Table 2.3; accordingly, the coefficient and F-statistic are small. The top right is the opposite extreme: I drop all flows and therefore obtain missing estimates by construction. From the bottom left to the top right, the focus on extreme flows increases, as do the coefficient estimates in the left panel. More dramatic fund flows trigger more drastic fire sales that result in larger price impacts. Using all flows, a 1% demand shock triggers a 0.2% return, but a 1% demand shock induced by the most extreme mutual fund flows triggers a 1% return.

For instrument strength in the right panel, a second force is at play. While focusing on more extreme flows increases the price impact of the ensuing fire sales, it also decreases the number of experiments. Hence, there is an interior optimum. I find choosing flows in the bottom 7.5% and the top 5% maximizes instrument strength, with an F-statistic of 27, implying that optimal flow cutoffs generate a strong instrument. These cutoffs are similar to, but more extreme than in Coval and

Stafford (2007), who use flows in the bottom and top 10%. Looking at the summary statistics in Table 2.1, the bottom 7.5% and the top 5% of the fund flow distribution correspond to fund flows of about -10% and 27%. This choice shares qualitative similarities with Edmans, Goldstein, and Jiang (2012) who use flows less than -5%. The negative cutoff is broadly similar and the effect of inflows and outflows is asymmetric. Inflows need to be substantially larger than outflows to induce a strong trading response that impacts asset prices. This is because funds have to sell stocks they own to accommodate outflows, but they may use inflows to initiate new positions. Appendix Figure 2.8 demonstrates that the results are qualitatively the same when scaling surprise flow-induced trading by shares outstanding instead of by total shares held by mutual funds, with a maximum F-statistic of 48.

Overall, the new fire sale function makes the instrument strong because it focuses on more extreme flows. This new, strong instrument could be used to reevaluate the multitude of findings referenced in the Introduction that are based on evidence from fund flows.

2.4.5 Robustness

Naturally, testing $(0.5/0.025)^2 = 400$ specifications raises data mining concerns. This is why the analysis above reports both: results from the data mined instrument and results from instruments as constructed in the literature. In addition, the maximum t-statistics of 5 and 7 are highly unlikely under the null, even with 400 tests. As a very conservative upper bound, a Bonferroni correction raises the 1% critical t-value to 4.2. Further, this multiple testing concern is mitigated by the fact that the tests are highly correlated and economically motivated. Finally, when the goal is to construct an instrument, there is no good alternative: either the instrument is constructed as in the literature but weak or the instrument is data mined but strong. Overall, even though correcting the reverse causality problem weakens the instrument, choosing the optimal fire sale function yields a strong instrument.

The new instrument is constructed from high-frequency surprise flows instead of from low-frequency flows and hence not driven by retail investor return chasing. While this is true by construction, I illustrate it further in Table 2.5. The table reports regressions of different versions

of the instrument on lags of stock returns and control variables. Column 1 uses the traditional instrument constructed using CRSP data. Naturally, the coefficients on the four lags of stock returns are positive and statistically significant. Column 2 also uses the traditional instrument, but now constructed using Morningstar data. I show this intermediate step to clarify that the results are not driven by the change in data source. Again, returns predict the instrument. The last column uses the new, corrected instrument. Now the coefficients on lagged stock returns have mixed signs and are statistically insignificant. As expected, stock returns do not predict the corrected instrument.

In addition, the instrument is uncorrelated with firm characteristics, mitigating the concerns raised in Berger (2019). All five firm characteristics corresponding to a Fama and French (2015) 5-factor asset pricing model, market beta, size, Tobin's Q, profitability and investment, predict the traditional, but not the corrected instrument. [4]

Next, the literature highlights the stock price reversal that follows fire sale events as evidence for temporary price pressure (e.g. Coval and Stafford, 2007). However, Wardlaw (2020) finds the reversal pattern is explained by risk factor loadings. I examine whether flow-induced trading and surprise flow induced-trading are followed by price reversal in Table 2.6, where I show cross-sectional, predictive regressions of stock returns on lags of different versions of flow-induced trading, controlling for risk factors and time fixed effects. Column 1 uses the traditional instrument constructed using CRSP data. The coefficients on the four lags of flow-induced trading have mixed signs and are statistically insignificant. This confirms the finding in Wardlaw (2020). Once accounting for risk factor loadings and time fixed effects, there is no stock price reversal.

Column 2 also uses the traditional instrument, but now constructed using Morningstar data. I show this intermediate step to clarify that the results are not driven by the change in data source. The results are qualitatively the same. The last column uses the new, corrected instrument. Again, the results are qualitatively the same. I do not find evidence for stock price reversal. However, this does not suggest the absence of price pressure effects. In fact, the index/additions deletions literature finds no reversals in response to asset demand shocks (e.g. Chang, Hong, and Liskovich,

[4] I do not include momentum because momentum is close to collinear with the lagged returns. Momentum is not exactly collinear with the four lagged quarterly returns because it omits the most recent month.

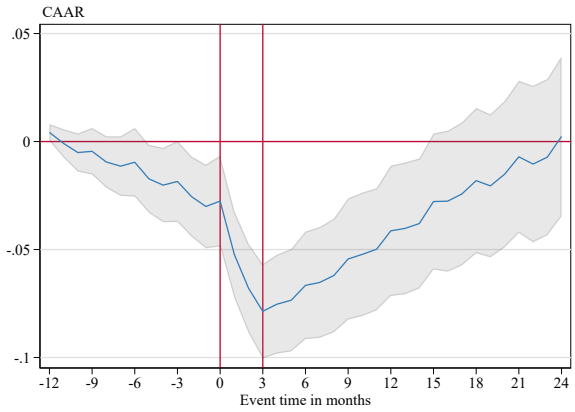
2015). Therefore, the absence of price reversals suggests that mutual fund fire sales have permanent price impact.

2.5 Conclusion

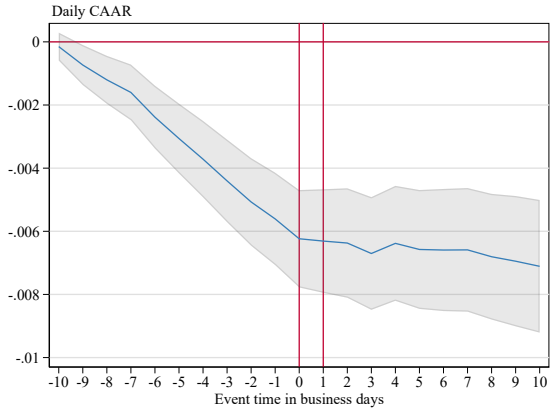
Based on evidence from mutual fund flows, an extensive literature finds equity markets are shallow and distort corporate decisions. This paper investigates to what extent these findings could be driven by reverse causality: instead of mutual fund outflows inducing fire sales, which drive down prices, poor stock returns reduce mutual fund returns, which in turn trigger outflows. I show this critique applies if retail investors chase within-quarter mutual fund returns and mutual funds hold few stocks and/or stocks with correlated factor loadings. Empirically, I show these conditions are met. The reverse causality issue also becomes intuitively apparent when increasing the frequency of the standard event study from quarterly to daily; returns precede flows. The solution is to construct a new instrument for returns from high-frequency surprise mutual fund flows. Correcting the instrument, the estimated price impact of mutual fund flow-induced fire sales are smaller but still positive and largely statistically significant. Markets are deeper than suggested.

Figures

Figure 2.1: Quarterly versus Daily Mutual Fund Flow-Induced Fire Sale Event Study



(a) Standard, low-frequency event study



(b) New, high-frequency event study

The left graph replicates the event study in Edmans, Goldstein, and Jiang (2012). An event is a stock level observation in which flow-induced trading (FIT) is in the bottom decile of its full sample distribution. The event lasts one quarter, from months 0 to 3. The authors compute cumulative average abnormal returns (CAAR) by deducting the CRSP equal-weighted index return and plot CAARs from one year before to two years after event onset. The sample is 1980 to 2017. The right figure is the analogue at the daily level, except that it takes into account the Wardlaw (2020) critique, meaning FIT is constructed as in equation 2.1 and CAARs are computed by demeaning returns cross-sectionally. Daily fund flows are an order of magnitude smaller than quarterly fund flows. To accomodate this while staying close to Edmans, Goldstein, and Jiang (2012), I define extreme outflows as flows less than -0.5% instead of -5% and use the bottom 1st percentile instead of the bottom decile. The results are robust with respect to these choices. The time is in business days, so the event window is one month. The sample is 2008 to 2017 because of the availability of daily Morningstar data.

Figure 2.2: Illustration of the Reverse Causality Problem

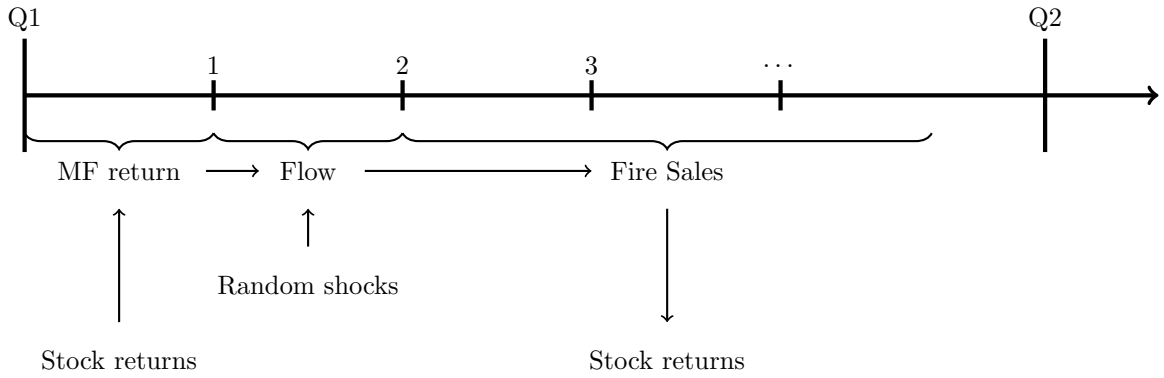
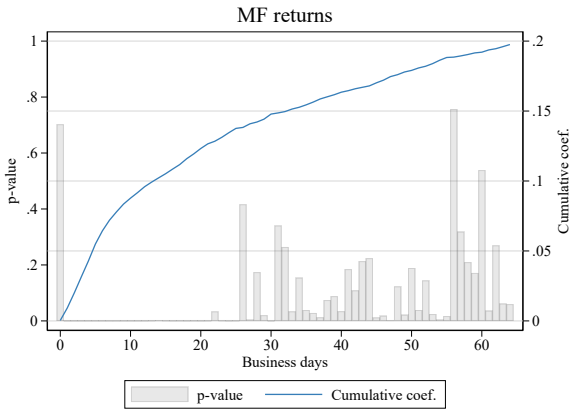
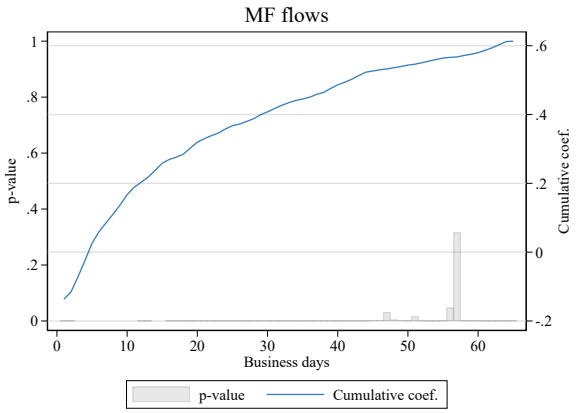


Illustration of the relationship between mutual fund flows and stock returns. I show events along a daily time series. Q1 and Q2 denote the beginning of a quarter. Mutual fund flows force mutual funds to trade which may have price impact. Flows can come from two sources. First, they may be generated by retail investors' random liquidity needs. In this case, the flow happens first, followed by stocks returns. Second, fund flows can result from performance chasing, and past performance is the result of the past returns of stocks in the mutual fund portfolio. In this case, stock returns happen first, followed by fund flows. As cumulating daily fund flows and stock returns over the quarter hides this temporal order, the impact of flow-induced trading on stock prices cannot be inferred from the contemporaneous, quarterly correlation.

Figure 2.3: Regression of Daily Mutual Fund Flows on Lags of Flows and Returns



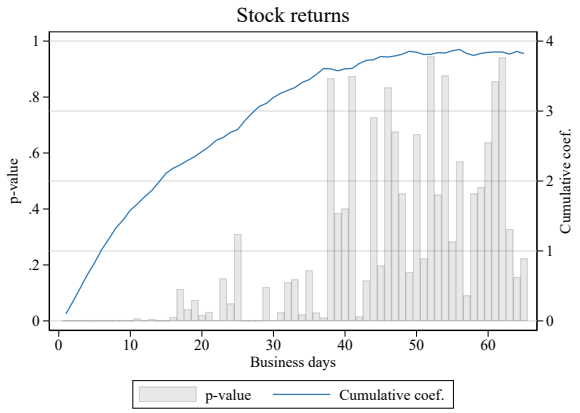
(a) Cum. coef. on lags of MF returns



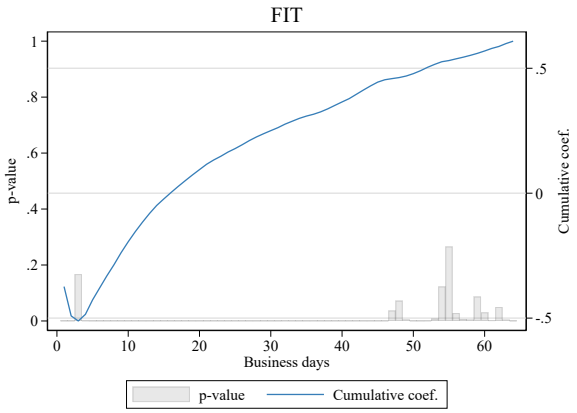
(b) Cum. coef. on lags of MF flows

This figure shows the results from a cross-sectional, predictive regression of daily mutual fund flows on lags of daily mutual fund returns and lags of daily mutual fund flows. It also controls for quarterly fund return and flow lags from one quarter to three years ago. I include time fixed effects and cluster standard errors by time. I include 65 daily lags, the maximum number of business days in a quarter. I plot cumulative coefficients in blue and display the p-values of each coefficient as gray bars. The sample excludes index mutual funds. The number of observations is 5,352,989. The R^2 is 0.0430. The within-R^2 is 0.0404. The sample is 2008 to 2017 because of the availability of daily Morningstar data.

Figure 2.4: Regression of Daily Flow-Induced Trading on Lags of FIT and Stock Returns



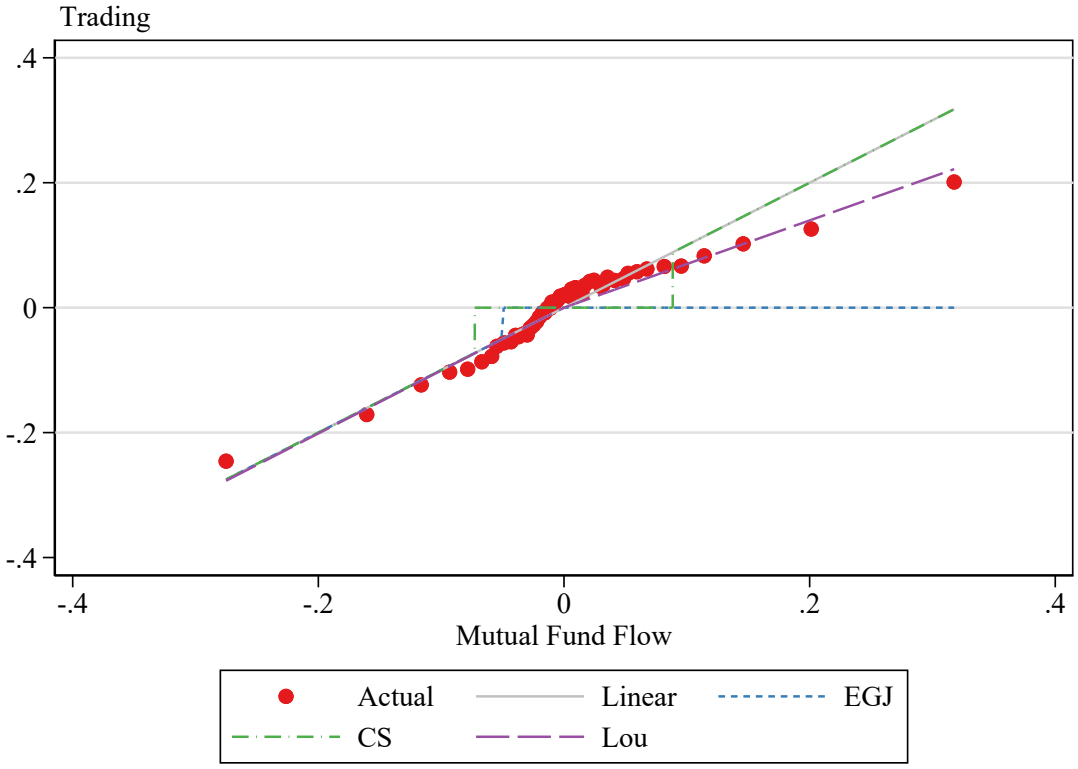
(a) Cum. coef. on lags of stock returns



(b) Cum. coef. on lags of FIT

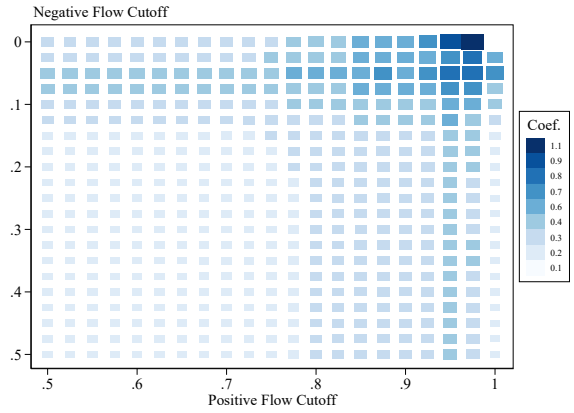
This figure shows the results from a cross-sectional, predictive regression of daily flow-induced trading (FIT) on lags of stock returns and FIT. This is the analogue to the regression in figure 2.3 at the stock level, rather than the mutual fund level. For ease of interpretation, daily FIT is normalized in the cross-section. I include time fixed effects and cluster standard errors by time. I include 65 lags, the maximum number of business days in a quarter. I plot cumulative coefficients in blue and display the p-values of each coefficient as gray bars. The number of observations is 5,664,001. The R^2 is 0.2279. The within-R^2 is 0.1398. The sample is 2008 to 2017 because of the availability of daily Morningstar data.

Figure 2.5: Mutual Fund Trading in Response to Mutual Fund Flows

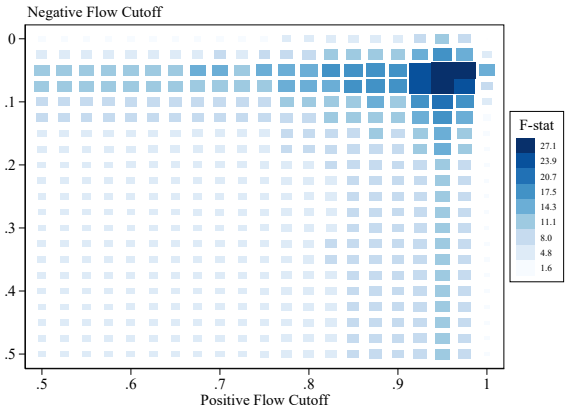


This figure shows how mutual funds trade in response to flows. Trading is the relative change in shares held. In red, I show demeaned position-weighted mean trading, for 50 mutual fund flow bins. I also plot four mappings of flows into trading from the literature. A basic approach is to use the identity which assumes mutual funds keep portfolio weights constant in response to flows. This corresponds to the grey 45 degree line. Following Edmans, Goldstein, and Jiang (2012), the market feedback effect literature makes this choice, except it discards flows greater than -5%. This is depicted in blue. Finally, I also show the mappings used in Coval and Stafford (2007) in green and Lou (2012) in purple. The sample is 1980 to 2017.

Figure 2.6: Regressions of Quarterly Stock Returns on Surprise Flow-Induced Trading for Different Flow Cutoffs



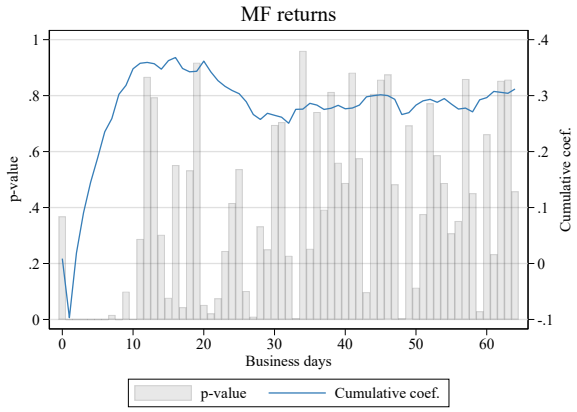
(a) Estimated coefficients



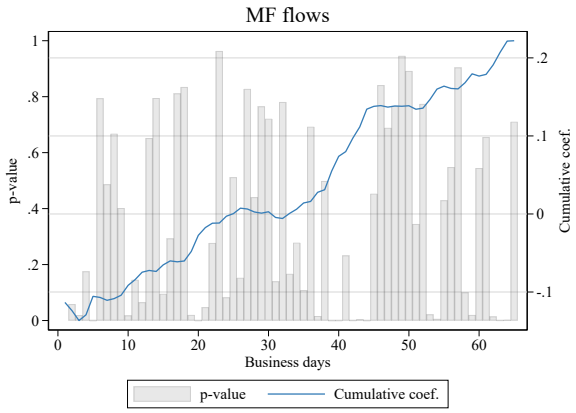
(b) F-statistics

The two heatmaps report coefficient estimates and Kleibergen and Paap (2006) F-statistics for quarterly cross-sectional regressions of stock returns on surprise flow-induced trading and control variables. Surprise flow-induced trading is constructed for all combinations of positive and negative flow cutoffs, according to equation 2.8. The axis give the percentile cutoffs; i.e. for each (x,y), I only use extreme surprise flows that are either in the top x or the bottom y percentile. The y-axis is reversed so that the bottom left corner drops all flows and the top right corner keeps all flows. The color and size of the squares indicate the magnitude of the F-statistic or estimated coefficient corresponding to the respective grid cell. Note that the heatmap values are not exact because they bin grid cells. I use a grid with a step size of 2.5%. The control variables are the characteristics corresponding to a standard 6-factor asset pricing model (Fama and French, 2018). The sample is 2008 to 2017 because constructing surprise flows requires daily Morningstar data.

Figure 2.7: Regression of Daily Mutual Fund Flows on Lags of Flows and Returns
Index Funds Only



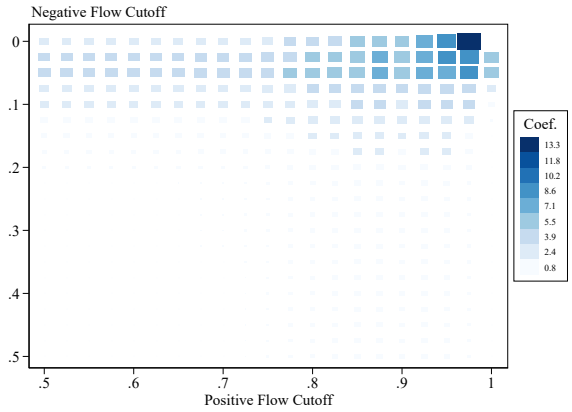
(a) Cum. coef. on lags of MF returns



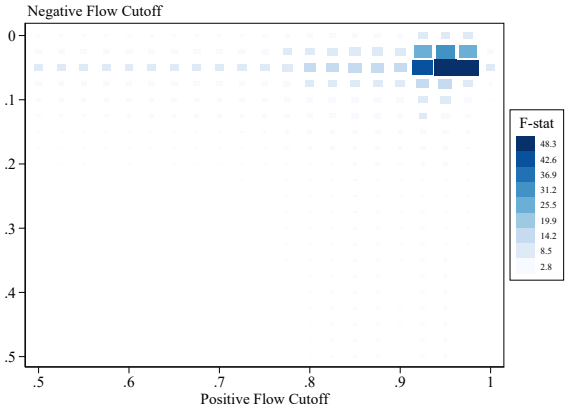
(b) Cum. coef. on lags of MF flows

This figure shows the results from a cross-sectional, predictive regression of daily mutual fund flows on lags of daily mutual fund flows and returns. This is the regression for index mutual funds. It also controls for quarterly fund return and flow lags from one quarter to three years ago. I include time fixed effects and cluster standard errors by time. I include daily 65 lags, the maximum number of business days in a quarter. I plot cumulative coefficients in blue and display the p-values of each coefficient as gray bars. The number of observations is 670,962. The R^2 is 0.0302. The within-R^2 is 0.0245. The sample is 2008 to 2017 because of the availability of daily Morningstar data.

Figure 2.8: Regressions of Quarterly Stock Returns on Surprise Flow-Induced Trading for Different Flow Cutoffs - Scaled by Shares Outstanding



(a) Estimated coefficients



(b) F-statistics

The two heatmaps report coefficient estimates and Kleibergen and Paap (2006) F-statistics for quarterly cross-sectional regressions of stock returns on surprise flow-induced trading and control variables. The difference to Figure 2.6 is that surprise flow-induced trading is constructed by scaling by shares outstanding instead of total shares held by mutual funds. Surprise flow-induced trading is constructed for all combinations of positive and negative flow cutoffs, according to equation 2.8. The axis give the percentile cutoffs; i.e. for each (x,y), I only use extreme surprise flows that are either in the top x or the bottom y percentile. The y-axis is reversed so that the bottom left corner drops all flows and the top right corner keeps all flows. The color and size of the squares indicate the magnitude of the F-statistic or estimated coefficient corresponding to the respective grid cell. Note that the heatmap values are not exact because they bin grid cells. I use a grid with a step size of 2.5%. The control variables are the characteristics corresponding to a standard 6-factor asset pricing model (Fama and French, 2018). The sample is 2008 to 2017 because constructing surprise flows requires daily Morningstar data.

Tables

Table 2.1: Summary Statistics

Panel A: CRSP versus Morningstar mutual fund coverage over time

	No. Funds		Median TNA ($M)		Total TNA ($T)		Total Flows ($B)	
Year	CRSP	Morningstar	CRSP	Morningstar	CRSP	Morningstar	CRSP	Morningstar
2008	3805	3221	110	97	3.1	2.1	220	121
2009	3533	2975	176	156	4.0	2.7	166	122
2010	3413	2866	224	201	4.4	3.0	199	137
2011	3328	2778	210	185	4.2	2.8	198	135
2012	3229	2690	244	206	4.7	3.0	212	126
2013	3287	2729	280	238	5.9	3.8	235	160
2014	3438	2877	275	231	6.3	4.1	232	164
2015	3604	3015	238	196	6.0	3.9	258	180
2016	3576	2986	241	199	6.3	4.2	268	189
2017	3417	2842	284	241	7.1	4.5	249	164

Panel B: Distribution of mutual fund flows and returns

	p5	p10	p25	p50	p75	p90	p95	sd
Flows (D)	-0.3	-0.2	-0.1	-0.0	0.0	0.2	0.4	1.1
Returns (D)	-1.9	-1.2	-0.4	0.1	0.5	1.2	1.8	1.3
Flows (Q)	-13.0	-8.0	-3.8	-1.0	3.2	13.4	27.4	19.1
Returns (Q)	-14.0	-8.7	-0.8	2.8	6.4	11.5	14.5	8.5

Panel C: Distribution of mutual fund portfolio characteristics in 2017 Q4

	p5	p10	p25	p50	p75	p90	p95	sd
# Stocks held	11.0	22.0	36.0	64.0	114.0	342.0	497.0	302.1
Market beta	0.9	0.9	1.0	1.1	1.2	1.3	1.4	0.2
Market equity	2.2	2.8	11.6	88.2	158.4	200.3	222.1	79.9
Tobin's Q	2.1	2.5	3.5	5.2	6.7	8.5	9.3	2.3
Profitability	23.4	31.6	42.1	51.1	57.1	62.4	65.6	13.5
Investment	2.4	4.0	6.5	9.4	13.2	16.7	19.1	6.0
Momentum	6.5	10.5	15.2	20.8	27.1	32.7	35.8	9.7

Panel A shows equity mutual fund summary statistics for each last quarter of a year, with CRSP and Morningstar data shown separately. I report the number of funds, their median total net assets (in million USD), the sum of their total net assets (in trillion USD) and the sum of absolute dollar mutual fund flows over the quarter (in billion USD). Panel B summarizes the distribution of mutual fund flows and returns at the daily and quarterly frequency, as indicated by (D) and (Q). For panels A and B, the sample is 2008 to 2017 because the daily mutual fund flow data from Morningstar start in 2008. Panel C shows the cross-sectional distribution of mutual fund portfolio characteristics corresponding to a standard 6-factor asset pricing model (Fama and French, 2018) and the number of stocks held. Portfolio characteristics are the portfolio weight-weighted mean of the characteristic across all stocks in a portfolio. I show these statistics for the last cross-section in my sample, 2017 Q4.

Table 2.2: Do Mutual Funds Hold Stocks with Correlated Characteristics?

	(1) Beta	(2) MV	(3) Q	(4) Profit	(5) Investment	(6) Momentum
PF mean (other)	0.654***	1.068***	0.863***	0.852***	0.736***	0.744***
	(0.00197)	(0.00372)	(0.00177)	(0.00275)	(0.00196)	(0.00226)
R-squared	0.0405	0.421	0.126	0.0998	0.0560	0.0562
N	26394754	26394754	26394754	26394754	26394754	26394754

This table shows regressions at the mutual fund portfolio position level. I regress the respective stock characteristic on the portfolio weight-weighted mean of the same characteristic over all other stocks in the same portfolio. The characteristics are market beta, market value, log Tobin's Q, profitability, investment and momentum, corresponding to a 6-factor asset pricing model (Fama and French, 2018). All regressions include time fixed effects. I report standard errors clustered by mutual fund × time in parentheses. The sample is 1980 to 2017. ***, **, and * denote significance at the 1%, 5% and 10% levels.

Table 2.3: Regressions of Quarterly Returns on (Surprise) Flow-Induced Trading

	(1) ret	(2) ret	(3) ret	(4) ret	(5) ret	(6) ret	(7) FIT
FIT(CRSP)	0.297***	0.287***					
	(0.0352)	(0.0332)					
FIT			0.381***	0.369***			
			(0.0818)	(0.0652)			
SFIT					0.266**	0.232**	1.215***
					(0.112)	(0.105)	(0.0662)
Controls	No	Yes	No	Yes	No	Yes	No
F-statistic	71.22	74.33	21.71	31.96	5.638	4.896	336.3
R-squared	0.197	0.207	0.245	0.247	0.243	0.246	0.515
N	409203	409203	108597	108597	108597	108597	110157

This table shows regressions of quarterly stock returns on different versions of flow-induced trading (FIT) and control variables. In columns 1 and 2, I construct FIT as in the literature. In columns 3 and 4, I change the data source from CRSP to Morningstar but use the same object. I show this intermediate step to clarify that the results are not driven by the change in data source. The number of observations decreases because there are only 10 years of Morningstar data in comparison to 37 years of CRSP data. In columns 5 to 7, I switch to the new, corrected instrument. The samples are 1980 to 2017 in columns 1 and 2, and 2008 to 2017 in columns 3 to 7, because of the availability of daily Morningstar data. The control variables are the characteristics corresponding to a standard 6-factor asset pricing model (Fama and French, 2018). All regressions include time fixed effects. I report standard errors clustered by time in parentheses and Kleibergen and Paap (2006) F-statistics in the table footer. ***, **, and * denote significance at the 1%, 5% and 10% levels.

Table 2.4: Regressions of Quarterly Returns on (Surprise) Flow-Induced Trading

	Flows		Surprise Flows	
	No Controls	Controls	No Controls	Controls
Panel A: FIT scaled by shares observed				
Flows	0.381***	0.369***	0.266**	0.232**
	(4.659)	(5.654)	(2.374)	(2.213)
Lou (2012)	0.448***	0.432***	0.291**	0.235*
	(4.782)	(5.591)	(2.318)	(1.973)
Edmans et al. (2012)	0.522***	0.487***	0.584***	0.493***
	(4.043)	(4.391)	(3.212)	(2.825)
Coval, Stafford (2007)	0.419***	0.416***	0.476***	0.452***
	(4.520)	(5.499)	(3.733)	(3.634)
Panel B: FIT scaled by shares outstanding				
Flows	2.777***	2.751***	-0.106	-0.348
	(5.568)	(6.149)	(-0.075)	(-0.262)
Lou (2012)	2.461***	2.690***	-1.196	-0.979
	(4.605)	(5.522)	(-0.719)	(-0.620)
Edmans et al. (2012)	1.360*	1.928***	0.050	1.000
	(1.892)	(3.129)	(0.018)	(0.388)
Coval, Stafford (2007)	3.732***	3.688***	3.127*	2.976*
	(5.406)	(6.258)	(1.952)	(1.902)

This table shows regressions of quarterly stock returns on different versions of flow-induced trading (FIT) and control variables. Columns 1 and 2 use mutual fund flows; columns 3 and 4 use surprise mutual fund flows to construct FIT. The regressions in columns 1 and 3 do not include control variables, while the other regressions do. The control variables are characteristics corresponding to a standard 6-factor asset pricing model (Fama and French, 2018). In each panel, I show results for different mappings of flows into trading: the identity, following Lou (2012), following Edmans, Goldstein, and Jiang (2012), and following Coval and Stafford (2007). The samples are 1980 to 2017 in columns 1 and 2, and 2008 to 2017 in columns 3 and 4, because of the availability of daily Morningstar data. All regressions include time fixed effects. Standard errors are clustered by time. I report t-statistics in parentheses to allow the reader to quickly determine instrument strength and statistical significance. ***, **, and * denote significance at the 1%, 5% and 10% levels.

Table 2.5: Predicting (Surprise) Flow-Induced Trading

	(1) FIT(CRSP)	(2) FIT	(3) SFIT
L.ret	0.252***	0.319***	0.0978
	(0.0679)	(0.0651)	(0.0596)
L2.ret	0.139***	0.133*	-0.0586
	(0.0449)	(0.0677)	(0.0484)
L3.ret	0.102**	0.0835	-0.0269
	(0.0496)	(0.0709)	(0.0358)
L4.ret	0.124**	0.0920*	-0.0183
	(0.0506)	(0.0540)	(0.0334)
Market beta	-0.0798***	-0.134***	-0.0169
	(0.0198)	(0.0153)	(0.0124)
Log Market equity	0.0936***	0.0355	0.0227
	(0.0218)	(0.0212)	(0.0189)
Log Tobin's Q	-0.0741***	-0.0754***	-0.000531
	(0.0221)	(0.0178)	(0.0138)
Profitability	-0.0188	-0.0617*	0.000697
	(0.0393)	(0.0319)	(0.0190)
Investment	-0.302***	-0.484***	-0.0488
	(0.0524)	(0.0426)	(0.0410)
R-squared	0.195	0.128	0.0742
N	335551	100824	100824

This table shows regressions of different versions of flow-induced trading (FIT) on lagged stock returns and control variables. Column 1 uses the traditional instrument constructed using CRSP data, which is subject to the reverse causality critique. In column 2, I estimate the same regression, except that FIT is now constructed using Morningstar data. I show this intermediate step to clarify that the results are not driven by the change in data source. The number of observations decreases because there are only 10 years of Morningstar data in comparison to 37 years of CRSP data. In the last column, I estimate the regression using the new, optimal, corrected instrument. The control variables are the characteristics corresponding to the 6-factor asset pricing model (Fama and French, 2018), excluding momentum, because momentum is close to collinear with lagged returns. All regressions include time fixed effects. I report standard errors clustered by time in parentheses. ***, **, and * denote significance at the 1%, 5% and 10% levels.

Table 2.6: Predicting Returns with (Surprise) Flow-Induced Trading

	(1) ret	(2) ret	(3) ret
L.FIT(CRSP)	0.0115		
	(0.0513)		
L2.FIT(CRSP)	-0.0223		
	(0.0632)		
L3.FIT(CRSP)	-0.0856		
	(0.113)		
L4.FIT(CRSP)	-0.0619		
	(0.0612)		
L.FIT		0.0493	
		(0.0739)	
L2.FIT		0.0520	
		(0.120)	
L3.FIT		-0.108	
		(0.0893)	
L4.FIT		0.102	
		(0.0992)	
L.SFIT			-0.0751
			(0.101)
L2.SFIT			0.199*
			(0.103)
L3.SFIT			-0.0901
			(0.114)
L4.SFIT			0.159
			(0.118)
R-squared	0.206	0.156	0.156
N	332148	88380	88380

This table shows regressions of quarterly stock returns on lags of different versions of flow-induced trading (FIT) and control variables. Column 1 uses the traditional instrument constructed using CRSP data, which is subject to the reverse causality critique. Column 2 estimates the same regression, except that FIT is now constructed using Morningstar data. I show this intermediate step to clarify that the results are not driven by the change in data source. The number of observations decreases because there are only 10 years of Morningstar data in comparison to 37 years of CRSP data. In the last column, I estimate the regression using the new, optimal, corrected instrument. The control variables are the characteristics corresponding to the 6-factor asset pricing model (Fama and French, 2018). All regressions include time fixed effects. I report standard errors clustered by time in parentheses. ***, **, and * denote significance at the 1%, 5% and 10% levels.

Chapter 3

Payout-Induced Trading, Asset Demand Elasticities, and Market Feedback Effects

3.1 Introduction

Asset demand shocks that are plausibly exogenous to firm fundamentals are the basis for many important findings in empirical finance research. Asset demand shocks, interpreted as residual supply shocks, allow researchers to identify asset demand elasticities, shedding light on the extensive debate about whether stocks are perfect substitutes, a core implication of the efficient market hypothesis (Shleifer, 1986; Harris and Gurel, 1986). Asset demand elasticities are also necessary inputs for counterfactual experiments that investigate the role of institutions in financial markets (Koijen and Yogo, 2019). Moreover, by generating variation in asset prices that is unrelated to firm fundamentals, asset demand shocks create a laboratory to study market feedback effects, the causal effects of stock prices on the real economy (surveyed in Bond, Edmans, and Goldstein, 2012; Baker and Wurgler, 2013).

The literature uses two prominent asset demand shocks to identify asset demand elasticities and market feedback effects: index additions/deletions (e.g. Chang, Hong, and Liskovich, 2015) and mutual fund flow-induced fire sales (e.g. Edmans, Goldstein, and Jiang, 2012). Index additions/deletions cleanly identify the pricing effects of index membership, but in addition to triggering trades by indexers, index membership may also affect liquidity and corporate governance. Mutual fund flow-induced fire sales are intuitively appealing shocks, but their pricing effects could potentially be driven by reverse causality, as I examine in Schmickler (2020).

This paper uses an alternative asset demand shock that addresses these drawbacks, allowing me to shine new light on two important questions in empirical finance: what is the slope of the demand curve for stocks? And how does a nonfundamental increase in stock prices affect corporate financing and investment?

I investigate these questions using the mechanism described by Kvamvold and Lindset (2018) and Chen (2020) and based on financial institutions' reinvestment of firms' cash payouts. On average, each year, US public companies pay out almost 6% of total market equity in cash to shareholders via dividends, share repurchases, and M&A deals. Dollar total cash payouts peaked in 2018, at almost $2 trillion. How do shareholders redeploy these payouts? Households largely consume cash payouts (Baker, Nagel, and Wurgler, 2007), but equities are primarily held by financial institutions, which do not consume and do not immediately pass-through payouts to households. Instead, institutional investors reinvest payouts, because cash is unproductive. However, since cash payouts do not increase the supply of shares, shareholders cannot reinvest into the payout stock itself on average. Instead, to avoid falling behind their benchmarks, most institutional investors reinvest payouts into their existing portfolios. This *payout-induced trading* is a positive demand shock. Hence, when firms pay out cash, payout-induced trading drives up prices of other stocks held in the same portfolios of financial institutions.

The demand shock aggregates the hypothetical number of shares institutions buy in response to payout flows, if institutions reinvest in proportion to portfolio weights. This construction uses the idea from the mutual fund flow-induced fire sale literature that *hypothetical* trades in proportion to existing portfolio weights are likely exogenous to firm fundamentals, while actual trades may well

be driven by fundamentals (e.g. Edmans, Goldstein, and Jiang, 2012). I construct this shock using US equity portfolio holdings and payout data from 1980 to 2017.

What is the slope of the demand curve for stocks? I address this question by examining the price impact of payout-induced trading. The key challenge in identifying this price pressure effect is that payouts reveal fundamental information which also affects asset prices. Therefore, if similar firms are held in the same portfolios, abnormal returns of connected stocks may be the result of news, not price pressure. I solve this identification problem by exploiting that demand pressure effects differ from news effects in terms of timing; news changes asset prices around announcements, while demand pressure changes asset prices after the payment, when financial institutions receive cash and reinvest. Dividends in particular generate payment date price pressure but no announcement date news spillover effects, suggesting that dividend-induced trading is plausibly exogenous to fundamentals. Further, the absence of pre-payment date trends and the long lag between the announcement and the payment, e.g. an average of 45 days for dividends, ensure that payment date effects are driven by price pressure, not post-announcement drifts.

I test whether payouts generate news and/or price pressure effects in a systematic analysis of returns around payout events. For each type of payout, I show daily abnormal returns of the firm itself, 3-digit SIC code industry peers, and connected stocks around announcements and payment dates. I start with ordinary cash dividends, because they allow for the cleanest identification. Dividend announcements reveal fundamental information. But once announced, dividends are deterministic. Dividends are then paid with a typical delay of one to three months and the payment reveals no fundamental information. Hence, the gap between the announcement and payment dates separates news from demand pressure effects.

I find the payout firm itself experiences positive abnormal returns after both the announcement and payment dates, as is well-known (Ogden, 1994; Hartzmark and Solomon, 2013). Next, I test for spillovers on industry peers to detect whether dividends generate news spillover effects. I find no spillover effects on industry peers, neither at the announcement nor the payment date. This suggests that dividends do not reveal substantial fundamental information about other firms. Finally, connected firms do not experience abnormal returns around announcement dates, but they

do exhibit persistent, positive abnormal returns after the payment date, with no pre-payment date trend. Specifically, the estimated price impact of dividend-induced trading identifies a demand elasticity of 1.25, consistent with the evidence from Russell 2000 additions and deletions in Chang, Hong, and Liskovich (2015).

While dividends provide the cleanest setting to separate news from price pressure, I also examine other types of payouts. Like dividends, M&A payouts drive up connected stocks but not peer stocks after the payment date, consistent with demand pressure effects. Unlike dividends, M&A announcements are informative about other stocks' fundamentals; specifically, M&A announcements lift peer and connected stocks. Hence, M&A-induced trading is not a candidate instrument for returns. Next, for share repurchases, high-frequency payment data are unavailable. Hence, I can only examine announcement returns, not payment date returns. I do not find statistically significant repurchase announcement spillover effects on connected stocks, likely because firms do no begin to repurchase shares immediately after announcements. As I cannot separate news from price pressure effects of share repurchases, share repurchase-induced trading is not a candidate instrument for returns either.

I also investigate heterogeneity across financial institutions and find the results are strongest for mutual funds, followed by investment advisors (e.g. asset managers), consistent with the notion that institutions engage in payout-induced trading because they are benchmarked.

How does a nonfundamental increase in stock prices affect corporate financing and investment? I address this question using dividend-induced trading as an instrument for stock returns. Dividends generate variation in returns of connected stocks that is plausibly exogenous to fundamentals, as evidenced by the high-frequency stock price reactions to payments, but not announcements. Hence, dividend-induced trading is a valid instrument for returns. As the pricing effects are persistent, this allows me to estimate the causal effect of an increase in stock prices on corporate outcomes, i.e. market feedback effects (surveyed in Bond, Edmans, and Goldstein, 2012; Baker and Wurgler, 2013).

Payout-induced trading drives up stock prices. However, this effect occurs in the secondary market, without capital flowing to firms. So why would secondary market prices impact real out-

comes? While the literature documents several channels, I find a releveraging market feedback effect on investment. Using dividend-induced trading as an instrumental variable for stock returns, I find firms respond to an exogenous stock price increase by issuing debt, i.e. moving back towards their target leverage ratio, and using the funds to increase investment. I estimate that firms undo about a quarter of a stock price increase's impact on their debt to equity ratio by issuing debt over the following year. Further, I estimate that a nonfundamental 1% return increases investment by almost 1% relative to its median.

As an additional line of defense against news as a confounding channel, I repeat the exercise using only the expected component of dividends because, by definition, only surprise dividends, not expected dividends, convey news. Constructing expected dividends is simple; exploiting that managers smooth split-adjusted dividends per share achieves an R^2 of 93%. Instrumenting returns with *expected* dividend-induced trading gives the same results as before, further supporting the notion that the financing and investment responses constitute market feedback effects.

From a policy maker's perspective, this mechanism is important because it informs the recurring policy debate on whether to restrict corporate payouts (see e.g. Boissel and Matray, 2019). The typical reasoning is that firms should invest instead of returning capital to shareholders. While Boissel and Matray (2019) do indeed find that firms increase investment after an exogenous decrease in payouts, my paper finds a new spillover effect of payouts: capital investment occurs despite payouts – it just happens at other firms.

Related literature. This paper uses payout-induced trading as an alternative to two prominent asset demand shocks in the literature: index additions/deletions (e.g. Chang, Hong, and Liskovich, 2015) and mutual fund flow-induced fire sales (e.g. Edmans, Goldstein, and Jiang, 2012). Index additions/deletions cleanly identify the pricing effects of index membership, but in addition to triggering trades by indexers, index membership may also affect liquidity and corporate governance. Further, index additions/deletions cannot be used to study heterogeneity in demand elasticities across firm size, because treated stocks are of similar size, by virtue of being close to index inclusion thresholds. Finally, while index additions/deletions generate price effects that are statistically significant, they are weak instruments for returns (see table 4 of Chang, Hong, and Liskovich,

2015), making it difficult to leverage them to identify market feedback effects. Next, mutual fund flow-induced fire sales are intuitively appealing shocks, but their pricing effects could potentially be driven by reverse causality, as I suggest in Schmickler (2020). Recent work also casts doubt on the mechanical construction of the fund flow-based shock (Wardlaw, 2020). Overall, in contrast to these two asset demand shocks, payout-induced trading has several desirable features: payout-induced trading is determined before returns, does not change liquidity, is a strong instrument for returns, and can be used to investigate heterogeneous and/or high-frequency price pressure and market feedback effects.

This paper is closely related to Kvamvold and Lindset (2018) and Chen (2020) who also examine the payout-induced trading mechanism. The main difference is that my paper examines returns of connected stocks around announcements and payments to separate news from demand pressure effects. In particular, I find that dividends generate price pressure but no news spillover effects then justifying the use of dividend-induced trading as an instrument for returns allowing me to identify the slope of the demand curve for stocks as well as to identify market feedback effects.

Next, this paper contributes to the market feedback effect literature. Since Edmans, Goldstein, and Jiang (2012), the literature instruments returns with mutual fund outflow-induced fire sales and finds that undervaluation reduces investment (e.g. Edmans, Goldstein, and Jiang, 2012; Derrien, Kecskes, and Thesmar, 2013; Phillips and Zhdanov, 2013; Bonaime, Gulen, and Ion, 2018; Eckbo, Makaew, and Thorburn, 2018; Lou and Wang, 2018; Dessaint et al., 2018).[1] While this literature focuses on how asset price decreases affect investment, the effect need not be symmetric. In fact, Binsbergen and Opp (2019) argue that overpricing leads to larger real inefficiencies than underpricing because capital adjustment costs are asymmetric; divesting is costly and firms rarely do. Hence, it is important to understand how stock price *increases* affect investment. While mutual fund outflow-induced trading is an experiment for asset price decreases, payout-induced trading is an instrument for asset price increases, allowing me to study the real effects of asset price increases.

Further, this paper contributes to the large literature that examines pricing effects of payout

[1]Earlier studies that examine how stock prices affect investment but do not instrument returns include Blanchard, Rhee, and Summers (1993), Baker, Stein, and Wurgler (2003), Gilchrist, Himmelberg, and Huberman (2005), Chen, Goldstein, and Jiang (2007), Polk and Sapienza (2009), and Bakke and Whited (2010).

events. The bulk of this literature focuses on how payout events affect the firm itself. For dividends, firms experience positive abnormal returns after announcements, before ex-dates, and after payments (e.g. Ogden, 1994; Hartzmark and Solomon, 2013). For share repurchases, firms experience positive abnormal returns after share repurchase program announcements (e.g. Vermaelen, 1981; Grullon and Michaely, 2004; Bargeron, Kulchania, and Thomas, 2011). For M&A, targets experience large, positive returns after announcements; acquirors experience muted returns (e.g. Asquith, Bruner, and Mullins, 1983; Jensen and Ruback, 1983; Mitchell, Pulvino, and Stafford, 2004). This paper contributes to this literature by systematically examining the returns of peer and connected stocks around announcement and payment dates of all types of payouts.

3.2 Data

Constructing payout-induced trading, which is the central object of this paper, requires two types of data: portfolio holdings of financial institutions and firm payout data. The empirical tests also require additional institution and firm characteristics.

3.2.1 Institutional Stock Holdings

Portfolio holdings come from two sources. First, stock holdings for all institutions except mutual funds are from the Thomson Reuters (TR) Institutional Holdings Database. This includes banks, insurance companies, investment advisors (e.g. asset managers), pension funds, and other investors (e.g. endowments). I apply the institution type correction from Koijen and Yogo (2019). Thomson Reuters' sources are SEC 13F filings. All financial institutions managing above $100 million must report long positions. 13F holdings are at the institution level. Second, more micro-level data are available for mutual funds, as the Thomson Reuters Mutual Fund Holdings database provides fund-level portfolios. For example, instead of Vanguard's holdings, the database provides the holdings of the Vanguard Dividend Growth Fund. The sources for this database are SEC-mandated disclosures in Forms N-30D, N-Q, and N-CSR as well as voluntary disclosures. Holdings data dictate my sample, which spans 1980 to 2017.

In addition, for mutual funds, detailed institution-level data are available. I take mutual fund AuM, returns, and distributions from the CRSP Mutual Fund Database and households' reinvestments of mutual fund distributions and portfolio equity shares from Morningstar.[2] For other institution types, I infer AuM and returns from portfolio holdings. I measure AuM as the total market value of all observed portfolio positions, and portfolio returns as the portfolio weight-weighted mean of stock returns, assuming that institutions only trade at the end of each quarter. I use AuM and portfolio returns to compute investment flows as a control variable. I follow the standard definition of investment flows. In time period t, Institution i receives investment flows of $flow_{i,t} = (A_{i,t} - A_{i,t-1}(1 + r_{i,t}))/A_{i,t-1}$, where $A_{i,t}$ are AuM and $r_{i,t}$ are institution returns (Coval and Stafford, 2007).

3.2.2 Firm Payouts

Ordinary cash dividends (distribution code *"distcd"* 1000-1399) and stock dividends (*distcd* 5530-5539) are from CRSP. The typical source for M&A data, SDC Platinum, does not provide payment dates; hence, I take M&A payment dates, and also announcement dates for consistency, from CRSP (*distcd* 3000-3399 for cash and *distcd* 3700-3799 for stock deals). CRSP payment dates assume M&A payment after delisting. Share repurchase program announcement dates are from SDC Platinum; payment date data are unavailable. Quarterly share repurchases are from Compustat North America Fundamentals Quarterly. Firms only have to report the actual number of shares repurchased since 2004. Accordingly, researchers must infer share repurchases for the pre-2004 period. I infer share repurchases following Banyi, Dyl, and Kahle (2008) who show that measures based on the Compustat item *purchases of common stock* provide the most accurate estimate of actual shares repurchased. Together, holdings and payout data allow me to construct dollar payout flows to each institution

[2] I preferentially take mutual fund data from the CRSP Mutual Fund Database, unless the data quality of Morningstar data is higher, as is the case for portfolio equity shares.

$$PayoutFlow_{i,t} = \sum_{n=1}^{N} Payout_t(n) Shares_{i,t-1}(n). \tag{3.1}$$

Institution i holds $Shares_{i,t-1}(n)$ shares of stock n, which pays out $Payout_t(n)$ per share, and dollar payout flows are the sum of the payouts from all N stocks. I construct payout flows separately for each type of payout, resulting in cash dividend, stock dividend, cash M&A, stock M&A, and share repurchase flows. For payouts in stock, I use the market value of the securities paid. I use stock payouts in placebo tests. For share repurchase flows, I do not observe which institutions sell to repurchase programs. Hence, I assume all institutions sell a fraction equal to the fraction of shares repurchased by the firm. Finally, I construct industry payout as the market capitalization-weighted mean payout to price ratio of firms with the same 3-digit SIC code.[3]

3.2.3 Firm Characteristics

Stock data are from CRSP, accounting data are from Compustat North America Fundamentals Annual and Quarterly. The main dependent variable is log returns. The main control variables are the characteristics corresponding to a standard six-factor asset pricing model (Fama and French, 2018), i.e. beta, log market equity, log Tobin's Q, profitability, investment, and momentum, because they are the most prominent drivers of expected returns. I also control for dividend to book equity as in Koijen and Yogo (2019), because of the outsize importance of dividends in this paper, though controlling for payouts is often redundant, because I exclude payout stocks in the main specifications to isolate spillover effects. The firm characteristics are constructed following Koijen and Yogo (2019). Accounting data are released with a delay, so I lag accounting data by 6 months. To construct market beta, I take the 1-month T-bill rate and the market return from Kenneth French's website. I calculate market betas using 60-month rolling window regressions.

When testing for market feedback effects on investment, I follow Dessaint et al. (2018). This

[3] I choose SIC codes over NAICS codes, which CRSP assigns starting in 2004, and over Hoberg and Phillips (2016) industry classifications, which start in 1996, because they cover the full 1980-2017 sample. I choose 3-digit SIC codes because this corresponds to the industry level. It is also the level of granularity targeted by Hoberg and Phillips (2016) industry classifications.

means I measure the investment rate as the ratio of capital expenditures and property plant and equipment and exclude financial firms (SIC codes 6000-6999) and utilities (SIC codes 4000-4999). I also test whether the investment response is financed by debt or equity. I measure debt as total liabilities and equity as common equity. In all exercises, I winsorize characteristics cross-sectionally at the 1 and 99% level, as for example in Green, Hand, and Zhang (2017) and Dessaint et al. (2018). I restrict the sample to US ordinary common stocks that trade on the NYSE, AMEX, and Nasdaq; have non-missing market values and returns; and for which the holdings data cover at least 1% of shares outstanding.

3.2.4 Summary Statistics

Firm Payouts

Figure 3.3 shows summary statistics. The first plot compares total cash payouts to two prominent, alternative sources of asset demand shocks: total absolute extreme mutual fund outflows and the total market value of firms that are added/deleted from the Russell 1000/2000 index. All three variables are scaled by total market equity. Index additions/deletions are from Chang, Hong, and Liskovich (2015). Total cash payouts, the sum of cash dividends, share repurchases, and cash M&A payouts, fluctuate around almost 6% of total market equity per year. This is the key number describing the magnitude of the shock. Firms pay out 6% of market equity in cash, so the average investor receives a 6% annual payout flow, and that translates into a 6% demand shock if all investors reinvest. At the end of the sample, total market equity is about $30 trillion, and total cash payouts are almost $2 trillion.

In comparison, total index/additions deletions are about 2%, and lower since Russell Inc. smoothed index transitions in 2007. Total extreme mutual fund outflows are around 1%, and are lower at the beginning of the sample, before the rise of mutual funds. Payouts are significantly larger than fund flows and index additions/deletions. That said, these shocks do not translate into demand shocks in the same way. How payouts translate into demand shocks depends on the fraction of investors that reinvests; for index additions/deletions it depends on the fraction of investors

tracking the Russell 1000/2000; and for fund flows it depends on how mutual funds accommodate flows. The second plot breaks cash payouts into their three components. From 1980 to 2017, dividends dropped from 4% to 2% of total market equity. Companies substituted dividends for share repurchases, which increased from 0% to 2%. Lastly, cash M&A transactions fluctuate between 0 and 3%. Overall, cash payouts are economically large, suggesting they have the potential to create large demand shocks and consequently large price impacts and real effects.

Financial Institutions

Table 3.1 summarizes financial institution information by institution type and decade. For mutual funds and investment advisors (e.g. asset managers), the number of institutions and market share increased steadily over the sample but remained largely stable for banks, pension funds, insurance companies, and unclassified institutions (e.g. endowments). Mutual funds and investment advisors are the most important institution types in terms of equity market share. In the most recent decade, mutual funds, investments advisors, banks, pension funds, insurance companies, and unclassified institutions hold 25, 21, 12, 3, 2, and 2% of all equities, respectively. The remaining share, about one third, is held by households and foreign investors.

Mutual funds, investment advisors, and unclassified institutions tend to be small in terms of AuM; banks, pension funds, and insurance companies tend to be larger. However, mutual funds and investment advisors are by far the most frequent institution type. For mutual funds, this is due to the availability of fund-level instead of institution-level data.

Most institutions hold concentrated portfolios. Mutual funds, investment advisors, and unclassified institutions are the least diversified, with a median of 59, 53, and 27 stocks held during the last decade, respectively. Even at the 90th percentile, they only hold 368, 270, and 441 out of approximately 4000 stocks. The median insurance company, bank, and pension fund are not diversified either, with 185, 187, and 512 stocks. However, they are diversified at the 90th percentile, with 2068, 1232, and 1581 stocks. The fact that most institutions hold concentrated portfolios is important for my analysis. If all investors held the market portfolio, every investor would reinvest payouts in proportion to market weights and there would be no cross-sectional variation in the

payout-induced trading demand shock. Finally, many institutions receive large payout flows. In the current decade, the median institution receives payout flows between 5 and 6%, with pension funds showing the largest tilt towards payout firms, likely because of their tax-exempt status. At the 90th percentile, payout flows range from 6.8% for banks to 9.2% for unclassified institutions.

3.3 Price Pressure Effects

This section estimates the price impact of payout-induced trading. The analysis is at the stock level, allowing me to investigate the effects at the daily frequency.

3.3.1 The Demand Shock

Institution i holds $Shares_{i,t}(n)$ split-adjusted shares of stock n at time t. I denote relative payout flows as $payoutflow_{i,t} = PayoutFlow_{i,t}/A_{i,t-1}$. Chen (2020) shows that institutions reinvest firm payouts into their existing portfolios. In addition, following the reasoning from the literature that constructs asset demand shocks from mutual fund flows (e.g. Edmans, Goldstein, and Jiang, 2012), hypothetical reinvestment in proportion to existing portfolio weights is likely exogenous to firm fundamentals, while actual trades may well be driven by firm fundamentals. If institutions reinvest in proportion to portfolio weights, they buy $payoutflow_{i,t}Shares_{i,t-1}(n)$ shares. This behavior can be used to construct an asset demand shock by aggregating the hypothetical number of shares institutions buy in response to payout flows. Scaling by total shares held by all institutions gives the relative demand shock

$$PIT_t(n) = \frac{\sum_{i=1}^{I} payoutflow_{i,t}Shares_{i,t-1}(n)}{\sum_{i=1}^{I} Shares_{i,t-1}(n)}. \tag{3.2}$$

I construct PIT separately for different payout types. This gives cash dividend-induced trading, stock dividend-induced trading, cash M&A-induced trading, stock M&A-induced trading, and share repurchase-induced trading. The definition of payout-induced trading is closely related to mutual

fund flow-induced trading (e.g. Lou, 2012; Edmans, Goldstein, and Jiang, 2012). Replacing payout flows with mutual fund flows yields the standard mutual fund flow-induced trading instrument as in Edmans, Goldstein, and Jiang (2012), though the exact construction here also takes the Wardlaw (2020) critique into account.

Two sources of variation drive the payout-induced trading shock: first, naturally, payouts; second, investor heterogeneity. If all investors held the market portfolio, every investor would reinvest payouts in proportion to market weights, so dollar payout-induced trading would be exactly proportional to market weights in each cross-section. In this world, every stock would get the same cross-sectional payout-induced trading shock, making it impossible to identify its price impact in the cross-section.

Figure 3.4 illustrates the deviation from this world using histograms of the ratio between a stocks' share of dollar payout-induced trading on a given day and its market weight. This construction eliminates time series variation. With homogeneous investors, all values would be equal to one. I winsorize the ratio at 5. Otherwise, the graph is dominated by whitespace. The left histogram illustrates the cash dividend shock. The distribution is approximately exponential, with most values falling between 0 and 2 and a small mass at 5. This means that on one hand, many stocks are close to untreated, while on the other hand, some stocks receive multiple times the shock they would receive in a homogeneous investor world. The right histogram corresponds to the cash M&A instead of the dividend shock. The distribution is similar, but the spread is even wider. More stocks are close to untreated, and more stocks receive more than five times the shock they would receive in the homogeneous investor world. This is because on an average day, 30 firms pay a dividend, but zero or one firm is acquired. Overall, Figure 3.4 shows that investor heterogeneity generates large cross-sectional dispersion in payout-induced trading.

3.3.2 Identification Strategy

I want to identify the price impact of payout-induced trading. The key challenge is that payouts reveal fundamental information which also changes asset prices. Therefore, if similar firms are held in the same portfolios, abnormal returns of connected stocks may be the result of news, not

price pressure. The identification strategy thus needs to separate price pressure from news effects. Of course, news and price pressure are not mutually exclusive and can both affect asset prices. However, demand pressure effects differ from news effects in terms of timing; news changes asset prices at announcement, while demand pressure changes asset prices at payment, when financial institutions receive cash and reinvest. Ordinary cash dividends in particular allow for the cleanest identification. Dividend announcements reveal fundamental information but once announced, dividends are deterministic. Dividends are then paid after a delay of one to three months. The payment reveals no fundamental information. Hence, the gap between the announcement and payment date separates the news from the demand pressure effect.

Figure 3.5 illustrates the setting I exploit. The left plot shows the distribution of the number of days between the dividend payment and announcement date. This gap separates the news from the demand pressure effect. Dividends are typically paid about one to three months after the announcement, with a mean delay of 45 days. The 1st percentile is 17 days, likely sufficient time for the market to incorporate any announcement date news into prices.

The right plot shows total cash dividends for each day in 2017. The plot highlights that on almost every business day of the year, at least one company pays a dividend. Each stock is held by many institutions and each institution holds many stocks. Consequently, institutions disperse payouts over many stocks. Hence, payout-induced trading affects almost every stock on almost every day, but with varying intensity. On most days, companies pay out about one billion dollars. On a few dozen days, companies pay out about five billion dollars. The biggest daily shocks are about ten billion dollars.

As almost every firm is treated on almost every day and with varying intensity, I investigate the price impact of payout-induced trading using a distributed lag model. In particular, I estimate the daily frequency, cross-sectional regression

$$r_t(n) = \alpha_t + \sum_{l=\underline{L}}^{\bar{L}} \gamma_l Z_{t-l}(n) + \beta' X_t(n) + \epsilon_t(n), \tag{3.3}$$

of stock returns on a vector of shocks Z, which contains the payout to price ratio, the industry payout ratio, and payout-induced trading. I use the payout to price ratio to measure the payout's impact on the firm itself, the industry payout ratio to measure spillovers on peers, and payout-induced trading to measure spillover effects on connected stocks. I include $\bar{L}$ lags and $\underline{L}$ leads. The coefficients on the lags measure the price impact of the shock; $\gamma_l > 0$ means stocks experience positive abnormal returns l days after the shock. I include the leads to test for pre-event trends. This empirical strategy is the analogue to an event study for continuous treatment. For the case when Z is one-dimensional and binary, the results are numerically identical to event study estimates (Schmidheiny and Siegloch, 2020). Lastly, X is a vector of control variables as described in the data section. I cluster standard errors by time because returns are highly correlated in the cross-section.

PIT depends on the payouts of the firm itself. This is consistent with the theoretical reasoning behind the payout-induced trading demand shock. When a firm pays out cash, investors want to reinvest some of this cash into that firm. However, it is known that the firm experiences positive abnormal returns after payouts (Ogden, 1994). The innovation of this paper is to show payouts' spillover effects on other stocks. While I control for firm payouts, this may not fully eliminate the concern. Further, firms held in the same portfolios are similar, so their payout schedules may be correlated and it is known that firms experience abnormal returns around dividend announcement and ex-dividend dates (Hartzmark and Solomon, 2013). I address both concerns by estimating spillover effects on the subsample that excludes payout firms, i.e. firms that pay out during the year centered around the date of the observation. Hence, the estimates are free of the known self-effects and isolate spillover effects.

3.3.3 The Return Pattern

I test whether payouts generate news and/or price pressure effects in a systematic analysis of returns around payout events. For each type of payout, I report daily abnormal returns of the firm itself, peer stocks, and connected stocks around announcement and around payment dates. I provide a compact summary of the findings in Figure 3.2 and go into full detail here.

Ordinary Dividends

I start with ordinary cash dividends, because they allow for the cleanest identification. Figure 3.6 investigates return patterns around cash dividend events at the daily frequency. On the left, I show returns around the announcement date, which capture the effects of news. On the right, I show returns around the payment date, which capture the effects of price pressure. The first row reports returns of the payout firm itself, the second row presents spillover effects on industry peers, and the last row displays spillover effects on connected stocks. The plots report estimates of equation 3.3. I plot cumulative coefficients and 95% confidence intervals from a Wald test as described in Schmidheiny and Siegloch (2020). Note that the coefficient estimates are comparable across columns, but not across rows, because the different rows report coefficients on different variables.

Plots (a) and (b) find the payout firm experiences abnormal returns after both the announcement and the payment. The cumulative coefficients reach 0.2 and 0.07, meaning a 1% dividend triggers a 20 basis point return after the announcement and a 7 basis point return after the payment. Both graphs exhibit modest, positive, pre-event trends, indicating that investors anticipate dividend events. These self-effects in panels (a) and (b) are well known (Ogden, 1994; Hartzmark and Solomon, 2013).

Plots (c) and (d) find no evidence for spillovers on industry peers, though there may be modest, positive, temporary spillover effects after the announcement date. If dividends conveyed news about other firms' fundamentals, we would expect to see spillover effects on industry peers. Therefore, the absence of spillover effects on industry peers, even after the announcement, suggests that dividends do not reveal substantial information about fundamentals of other firms. This is unsurprising, because in the vast majority of cases, firms simply keep dividends constant on a split-adjusted dividend per share basis. In fact, I exploit this behavior in section 3.4.3 where I construct expected dividends and find that simply predicting that past dividend behavior continues yields an R^2 of 93%.

Lastly, panels (e) and (f) examine spillover effects on other stocks held in the same portfolios of financial institutions. To test for spillover effects on connected firms after the announcement,

I construct hypothetical dividend-induced trading if dividends were paid on the announcement date. I find no evidence for spillover effects after the dividend announcement. While there is a modest increase from day three to day ten, the effect is almost precisely 0 over the three days following the shock, as well as over the full $\pm$ 10 day window. This indicates that dividends do not reveal fundamental information about connected stocks. This is expected, as dividends do not even generate spillover effects on industry peers.

Yet, panel (f) documents that connected stocks rise after the payment. This return response is the price impact of payout-induced trading. Connected stocks experience positive, abnormal returns the day of and for two days after the payment. The coefficients on the following lags, as well as the estimated pre-event trends, are close to zero and statistically insignificant. The reaction is strongest one day after the payment. I estimate that a 1% demand shock translates into a 0.8% return. This implies an asset demand elasticity of 1.25, consistent with the evidence from Russell 2000 additions and deletions in Chang, Hong, and Liskovich (2015). Overall, I find dividends generate no news spillover effects but do produce large price pressure spillover effects on connected stocks.

I find no reversal after the demand shock. Chen (2020) provides related results at low-frequency, documenting that the price impact of dividend- and share repurchase-induced trading is persistent. However, he does find reversal after M&A-induced trading price pressure effects. In the literature, whether price pressure effects are followed by reversal depends on the setting. Ogden (1994) finds no reversal of the dividend payment date effect. Neither do newer studies exploiting index/additions deletions, including Kaul, Mehrotra, and Morck (2000) and Chang, Hong, and Liskovich (2015). However, early studies of S&P 500 index additions do find price reversal (Shleifer, 1986; Harris and Gurel, 1986). Similarly, Greenwood (2005) finds partial reversal. Studies of mutual fund flow-induced fire sales also document reversal (e.g. Edmans, Goldstein, and Jiang, 2012). However, Wardlaw (2020) finds that factor loadings explain the reversal following flow-induced fire sales.

Cash M&A Payouts

So far, I have estimated the price impact of PIT using ordinary cash dividends. While dividends provide a clean setting to separate news from price pressure, there are other payout types. However,

there are no high-frequency payment data for share repurchases and for M&A payouts, announcement and payment are less cleanly separated than for dividend payouts, because in rare cases, M&A deals fail even after a merger agreement is signed. While this setting is not ideal for identification, it does provide additional evidence.

Figure 3.7 repeats the analysis in Figure 3.6 for cash M&A instead of cash dividend payouts. Note that the top right panel, "Payment, payout stock", is missing because the M&A target ceases to exist. All other panels are analogous. Panel (a) shows the well-known empirical fact that M&A target stocks experience positive abnormal returns after deal announcements, with a statistically significant pre-event trend, suggestive of information leakage (e.g. Jensen and Ruback, 1983). The second row looks at spillover effects on peer firms. In contrast to dividend payouts, M&A payouts reveal fundamental information about other stocks. Peer stocks rise after announcements but are flat around payments. The latter suggests that failures of closed M&A deals are so rare that payments do not reveal fundamental information about other firms. Accordingly, any payment date spillover effects on connected stocks can be interpreted as price pressure effects. In the last row, I show spillover effects on connected stocks, which exhibit gains after M&A announcements and payments, consistent with price pressure effects, particularly as there are no payment date spillover effects on peer stocks.

M&A cash payments generate qualitatively the same payment date pattern as cash dividend payments. Connected stocks experience positive abnormal returns the day of and the two days after payments. Again, the cumulative coefficient flattens after that. In addition, while largely statistically insignificant, there is a positive pre-event trend. This is expected, because many investors sell their positions in M&A targets and reinvest early. The coefficients sum to 0.25. This is smaller than in the dividend-induced trading exercise, likely because M&A deals trigger trading frenzies. Consequently, the quarterly holdings snapshots mismeasure payout recipients, biasing down the estimates. The M&A results are less reliable quantitatively but still confirm my finding qualitatively. Overall, M&A deals generate news and price pressure spillover effects. Therefore, while they provide additional evidence for payout-induced trading price pressure effects, they are not ideal to construct instrumental variables for stock returns.

Share Repurchases

For completeness, Figure 3.8 repeats the analysis in Figure 3.6 for share repurchases. Note that the right column, "Payment, ...", is missing because there are no high-frequency payment data. Nevertheless, examining share repurchase announcement spillover effects allows me to investigate the speed of spillover effects. Fast spillover effects are likely driven by news; slow spillover effects are likely the result of price pressure. This is because SEC rule 10b-18 requires the repurchasing company to delay actual repurchases to give the market sufficient time to absorb the new information. The rule also imposes volume limits, meaning most repurchase programs are not completed immediately. Hence, examining the speed of spillover effects helps gauge share repurchase-induced trading's suitability as an instrument for stock returns.

Panel (a) shows the well-known empirical fact that stocks experience positive abnormal returns after share repurchase program announcements (e.g. Vermaelen, 1981). Panel (b) demonstrates that peer firms largely do not experience abnormal returns, suggesting that repurchase announcements reveal predominantly idiosyncratic information. Lastly, panel (c) does not find statistically significant spillover effects on connected stocks, though there is a small positive response with a two day delay. It is not clear whether this response is driven by news or price pressure. Either markets incorporate the announcement information about connected stocks with a delay of two days, or institutions start reinvesting share repurchase proceeds two days after repurchase announcements. While both explanations are certainly possible, connected stocks' gains appear to coincide with the beginning of the actual repurchases, not the announcements. The cumulative coefficient reaches 0.1, smaller than after dividend or M&A events. This is unsurprising, because most share repurchase programs are likely not completed 10 business days after the announcement.

3.3.4 Institutional Heterogeneity

One reason institutions engage in payout-induced trading is likely that many are benchmarked. This implies institutional heterogeneity: the more an institution is benchmarked, the more likely it is to reinvest into its existing portfolio. To test this, I construct payout-induced trading using

only one institution type at a time. I then estimate equation 3.3 for each institution type. This ranks institution types by price impact, and benchmarked institution types should lead the list. In particular, many mutual funds are closet indexers (Cremers and Petajisto, 2009) or index funds, so mutual funds are likely the most benchmarked institution type.

Figure 3.9 reports the same analysis as panel (f) of Figure 3.6 but by institution type. Broadly, I find price pressure effects for mutual funds and investment advisors (e.g. hedge funds and asset managers) but not for banks, pension funds, insurance companies, or unclassified institutions (e.g. endowments). This is likely because mutual funds and investment advisors are large and benchmarked, while other institution types are small and/or not benchmarked. Panel (a) shows that dividend-induced trading by mutual funds generates largely the same price response pattern as aggregate payout-induced trading. Connected stocks experience positive abnormal returns after the payment. The cumulative coefficient is even larger, 1 instead of 0.7, consistent with mutual funds being the most benchmarked institution type. Panel (b) displays the results for investment advisors. Again, the price response is positive, concave and statistically significant. However, the cumulative coefficient is now smaller, 0.3 instead of 0.7. There are three potential explanations for this decrease. First, investment advisors are less benchmarked than mutual funds. Second, investment advisors are more price elastic than mutual funds; they are less likely to continue buying as prices are rising. Finally, investment advisors, hedge fund in particular, may trade more frequently, meaning their quarterly portfolio snapshots mismeasure payout recipients more.

Panels (c) and (d) present the analogous analysis for cash M&A payouts. The overall pattern is the same. The price impact of M&A-induced trading is positive and statistically significant for mutual funds and investment advisors. Again, the price impact is larger for mutual funds than for investment advisors. Further, section 3.3.3 argues that the estimated price impact of M&A-induced trading is likely lower than that of dividend-induced trading because many institutions sell their shares of M&A targets to merger arbitrageurs, implying quarterly holdings mismeasure who receives M&A payouts. This suggests further institutional heterogeneity: there should be no price impact for institution types that sell to merger arbitrageurs but there should be a price response for institution types that *are* merger arbitrageurs. As mutual funds and hedge funds (included in

investment advisors) are the main institution types that run merger arbitrage strategies, the results are consistent with this reasoning.

3.4 Market Feedback Effects

This section examines how payout-induced trading, by driving up stock prices, affects corporate financing and investment. I investigate these effects at the annual frequency to accommodate the seasonality of investment as well as that managers may respond slowly because of financial frictions (e.g. Fazzari et al., 1988; Kaplan and Zingales, 1997).

An extensive literature asks whether stock prices have causal effects on real corporate outcomes, predominantly investment (e.g. Edmans, Goldstein, and Jiang, 2012; Derrien, Kecskes, and Thesmar, 2013; Phillips and Zhdanov, 2013; Bonaime, Gulen, and Ion, 2018; Eckbo, Makaew, and Thorburn, 2018; Lou and Wang, 2018; Dessaint et al., 2018). To answer this question, researchers need a natural experiment that generates variation in stock returns that is independent of fundamentals. The state-of-the-art instrument for returns is mutual fund outflow-induced trading. The idea is that mutual fund redemptions are as good as random but force mutual funds to liquidate assets, driving down prices (Edmans, Goldstein, and Jiang, 2012). Hence, outflow-induced trading is an experiment for decreases in asset prices. All of the market feedback effect papers cited above use this instrument.[4] The only market feedback effect paper investigating the impact of increases in asset prices that I am aware of is Khan, Kogan, and Serafeim (2012). They instrument returns with mutual fund inflow-induced trading. However, they investigate effects on SEOs, not investment.

Even though the literature focuses on how asset prices decreases affect investment, the effect need not be symmetric. In fact, Binsbergen and Opp (2019) argue that overpricing leads to larger real inefficiencies than underpricing because capital adjustment costs are asymmetric; firms rarely divest because it is costly. Hence, it is important to understand how asset price *increases* impact investment. As an alternative to outflow-induced trading, payout-induced trading provides an

[4]Earlier studies that examine how stock prices affect investment but do not instrument returns include Blanchard, Rhee, and Summers (1993), Baker, Stein, and Wurgler (2003), Gilchrist, Himmelberg, and Huberman (2005), Chen, Goldstein, and Jiang (2007), Polk and Sapienza (2009), and Bakke and Whited (2010).

experiment for asset price increases and allows me to estimate the other side of the market feedback effect. In addition, payout-induced trading is a new experiment for changes in asset prices. This is valuable, because recent work casts doubt on the mechanical construction of the mutual fund flow-induced instrument (Wardlaw, 2020) as well as on the direction of causality (Schmickler, 2020).

3.4.1 Identification Strategy

I start with a standard investment-Q regression, closely following the empirical setup in Dessaint et al. (2018)

$$\frac{I_t(n)}{K_{t-1}(n)} = \alpha_t + \alpha(n) + \beta Q_{t-1}(n) + \gamma' X_{t-1}(n) + \xi_t(n), \tag{3.4}$$

a regression of the investment rate on Tobin's Q, control variables, time, and firm fixed effects. The regression is motivated by a Q-theory of investment model which relates investment to marginal Q (e.g. Almeida, Campello, and Galvao, 2010). Marginal Q is unobservable but equals average Q under constant returns to scale and perfect competition (Hayashi, 1982).

Payout-induced trading is an instrument for returns, approximately first differences of log Tobin's Q. Hence, I estimate the equation in first differences

$$\Delta \frac{I_t(n)}{K_{t-1}(n)} = \eta_t + \beta r_{t-1}(n) + \tilde{\gamma}' \tilde{X}_{t-1}(n) + \epsilon_t(n) \tag{3.5}$$

and instrument returns with payout-induced trading. The control variables are the risk factor characteristics as described in the data section, as well as the payout ratios for all payout types, i.e. cash dividends, stock dividends, cash M&A, stock M&A, and share repurchases. I also estimate alternative versions of this equation. I show estimates excluding control variables as a simple baseline and estimates replacing time fixed effects with time x industry fixed effects which mitigates the potential concern that payout-induced trading captures news instead of price pressure spillover

effects.

I illustrate this identification strategy in Figure 3.1. Pfizer pays a dividend; Pfizer is held in the same portfolio as Lyft, but not as Uber. Hence, Lyft experiences payout-induced trading, driving up its stock price relative to Uber. This section tests whether Lyft increases its investment rate over the following year relative to Uber and attributes this investment response to the price increase. The threat to identification is that since they are held in the same portfolio, Lyft may be more similar to Pfizer than Uber. Accordingly, Pfizer and Lyft may be hit by common shocks that do not hit Uber. Further, market feedback effect regressions require a low-frequency instrument because managers make corporate decisions at low frequency. Therefore, I cannot use the same, daily frequency empirical strategy that identified the payout-induced trading price pressure effect. However, section 3.3 shows that dividends only generate spillover effects on asset prices of connected stocks after the payment date, not after the announcement date. This suggests that payout-induced trading only affects asset prices of connected stocks via a demand pressure channel and not via a news channel. Therefore, it is a plausibly exogenous shock to asset prices of connected stocks.

As discussed in section 3.3, PIT depends on the payouts of the firm itself. At high frequency, where data are abundant, a simple way to separate the spillover effect from the self-effect was to exclude payout firms. At low frequency, however, data are not abundant. Therefore, while I examine the subsample that excludes payout firms in a robustness check, I begin with an alternative approach to address this concern. I define $PayoutFlowEx_{i,t}(n) = \sum_{m \neq n} Payout_t(m) Shares_{i,t-1}(m)$ and substitute *PayoutFlow* with *PayoutFlowEx* in the definition of PIT. Now PIT is trading induced by other firms' payouts and is thus free of the known self-effect. In addition, I control for the payout ratios corresponding to all payout types.

3.4.2 The Investment Response

Table 3.2 reports estimates of equation 3.5. I show the first and second stage of 2SLS estimates of equation 3.5, as well as the OLS estimate as a baseline. In all specifications, I cluster standard errors by time because returns are highly correlated in the cross-section. Appendix Table 3.7 demonstrates that clustering by time is more conservative than clustering by firm and as conservative as double

clustering by time and firm. The first column reports the results of the OLS regression. There is a strong positive relationship between returns and investment over the next year. The coefficient is 0.08, meaning a 1% return is associated with a 8 basis point increase in investment. This corresponds to about 0.5% of the median annual investment rate of 17%. However, this coefficient estimate is not the causal effect of stock prices on investment. One prominent reason for this is that measurement error creates attenuation bias (e.g. Erickson and Whited, 2000; Almeida, Campello, and Galvao, 2010; Bakke and Whited, 2010). Specifically, Bakke and Whited (2010) argue returns that are ignored by managers can be treated econometrically as measurement error and show that applying errors-in-variables techniques significantly increases the estimated investment-to-Q sensitivity. Another attenuation bias comes from the investment factor, i.e. the empirical fact that high-investment firms have lower expected returns, likely because firms with a lower cost of capital invest more (Titman, Wei, and Xie, 2004; Hou, Xue, and Zhang, 2015).

The next columns report results from instrumental variable regressions. Columns 2 and 3 present the first and second stage of a baseline 2SLS estimate that does not include control variables. The first stage regression shows that dividend-induced trading is a strong instrument, with a Kleibergen and Paap (2006) F-statistic of 50. In the second stage, I estimate a highly statistically significant coefficient of 0.1. This estimate is higher than the OLS coefficient, consistent with the common finding that instrumenting returns addresses the mismeasurement problem and mitigates the attenuation bias (e.g. Erickson and Whited, 2000; Almeida, Campello, and Galvao, 2010; Bakke and Whited, 2010).

In columns 4 and 5, I add control variables. Again, dividend-induced trading is a strong instrument with an F-statistic of over 50. In the second stage, the coefficient remains statistically significant and increases from 0.1 to 0.17. In the last two columns, I include time x industry fixed effects. News spillovers should be strongest within an industry. Hence, if the results were driven by news spillovers, the coefficients in the first and second stage should decrease significantly with the inclusion of time x industry fixed effects. This is not the case. The instrument remains strong with an F-statistic of 30, and the coefficient in the second stage decreases slightly to 0.15. The investment-return elasticity estimates range from 0.1 to 0.17, but their difference is not statistically

significant. Overall, a 1% return translates into a 0.1 to 0.17% increase in investment, up to 1% of the median annual investment rate. Returns translate into investment almost one-for-one.

3.4.3 Expected payout-induced trading

As an additional line of defense against news as a confounding channel, I repeat the exercise using only the expected component of dividends because, by definition, only surprise dividends, not expected dividends, convey news. At the end of each year, I predict split-adjusted dividends per share for each quarter of the next year. I take this approach because the market feedback effect regressions are at the annual frequency, but the shock is constructed using quarterly portfolio holdings data. Constructing expected dividends is simple. For dividends announced in the previous year, actual and expected dividends are the same. If no dividend is announced yet, I exploit that managers smooth split-adjusted dividends per share. I predict that the last dividend behavior continues in the next year, for each dividend frequency. This simple dividend prediction achieves an R^2 of 93%. I illustrate this in Appendix Figure 3.10 using a histogram of the relative prediction error. For the vast majority of dividends, the simple prediction is exactly accurate, as evidenced by the large mass at 0. However, the prediction misses dividend initiations and omissions, which correspond to the small bars at -1 and 1. In addition, it misses dividend increases, which correspond to the small bars between -1 and 0. Finally, dividend decreases are too infrequent to be visible.

Table 3.3 reports estimates of equation 3.5 using expected dividend-induced trading as the instrument for returns. The results are close to identical to the estimates in table 3.2. The instrument is slightly weaker. Looking at the strictest specification in column 5, the F-statistic is 26 instead of 32 and the estimated market feedback effects are slightly larger, but statistically indistinguishable. Again, considering the strictest specification in column 6, the coefficient on returns is 0.17 instead of 0.15. This shows that the results are not driven by the news component of dividend-induced trading and therefore suggests that the increase in investment is driven by price pressure.

3.4.4 Capital Structure Rebalancing Effect

How do firms finance the investment increases? Taking a step back, how do stock prices affect corporate financing? Baker and Wurgler (2002) find firms issue equity to take advantage of higher stock prices. In contrast, Leary and Roberts (2005), Flannery and Rangan (2006), and Kayhan and Titman (2007) show firms slowly counteract stock price changes to move towards their target debt ratio, consistent with survey evidence in Graham and Harvey (2001) that most CFOs have a target capital structure. However, these studies examine how firms' financing decisions respond to changes in stock prices, whereas I ask how firms react to *exogenous* changes in stock prices. The difference may be large, because firms experience great returns when their investment opportunities improve and vice versa. Accordingly, firms may issue equity to raise capital for new investment opportunities or repurchase stock due to a lack of investment opportunities. In contrast, I explore how firms react to returns that are unrelated to fundamentals.

I test this by estimating equation 3.5 but substituting the investment rate for the debt or equity issuance rate. I show the results in Table 3.4 which reports results corresponding to columns 3 and 4 in Table 3.2, i.e. the specification including controls and time fixed effects. I report results for alternative specifications in the Appendix. I find firms increase debt, not equity. For debt, the 2nd stage coefficient is 0.28 and statistically significant at the 5% level. In response to a 1% return, firms increase debt by 0.28%. This means a firm undoes about one quarter of the stock return's impact on its debt to equity ratio. For equity, I find a positive but statistically insignificant coefficient of 0.06. Together, these findings mean firms partially undo *exogenous* stock returns' effects on leverage. They issue debt to rebalance their capital structure and use the funds for real investment.

3.4.5 Robustness

The Appendix contains further robustness results. In Table 3.6, I find the same results qualitatively when excluding payout firms. In fact, the coefficients in the first and second stage increase. I also show that the results are robust to omitting within-industry flows. I implement this by constructing payout flows assuming that investors reinvest into their existing portfolio, but not if a stock is in

the same industry. While this is a bad description of investor behavior, it allows me to test whether within-industry flows drive the results, which would be inconsistent with the price pressure channel. I find that this is not the case. The estimated coefficients are virtually unchanged. Next, Appendix Table 3.7 shows that clustering by time is substantially more conservative than clustering by firm and as conservative as double clustering by time and firm. This indicates that error terms are correlated in the cross-section, but less in the time series.

Finally, Appendix Tables 3.8 and 3.9 report results for alternative specifications of the regressions that test for market feedback effects on corporate financing. Appendix Table 3.8 reports results for the stricter specification that includes time x industry fixed effects. Appendix Table 3.9 also includes time x industry fixed effects, but in addition, it uses *expected* dividend-induced trading as the instrument for returns. The results are robust to either specification.

3.5 Conclusion

Cash payouts by US public companies are economically large: the average annual cash payout is almost 6% of market equity and dollar total cash payouts peaked in 2018, at almost $2 trillion. This paper uses the reinvestment of cash payouts by financial institutions as a nonfundamental shock to asset prices to estimate the slope of the demand curve for stocks as well as the real effects of stock price increases on corporate financing and investment.

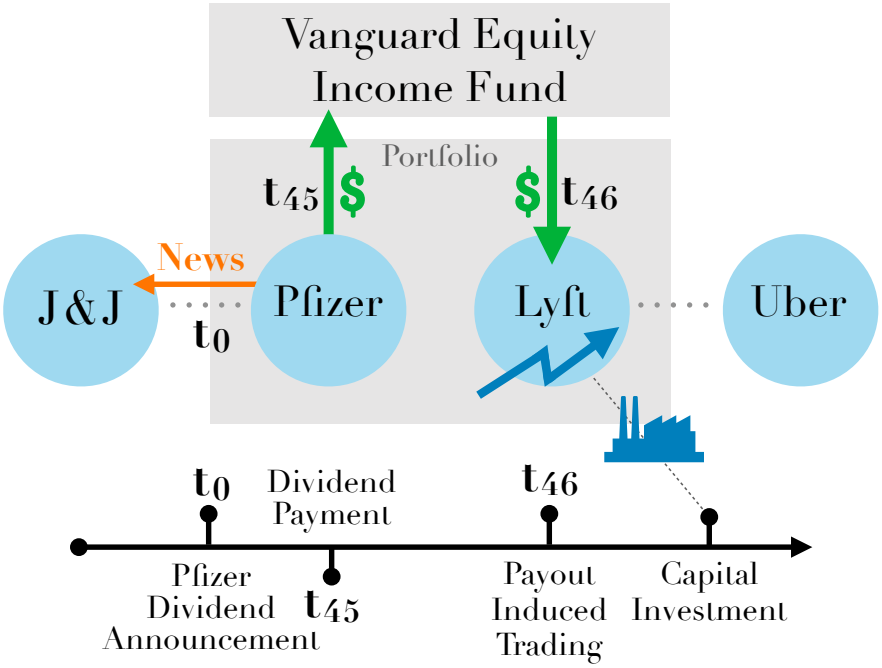
Exploiting the separation of announcement and payment at the daily frequency, I find price pressure spillover effects of firm payouts on other stocks held in the same portfolios of financial institutions which identify an asset demand elasticity of 1.25. Dividends in particular generate payment date price pressure but no announcement date news spillover effects, suggesting that dividend-induced trading is plausibly exogenous to fundamentals. Therefore, I use dividend-induced trading as an instrument for stock returns and document a releveraging market feedback effect on investment where firms respond to an exogenous stock price increase by issuing debt and use the funds to invest. I estimate that firms undo about a quarter of a nonfundamental stock price increase's impact on their debt to equity ratio by issuing debt over the following year. Further, I

estimate that a nonfundamental 1% return increases investment by almost 1% relative to its median. This informs the recurring policy debate on whether to restrict payouts so firms invest instead. My paper finds a new channel which implies that capital investment occurs despite payouts – it just happens at other firms.

In future work, payout-induced trading could be used to shed light on new feedback effects of financial markets. In particular, payout-induced trading is a shock that hits all stocks almost every day. Therefore, unlike existing natural experiments for stock returns, it could be used to investigate heterogeneity in price pressure and market feedback effects, as well as high-frequency market feedback effects.

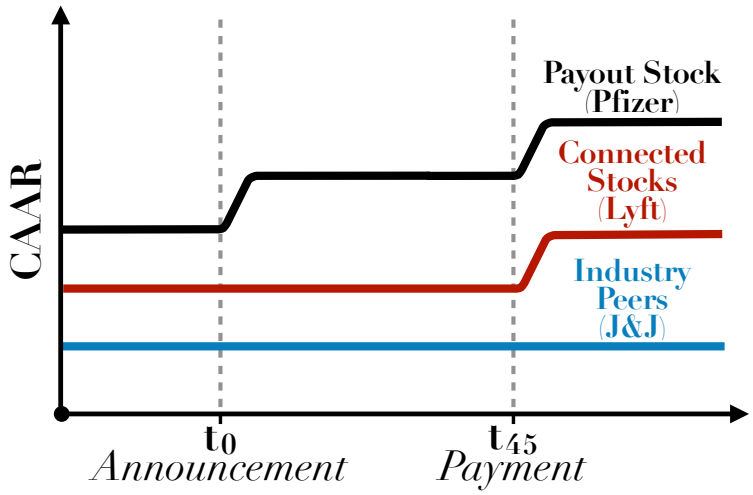
Figures

Figure 3.1: Illustration of Payout-Induced Trading Mechanism



This figure illustrates the payout-induced trading effect. At t_0, Pfizer announces a cash dividend. This reveals fundamental information about Pfizer and potentially also about peer firms such as Johnson & Johnson. The dividend is paid with a 45-day lag at t_{45}. Vanguard Equity Income Fund holds Pfizer and hence receives a dividend check. As the fund is benchmarked and/or a closet-indexer, it reinvests the cash into its portfolio. I call these purchases *payout-induced trading*. Payout-induced trading may occur immediately on t_{45} and/or over the following days. Here, Lyft is held in the same portfolio as Pfizer and is thus subject to payout-induced trading. This demand shock pushes up Lyft's price relative to industry peers. Then, in response to the lower cost of capital, Lyft increases investment.

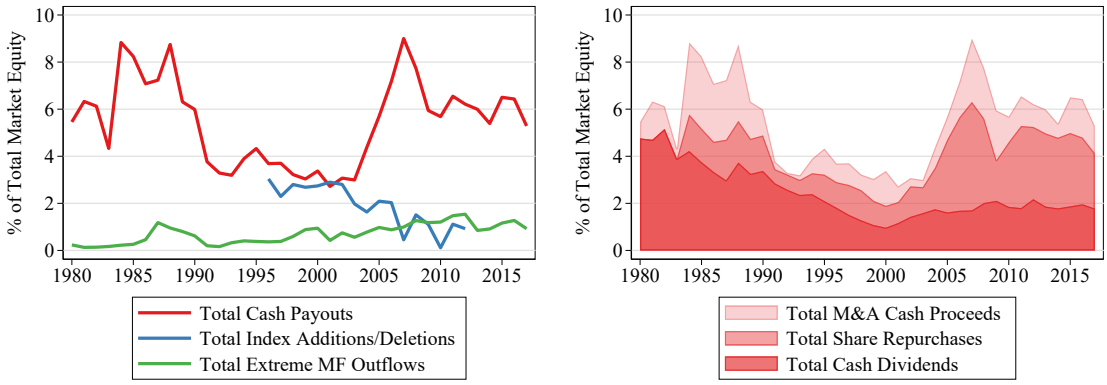
Figure 3.2: Illustration of Identification Strategy



(a) Returns around Dividend Events

This figure illustrates the main identification strategy. Demand pressure effects differ from news effects in terms of timing; news change asset prices at announcement, while demand pressure changes asset prices at payment, when financial institutions receive cash and reinvest. I examine returns of the firm itself, peer stocks, and connected stocks around announcement and payment dates. As is widely known, the payout firm itself always experiences positive abnormal returns after dividend announcements and payments (Ogden, 1994; Hartzmark and Solomon, 2013). Next, I find no spillover effects on industry peers, neither at the announcement nor the payment date. This suggests that dividends do not reveal substantial fundamental information about other firms. Finally, connected firms do not experience abnormal returns around announcement dates, but they do exhibit persistent, positive abnormal returns after the payment date, suggesting that dividend-induced trading is plausibly exogenous to fundamentals and motivating its use as an instrument for stock returns.

Figure 3.3: Total Cash Payouts vs. Alternative Sources of Asset Demand Shocks

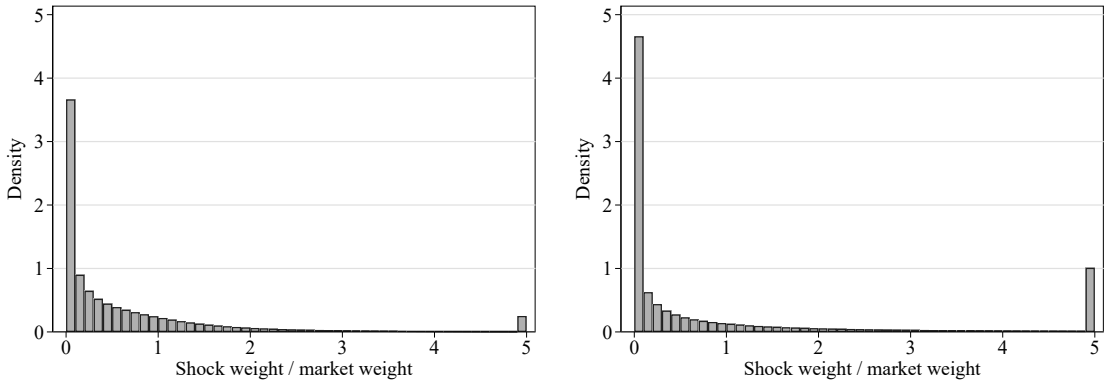


(a) Sources of Asset Demand Shocks

(b) Payout Type Breakup

This figure shows summary statistics. The first plot compares total cash payouts to two prominent, alternative sources for asset demand shocks: total absolute extreme mutual fund outflows and total market value of firms that are added/deleted from the Russell 1000/2000 index. All three are scaled by total market equity. Total cash payouts are the sum of cash dividends, share repurchases, and cash M&A payouts. Total extreme mutual fund outflows are the absolute value of the sum of quarterly mutual fund level flows less than -5%. These are the source of the current state-of-the-art shock for stock returns (Wardlaw, 2018). Total index additions/deletions are the sum of the market values of all firms that switch from the Russell 2000 to the Russell 1000 or vice versa. This shock is as in Chang, Hong, and Liskovich (2015). The second plot breaks payout into its three components.

Figure 3.4: The Payout-Induced Trading Shock

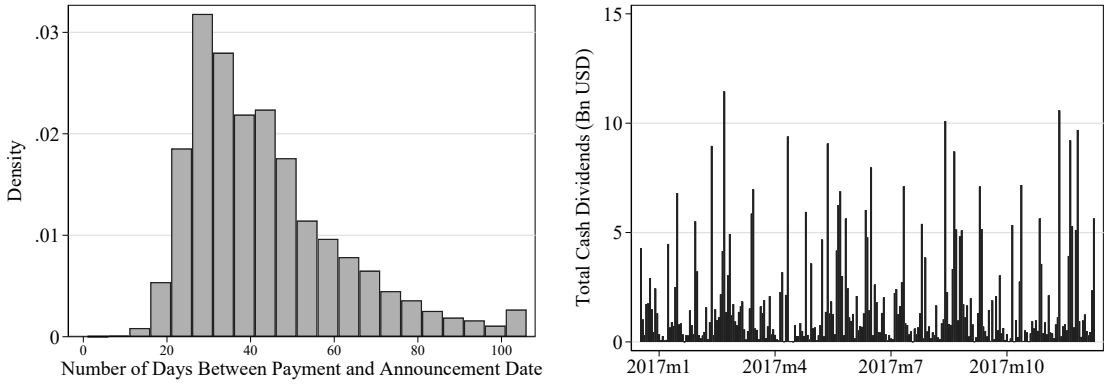


(a) Dividend-Induced Trading

(b) M&A-Induced Trading

The two plots show histograms for the ratio between a stocks' share of dollar payout-induced trading on a given day and its market weight. This construction eliminates time series variation. If all investors held the market portfolio, the ratio would be exactly 1 for all observations. The ratio is not bounded above, so I winsorize it at 5 for this figure. (a) shows the histogram for the cash dividend shock; (b) shows the histogram for the cash M&A shock. The sample is the daily stock level panel from 1980 to 2017.

Figure 3.5: Ordinary Dividends

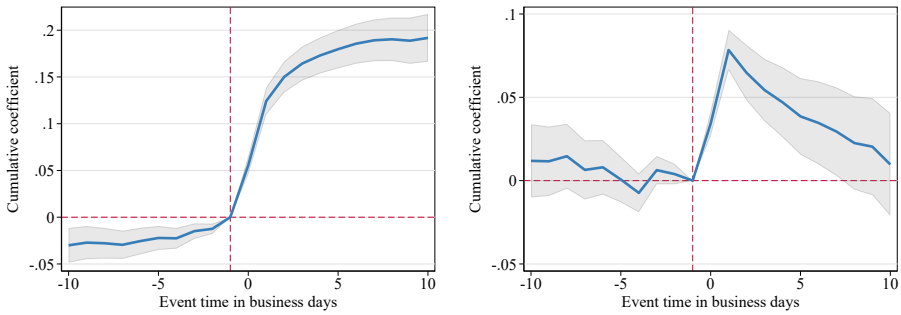


(a) Payment - Announcement Date Lag

(b) Daily Cash Dividends in 2017

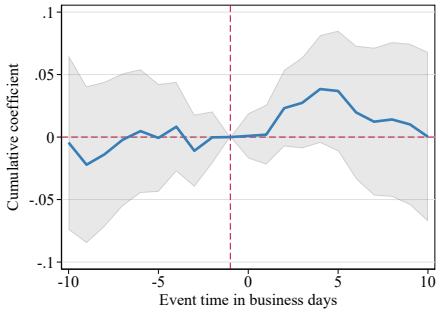
This figure illustrates cash dividend payouts. The first plot shows the distribution of the number of days between the dividend payment and announcement date. The number of days can be very large, so I winsorize it at the 99th percentile for this figure. I single out dividends because they allow for the cleanest identification of the payout-induced trading effect because of the gap between the announcement and payment date. Plot (b) shows total cash dividends for each day in 2017.

Figure 3.6: Abnormal Returns Around Dividend Announcement and Payment Dates

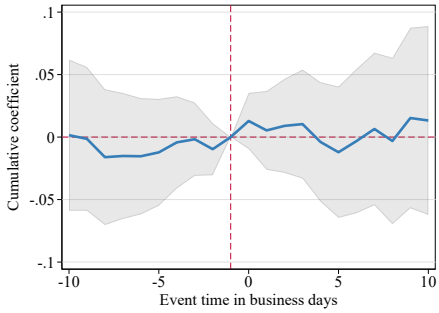


(a) Announcement, payout stock

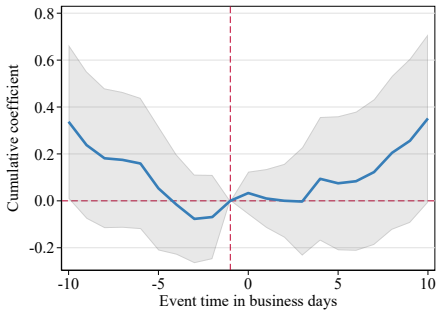
(b) Payment, payout stock



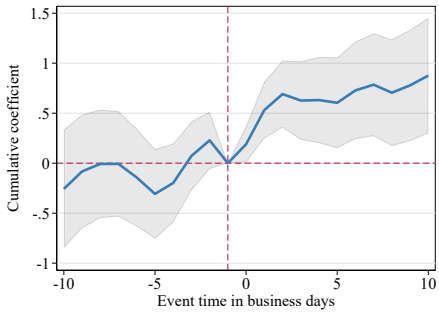
(c) Announcement, industry peers



(d) Payment, industry peers



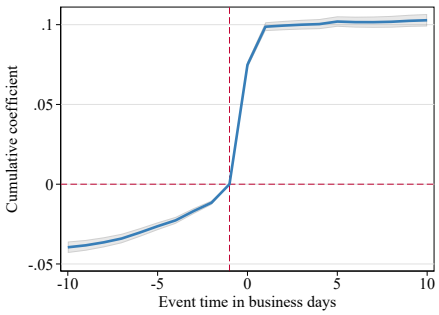
(e) Announcement, connected stocks



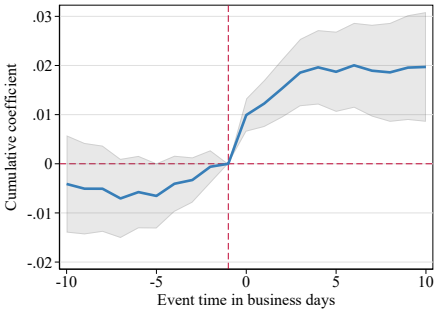
(f) Payment, connected stocks

The plots show percent cumulative abnormal returns in response to a 1% shock around dividend announcement and payment dates of the firm itself (a and b), 3-digit SIC code industry peers (c and d), and other stocks held in the same portfolios of financial institutions (e and f). Coefficients are comparable across columns but not across rows, because different rows report coefficients on different variables. The plots are based on estimates of equation 3.3, with standard errors clustered by time. I show cumulative coefficients and 95% confidence intervals from a Wald test. Following Schmidheiny and Siegloch (2020), I sum the lead and lag coefficients separately. The regressions in graph (a) and (b) use the full sample. Graphs (c) to (f) use the sample that excludes payout firms to isolate spillover effects. The controls are described in the data section. The sample is daily from 1980 to 2017.

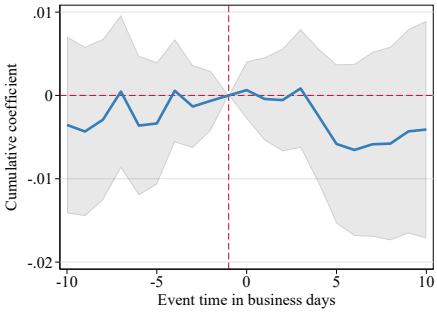
Figure 3.7: Abnormal Returns Around M&A Announcement and Payment Dates



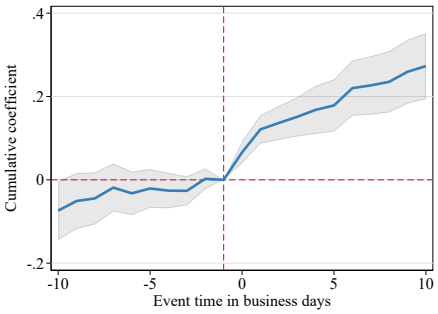
(a) Announcement, payout stock



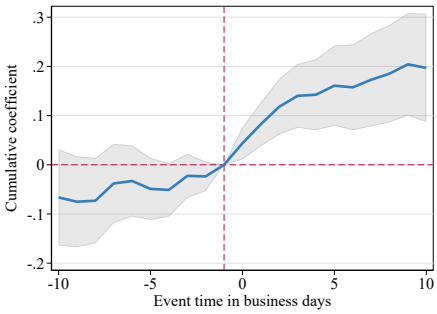
(b) Announcement, industry peers



(c) Payment, industry peers



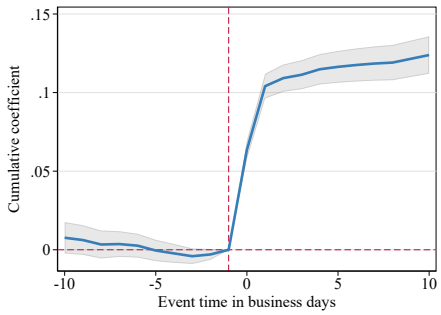
(d) Announcement, connected stocks



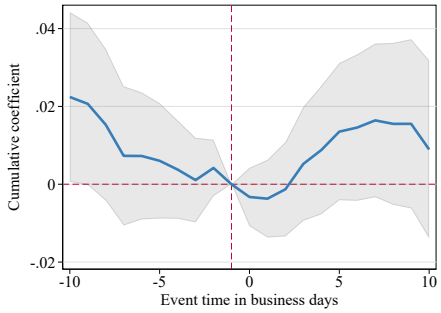
(e) Payment, connected stocks

The plots show percent cumulative abnormal returns in response to a 1% shock around cash M&A announcement and payment dates of the firm itself (a), 3-digit SIC code industry peers (b and c), and other stocks held in the same portfolios of financial institutions (d and e). Coefficients are comparable across columns but not across rows, because different rows report coefficients on different variables. The plots are based on estimates of equation 3.3, with standard errors clustered by time. I show cumulative coefficients and 95% confidence intervals from a Wald test. Following Schmidheiny and Siegloch (2020), I sum the lead and lag coefficients separately. The regression in graph (a) uses the full sample. Graphs (b) to (e) use the sample that excludes payout firms to isolate spillover effects. The controls are described in the data section. The sample is daily from 1980 to 2017.

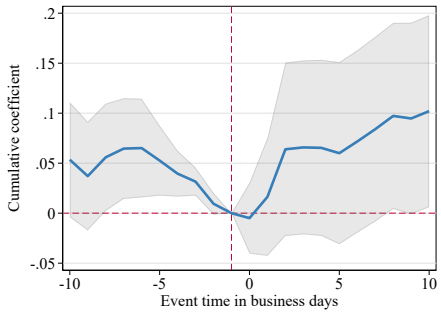
Figure 3.8: Abnormal Returns Around Share Repurchase Announcement Dates



(a) Announcement, payout stock



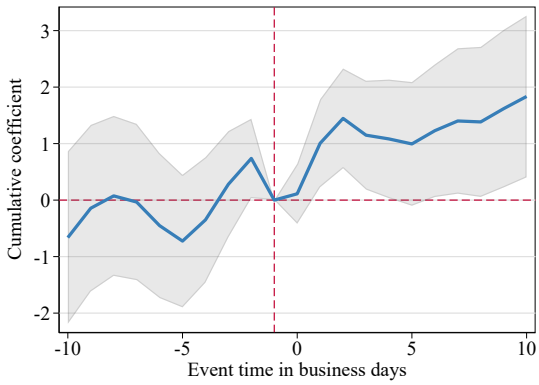
(b) Announcement, industry peers



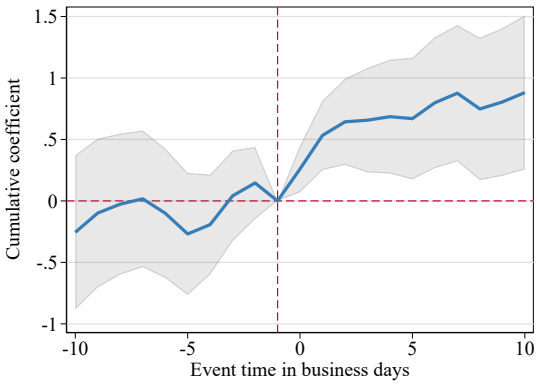
(c) Announcement, connected stocks

The plots show percent cumulative abnormal returns in response to a 1% shock around share repurchase announcement dates of the firm itself (a), 3-digit SIC code industry peers (b), and other stocks held in the same portfolios of financial institutions (c). Coefficients are not comparable across plots, because different plots report coefficients on different variables. The plots are based on estimates of equation 3.3, with standard errors clustered by time. I show cumulative coefficients and 95% confidence intervals from a Wald test. Following Schmidheiny and Siegloch (2020), I sum the lead and lag coefficients separately. The regression in graph (a) uses the full sample. Graphs (b) and (c) use the sample that excludes payout firms to isolate spillover effects. The controls are described in the data section. The sample is daily from 1980 to 2017.

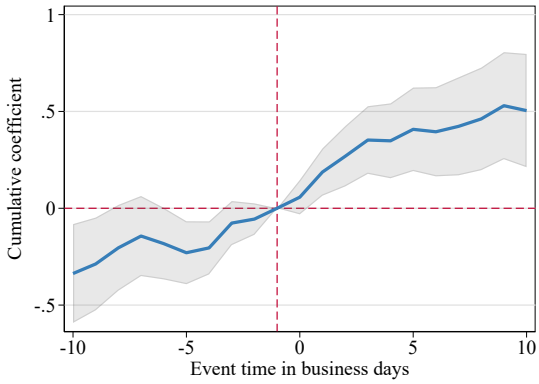
Figure 3.9: Price Impact of Payout-Induced Trading – Institutional Heterogeneity



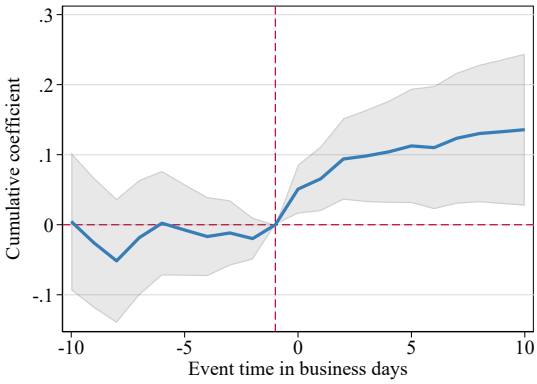
(a) Dividends, Mutual Funds



(b) Dividends, Investment Advisors



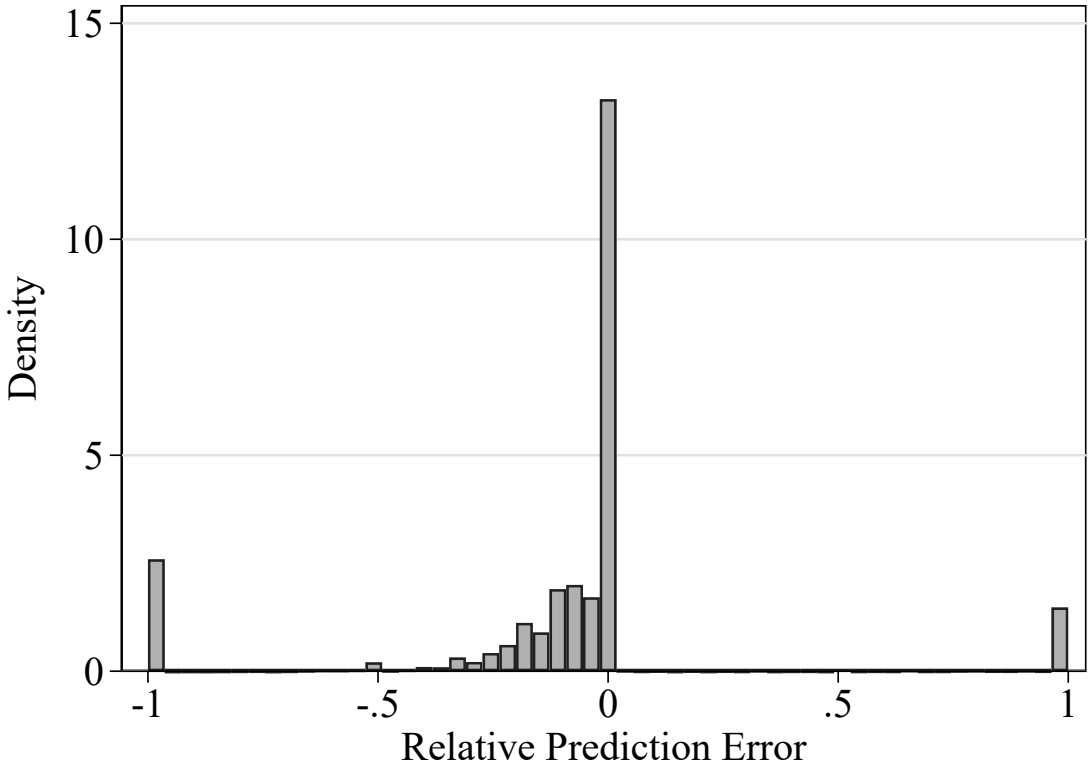
(c) M&A, Mutual Funds



(d) M&A, Investment Advisors

The plots show percent cumulative abnormal returns in response to a 1% payout-induced trading shock, for different institutions and payout types. The top row reports results for cash dividend-induced trading, the bottom row for cash M&A-induced trading. The left column reports results for mutual funds, the right column for investment advisors (e.g. hedge funds). The plots are based on estimates of equation 3.3, with standard errors clustered by time. I show cumulative coefficients and 95% confidence intervals from a Wald test. Following Schmidheiny and Siegloch (2020), I sum the lead and lag coefficients separately. The underlying regressions use the sample that excludes payout firms to isolate spillover effects. The controls are described in the data section. The sample is daily from 1980 to 2017.

Figure 3.10: Relative Dividend Prediction Error



This histogram shows the distribution of the relative dividend prediction error, i.e. the prediction error scaled by the prediction target. -1 and 1 are dividend initiations and dividend discontinuations, respectively. At the end of each year, I predict split-adjusted dividends per share for each quarter of the next year. This is because the market feedback effect regressions are at the annual frequency, but the shock is constructed using quarterly frequency portfolio holdings data. For dividends announced in the previous year, actual and expected dividends are the same. If no dividend is announced yet, I exploit that managers smooth split-adjusted dividends per share. I predict that the last dividend behavior continues in the next year, for each dividend frequency. The sample is the daily stock level panel from 1980 to 2017.

Tables

Table 3.1: Summary Statistics for Financial Institutions by Type

			AuM ($million)		# stocks held		Payout flows	
Period	# Inst.	Mkt share	50th	90th	50th	90th	50th	90th
Panel A: Mutual Funds								
1980-1989	401	4	63	486	46	103	4.5	9.6
1990-1999	1366	10	85	1008	56	170	3.0	6.0
2000-2009	3167	20	114	1525	64	312	4.0	7.8
2010-2017	3426	25	168	2353	59	368	5.1	8.7
Panel B: Investment Advisors								
1980-1989	224	7	244	1154	81	240	4.8	11.3
1990-1999	588	9	217	1255	73	221	3.2	6.8
2000-2009	1621	13	227	1762	66	247	3.9	9.1
2010-2017	2857	21	204	2295	53	270	5.1	9.1
Panel C: Banks								
1980-1989	220	15	386	3393	185	580	5.8	7.9
1990-1999	204	13	493	10250	213	960	3.7	4.6
2000-2009	167	12	396	15308	205	1322	4.5	5.9
2010-2017	160	12	404	18676	187	1232	5.7	6.8
Panel D: Pension Funds								
1980-1989	32	3	871	5613	154	549	5.8	8.3
1990-1999	36	4	1376	23765	367	1291	3.7	4.9
2000-2009	39	3	3536	37969	633	2229	4.9	5.7
2010-2017	53	3	4032	25819	512	1581	5.9	7.6
Panel E: Insurance companies								
1980-1989	67	3	389	2293	101	412	5.3	9.6
1990-1999	75	4	827	5307	136	800	3.4	4.8
2000-2009	58	4	1126	15097	203	1780	4.5	6.3
2010-2017	49	2	903	35447	185	2068	5.7	8.6
Panel F: Other financial institutions								
1980-1989	37	1	213	1116	55	192	4.9	10.0
1990-1999	32	1	251	2379	61	144	3.5	5.0
2000-2009	144	1	137	1401	37	236	3.7	11.7
2010-2017	179	2	168	3225	27	441	4.8	9.2

This table summarizes financial institution information by institution type and decade, reporting time-series means by institution type and within the given period. I report the number of institutions, the market share of the institution type in %, the median and 90th percentile of assets under management in $ million, the number of stocks held in a portfolio, and payout flows in %. The investment advisor type includes e.g. hedge funds. The "other" type includes e.g. endowments. The sample is 1980 to 2017.

Table 3.2: The Market Feedback Effect on Investment

	(1) OLS	(2) FS	(3) 2SLS	(4) FS	(5) 2SLS	(6) FS	(7) 2SLS
return	0.0826***		0.100***		0.171***		0.151***
	(0.00605)		(0.0257)		(0.0409)		(0.0407)
DIT		0.0210***		0.0140***		0.0118***	
		(0.00288)		(0.00187)		(0.00209)	
Controls				✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓		
Time x ind. FE						✓	✓
F-statistic		53.07		56.06		31.75	
N	112444	112444	112444	112076	112076	109725	109725

This table shows estimates of equation 3.5, the firm-level regression of the change in the investment rate on returns and controls and fixed effects depending on the specification. The control variables are stock characteristics as described in the data section. The first column shows the results of an OLS regression. The next columns show first and second stages of different versions of 2SLS regressions. The instrument is normalized cash dividend-induced trading (DIT). The frequency is annual. I report standard errors clustered by time in parentheses and Kleibergen and Paap (2006) F-statistics in the table footer. ***, **, and * denote significance at the 10%, 5%, and 1% levels.

Table 3.3: The Market Feedback Effect on Investment – Instrument with *Expected* DIT

	(1) FS	(2) 2SLS	(3) FS	(4) 2SLS	(5) FS	(6) 2SLS
return		0.107***		0.192***		0.172***
		(0.0269)		(0.0457)		(0.0452)
DIT	0.0186***		0.0119***		0.00994***	
	(0.00271)		(0.00175)		(0.00193)	
Controls			✓	✓	✓	✓
Time FE	✓	✓	✓	✓		
Time x ind. FE					✓	✓
F-statistic	46.73		46.21		26.53	
N	112444	112444	112076	112076	109725	109725

This table is analogous to table 3.2, except that the instrumental variable is *expected* dividend-induced trading. It shows estimates of equation 3.5, the firm-level regression of the change in the investment rate on returns and controls and fixed effects depending on the specification. The control variables are stock characteristics as described in the data section. The columns show first and second stages of different versions of 2SLS regressions. The instrument is normalized for ease of comparison. The frequency is annual. I report standard errors clustered by time in parentheses and Kleibergen and Paap (2006) F-statistics in the table footer. ***, **, and * denote significance at the 10%, 5%, and 1% levels.

Table 3.4: Capital Structure Rebalancing Effect

	Debt			Equity		
	(1) OLS	(2) FS	(3) 2SLS	(4) OLS	(5) FS	(6) 2SLS
return	0.0487***		0.282**	-0.0401**		0.0625
	(0.0123)		(0.107)	(0.0170)		(0.0831)
DIT		0.0140***			0.0154***	
		(0.00181)			(0.00219)	
F-statistic		59.78			49.11	
N	100845	100845	100845	94264	94264	94264

This table is analogous to table 3.2, though with different dependent variables. I replace investment (CAPX, the change in property plant and equipment) with the change in total liabilities in columns 1 to 3 and the change in common equity in columns 4 to 6. The table shows estimates of equation 3.5, the firm-level regression of the respective dependent variable on returns, controls, and time fixed effects. The control variables are stock characteristics as described in the data section, except that I drop Tobin's Q from the equity issuance regression because it contains book equity which is part of the dependent variable. The columns show first and second stages of different versions of 2SLS regressions. The instrument is normalized cash dividend-induced trading (DIT). The frequency is annual. I report standard errors clustered by time in parentheses and Kleibergen and Paap (2006) F-statistics in the table footer. ***, **, and * denote significance at the 10%, 5%, and 1% levels.

Table 3.5: Firms Subject to Dividend-Induced Trading are Similar to Dividend-Paying Firms

	DIT	Div/P
Market beta	-0.104**	-0.290***
	(0.0444)	(0.0335)
Log market equity	0.990***	0.242***
	(0.0768)	(0.0230)
Log Tobin's Q	-0.649***	-0.259***
	(0.0525)	(0.0335)
Profitability	-0.673***	0.0457
	(0.240)	(0.120)
Investment	-1.570***	-0.855***
	(0.122)	(0.0763)
R-squared	0.467	0.233
N	149313	149313

This table shows estimates of regressions of dividend-induced trading and of the dividend yield on firm characteristics and time fixed effects. The frequency is annual. The sample is 1980 to 2017. I report standard errors clustered by time in parentheses. ***, **, and * denote significance at the 10%, 5%, and 1% levels.

Table 3.6: The Market Feedback Effect on Investment – Robustness Tests

	Exclude payout firms		No within-industry flows	
	(1) FS	(2) 2SLS	(3) FS	(4) 2SLS
DIT	0.0245***			
	(0.00382)			
return		0.199***		0.187***
		(0.0415)		(0.0425)
DIT (ex ind.)			0.0128***	
			(0.00171)	
F-statistic	41.18		55.94	
N	72637	72637	112076	112076

This table is analogous to table 3.2, except for an alternative sample or instrument. Columns 1 and 2 exclude dividend-paying firms. Columns 3 and 4 use a version of the dividend-induced trading instrument that assumes investors do not reinvest into stocks with the same 3-digit SIC code. The table shows estimates of equation 3.5, the firm-level regression of the change in the investment rate on returns and controls and time fixed effects. The control variables are stock characteristics as described in the data section. The columns show first and second stages of different versions of 2SLS regressions. The instruments are normalized for ease of comparison. The frequency is annual. I report standard errors clustered by time in parentheses and Kleibergen and Paap (2006) F-statistics in the table footer. ***, **, and * denote significance at the 10%, 5%, and 1% levels.

Table 3.7: The Market Feedback Effect on Investment – Alternative Clustering

	Time		Firm		Firm & time	
	(1) FS	(2) 2SLS	(3) FS	(4) 2SLS	(5) FS	(6) 2SLS
DIT	0.0140***		0.0140***		0.0140***	
	(0.00187)		(0.000643)		(0.00189)	
return		0.171***		0.171***		0.171***
		(0.0409)		(0.0188)		(0.0369)
F-statistic	56.06		474.9		54.96	
N	112076	112076	112076	112076	112076	112076

This table is analogous to table 3.2, except for alternative assumptions about the distribution of the residual. In columns 1 and 2, I cluster standard errors by time (repeated for comparison). In columns 3 and 4, I cluster standard errors by firm. Lastly, in columns 5 and 6, I cluster standard errors by firm and time. The table shows estimates of equation 3.5, the firm-level regression of the change in the investment rate on returns, controls, and time fixed effects. The control variables are stock characteristics as described in the data section. The columns show first and second stages of different versions of 2SLS regressions. The instrument is normalized cash dividend-induced trading (DIT). The frequency is annual. I report standard errors in parentheses and Kleibergen and Paap (2006) F-statistics in the table footer. ***, **, and * denote significance at the 10%, 5%, and 1% levels.

Table 3.8: Capital Structure Rebalancing Effect – Time x Industry Fixed Effects

	Debt			Equity		
	(1) OLS	(2) FS	(3) 2SLS	(4) OLS	(5) FS	(6) 2SLS
return	0.0464***		0.257*	-0.0464***		0.0205
	(0.0121)		(0.134)	(0.0161)		(0.0955)
DIT		0.0118***			0.0128***	
		(0.00204)			(0.00219)	
F-statistic		33.16			34.42	
N	98574	98574	98574	91953	91953	91953

This table is analogous to table 3.4, except that I control for time x industry fixed effects instead of time fixed effects. The table shows estimates of equation 3.5, the firm-level regression of the respective dependent variable on returns, controls, and time x industry fixed effects. The control variables are stock characteristics as described in the data section, except that I drop Tobin's Q from the equity issuance regression because it contains book equity which is part of the dependent variable. The columns show first and second stages of different versions of 2SLS regressions. The instrument is normalized cash dividend-induced trading (DIT). The frequency is annual. I report standard errors clustered by time in parentheses and Kleibergen and Paap (2006) F-statistics in the table footer. ***, **, and * denote significance at the 10%, 5%, and 1% levels.

Table 3.9: Capital Structure Rebalancing Effect – Instrument with *Expected* DIT

	Debt			Equity		
	(1) OLS	(2) FS	(3) 2SLS	(4) OLS	(5) FS	(6) 2SLS
return	0.0475***		0.274**	-0.0464***		0.00991
	(0.0120)		(0.115)	(0.0161)		(0.105)
DIT		0.0124***			0.0111***	
		(0.00211)			(0.00199)	
F-statistic		34.72			31.02	
N	98574	98574	98574	91953	91953	91953

This table is analogous to table 3.8, except that the instrumental variable is *expected* dividend-induced trading. The table shows estimates of equation 3.5, the firm-level regression of the respective dependent variable on returns, controls, and time x industry fixed effects. The control variables are stock characteristics as described in the data section, except that I drop Tobin's Q from the equity issuance regression because it contains book equity which is part of the dependent variable. The columns show first and second stages of different versions of 2SLS regressions. The instrument is normalized expected cash dividend-induced trading (DIT). The frequency is annual. I report standard errors clustered by time in parentheses and Kleibergen and Paap (2006) F-statistics in the table footer. ***, **, and * denote significance at the 10%, 5%, and 1% levels.

Bibliography

Adrian, Tobias, Erkko Etula, and Tyler Muir (2014). "Financial Intermediaries and the Cross-Section of Asset Returns". In: *The Journal of Finance* 69.6, pp. 2557–2596.

Almeida, Heitor, Murillo Campello, and Antonio F. Galvao (2010). "Measurement Errors in Investment Equations". In: *The Review of Financial Studies* 23.9, pp. 3279–3328.

Amihud, Yakov and Haim Mendelson (1980). "Dealership Market: Market-making with Inventory". In: *Journal of Financial Economics* 8.1, pp. 31–53.

Anand, Amber, Carsten Tanggaard, and Daniel G. Weaver (2009). "Paying for Market Quality". In: *The Journal of Financial and Quantitative Analysis* 44.6, pp. 1427–1457.

Anand, Amber and Kumar Venkataraman (2016). "Market Conditions, Fragility, and the Economics of Market Making". In: *Journal of Financial Economics* 121.2, pp. 327–349.

Andrade, Sandro C., Charles Chang, and Mark S. Seasholes (2008). "Trading Imbalances, Predictable Reversals, and Cross-Stock Price Pressure". In: *Journal of Financial Economics* 88.2, pp. 406–423.

Anton, Miguel and Christopher Polk (2014). "Connected Stocks". In: *The Journal of Finance* 69.3, pp. 1099–1127.

Aquilina, Matteo, Eric Budish, and Peter O'Neill (2021). "Quantifying the High-Frequency Trading "Arms Race"". In: *The Quarterly Journal of Economics*.

Asquith, Paul, Robert F. Bruner, and David W. Mullins (1983). "The Gains to Bidding Firms from Merger". In: *Journal of Financial Economics* 11.1, pp. 121–139.

Baker, Malcolm, Stefan Nagel, and Jeffrey Wurgler (2007). "The Effect of Dividends on Consumption". In: *Brookings Papers on Economic Activity* 38.1, pp. 231–292.

Baker, Malcolm, Jeremy C. Stein, and Jeffrey Wurgler (2003). "When Does the Market Matter? Stock Prices and the Investment of Equity-Dependent Firms". In: *The Quarterly Journal of Economics* 118.3, pp. 969–1005.

Baker, Malcolm and Jeffrey Wurgler (2002). "Market Timing and Capital Structure". In: *The Journal of Finance* 57.1, pp. 1–32.

Baker, Malcolm and Jeffrey Wurgler (2013). "Chapter 5 - Behavioral Corporate Finance: An Updated Survey". In: *Handbook of the Economics of Finance*. Ed. by George M. Constantinides, Milton Harris, and Rene M. Stulz. Vol. 2. Elsevier, pp. 357–424.

Bakke, Tor-Erik and Toni M. Whited (2010). "Which Firms Follow the Market? An Analysis of Corporate Investment Decisions". In: *The Review of Financial Studies* 23.5, pp. 1941–1980.

Banyi, Monica L., Edward A. Dyl, and Kathleen M. Kahle (2008). "Errors in Estimating Share Repurchases". In: *Journal of Corporate Finance* 14.4, pp. 460–474.

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler (2005). "Comovement". In: *Journal of Financial Economics* 75.2, pp. 283–317.

Bargeron, Leonce, Manoj Kulchania, and Shawn Thomas (2011). "Accelerated Share Repurchases". In: *Journal of Financial Economics* 101.1, pp. 69–89.

Bellia, Mario, Loriana Pelizzon, Marti G. Subrahmanyam, and Darya Yuferova (2020). *Designated Market Makers: Competition and Incentives*. SSRN Scholarly Paper ID 3354400. Rochester, NY: Social Science Research Network.

Ben-Rephael, Azi, Shmuel Kandel, and Avi Wohl (2011). "The Price Pressure of Aggregate Mutual Fund Flows". In: *The Journal of Financial and Quantitative Analysis* 46.2, pp. 585–603. JSTOR: 23018420.

Ben-Rephael, Azi, Shmuel Kandel, and Avi Wohl (2012). "Measuring Investor Sentiment with Mutual Fund Flows". In: *Journal of Financial Economics* 104.2, pp. 363–382.

Berger, Elizabeth (2019). "Selection Bias in Mutual Fund Flow-Induced Fire Sales: Causes and Consequences". In: *Unpublished Working Paper*.

Berk, Jonathan and Richard Green (2004). "Mutual Fund Flows and Performance in Rational Markets". In: *Journal of Political Economy* 112.6, pp. 1269–1295. JSTOR: 10.1086/424739.

Bessembinder, Hendrik, Jia Hao, and Kuncheng Zheng (2020). "Liquidity Provision Contracts and Market Quality: Evidence from the New York Stock Exchange". In: *The Review of Financial Studies* 33.1, pp. 44–74.

Binsbergen, Jules H. van and Christian C. Opp (2019). "Real Anomalies". In: *The Journal of Finance* 74.4, pp. 1659–1706.

Blanchard, Olivier, Changyong Rhee, and Lawrence Summers (1993). "The Stock Market, Profit, and Investment". In: *The Quarterly Journal of Economics* 108.1, pp. 115–136.

Boissel, Charles and Adrien Matray (2019). "Higher Dividend Taxes, No Problem! Evidence from Taxing Entrepreneurs in France". In: *Unpublished Working Paper*.

Bonaime, Alice, Huseyin Gulen, and Mihai Ion (2018). "Does Policy Uncertainty Affect Mergers and Acquisitions?" In: *Journal of Financial Economics* 129.3, pp. 531–558.

Bond, Philip, Alex Edmans, and Itay Goldstein (2012). "The Real Effects of Financial Markets". In: *Annual Review of Financial Economics* 4.1, pp. 339–360.

Boulatov, Alex, Terrence Hendershott, and Dmitry Livdan (2013). "Informed Trading and Portfolio Returns". In: *The Review of Economic Studies* 80.1, pp. 35–72.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan (2014). "High-Frequency Trading and Price Discovery". In: *Review of Financial Studies* 27.8, pp. 2267–2306.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan (2019). "Price Discovery without Trading: Evidence from Limit Orders". In: *The Journal of Finance* 74.4, pp. 1621–1658.

Brogaard, Jonathan, Allen Carrion, Thibaut Moyaert, Ryan Riordan, Andriy Shkilko, and Konstantin Sokolov (2018). "High Frequency Trading and Extreme Price Movements". In: *Journal of Financial Economics* 128.2, pp. 253–265.

Brunnermeier, Markus K. and Lasse Heje Pedersen (2009). "Market Liquidity and Funding Liquidity". In: *The Review of Financial Studies* 22.6, pp. 2201–2238.

Chang, Yen-Cheng, Harrison Hong, and Inessa Liskovich (2015). "Regression Discontinuity and the Price Effects of Stock Market Indexing". In: *The Review of Financial Studies* 28.1, pp. 212–246.

Chen, Honghui, Vijay Singal, and Robert F. Whitelaw (2016). "Comovement Revisited". In: *Journal of Financial Economics* 121.3, pp. 624–644.

Chen, Huaizhi (2020). "Cash Induced Demand". In: *Unpublished Working Paper*, p. 62.

Chen, Qi, Itay Goldstein, and Wei Jiang (2007). "Price Informativeness and Investment Sensitivity to Stock Price". In: *The Review of Financial Studies* 20.3, pp. 619–650.

Chen, Qi, Itay Goldstein, and Wei Jiang (2010). "Payoff Complementarities and Financial Fragility: Evidence from Mutual Fund Outflows". In: *Journal of Financial Economics* 97.2, pp. 239–262.

Chevalier, Judith and Glenn Ellison (1997). "Risk Taking by Mutual Funds as a Response to Incentives". In: *Journal of Political Economy* 105.6, pp. 1167–1200.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam (2000). "Commonality in Liquidity". In: *Journal of Financial Economics* 56.1, pp. 3–28.

Claessens, Stijn and Yishay Yafeh (2013). "Comovement of Newly Added Stocks with National Market Indices: Evidence from Around the World". In: *Review of Finance* 17.1, pp. 203–227.

Clark-Joseph, Adam D., Mao Ye, and Chao Zi (2017). "Designated Market Makers Still Matter: Evidence from Two Natural Experiments". In: *Journal of Financial Economics* 126.3, pp. 652–667.

Cochrane, John H. and Monika Piazzesi (2002). "The Fed and Interest Rates - A High-Frequency Identification". In: *American Economic Review* 92.2, pp. 90–95.

Comerton-Forde, Carole, Terrence Hendershott, Charles M. Jones, Pamela C. Moulton, and Mark S. Seasholes (2010). "Time Variation in Liquidity: The Role of Market-Maker Inventories and Revenues". In: *The Journal of Finance* 65.1, pp. 295–331.

Coughenour, Jay F. and Mohsen M. Saad (2004). "Common Market Makers and Commonality in Liquidity". In: *Journal of Financial Economics* 73.1, pp. 37–69.

Coval, Joshua and Erik Stafford (2007). "Asset Fire Sales (and Purchases) in Equity Markets". In: *Journal of Financial Economics* 86.2, pp. 479–512.

Cremers, K. J. Martijn and Antti Petajisto (2009). "How Active Is Your Fund Manager? A New Measure That Predicts Performance". In: *The Review of Financial Studies* 22.9, pp. 3329–3365.

Dannhauser, Caitlin D. and Jeffrey Pontiff (2019). *FLOW*. SSRN Scholarly Paper ID 3428702. Rochester, NY: Social Science Research Network.

Derrien, Francois, Ambrus Kecskes, and David Thesmar (2013). "Investor Horizons and Corporate Policies". In: *Journal of Financial and Quantitative Analysis* 48.6, pp. 1755–1780.

Dessaint, Olivier, Thierry Foucault, Laurent Fresard, and Adrien Matray (2018). "Noisy Stock Prices and Corporate Investment". In: *The Review of Financial Studies* 32.7, pp. 2625–2672.

Duffie, Darrell (2010). "Presidential Address: Asset Price Dynamics with Slow-Moving Capital". In: *The Journal of Finance* 65.4, pp. 1237–1267.

Easley, David, Marcos López de Prado, Maureen O'Hara, and Zhibai Zhang (2020). "Microstructure in the Machine Age". In: *The Review of Financial Studies*.

Eckbo, B. Espen, Tanakorn Makaew, and Karin S. Thorburn (2018). "Are Stock-financed Takeovers Opportunistic?" In: *Journal of Financial Economics* 128.3, pp. 443–465.

Edelen, Roger M and Jerold B Warner (2001). "Aggregate Price Effects of Institutional Trading: A Study of Mutual Fund Flow and Market Returns". In: *Journal of Financial Economics* 59.2, pp. 195–220.

Edmans, Alex, Itay Goldstein, and Wei Jiang (2012). "The Real Effects of Financial Markets: The Impact of Prices on Takeovers". In: *The Journal of Finance* 67.3, pp. 933–971.

Erickson, Timothy and Toni M. Whited (2000). "Measurement Error and the Relationship between Investment and q". In: *Journal of Political Economy* 108.5, pp. 1027–1057.

Ernst, Thomas (2021). "Stock-Specific Price Discovery From ETFs". In: *Working Paper*.

Fama, Eugene F. and Kenneth R. French (2015). "A Five-factor Asset Pricing Model". In: *Journal of Financial Economics* 116.1, pp. 1–22.

Fama, Eugene F. and Kenneth R. French (2018). "Choosing Factors". In: *Journal of Financial Economics* 128.2, pp. 234–252.

Fama, Eugene F. and James D. MacBeth (1973). "Risk, Return, and Equilibrium: Empirical Tests". In: *Journal of Political Economy* 81.3, pp. 607–636.

Fazzari, Steven M., R. Glenn Hubbard, Bruce C. Petersen, Alan S. Blinder, and James M. Poterba (1988). "Financing Constraints and Corporate Investment". In: *Brookings Papers on Economic Activity* 1988.1, pp. 141–206. JSTOR: 2534426.

Flannery, Mark J. and Kasturi P. Rangan (2006). "Partial Adjustment toward Target Capital Structures". In: *Journal of Financial Economics* 79.3, pp. 469–506.

Frazzini, Andrea and Owen A. Lamont (2008). "Dumb Money: Mutual Fund Flows and the Cross-Section of Stock Returns". In: *Journal of Financial Economics* 88.2, pp. 299–322.

Frazzini, Andrea and Lasse Heje Pedersen (2014). "Betting Against Beta". In: *Journal of Financial Economics* 111.1, pp. 1–25.

Gider, Jasmin, Simon Schmickler, and Christian Westheide (2019). "High-Frequency Trading and Price Informativeness". In: *Working Paper*.

Gilchrist, Simon, Charles P. Himmelberg, and Gur Huberman (2005). "Do Stock Price Bubbles Influence Corporate Investment?" In: *Journal of Monetary Economics* 52.4, pp. 805–827.

Goldstein, Itay, Hao Jiang, and David T. Ng (2017). "Investor Flows and Fragility in Corporate Bond Funds". In: *Journal of Financial Economics* 126.3, pp. 592–613.

Gorodnichenko, Yuriy and Michael Weber (2016). "Are Sticky Prices Costly? Evidence from the Stock Market". In: *American Economic Review* 106.1, pp. 165–199.

Graham, John R and Campbell R Harvey (2001). "The Theory and Practice of Corporate Finance: Evidence from the Field". In: *Journal of Financial Economics*, p. 57.

Green, Jeremiah, John R. M. Hand, and X. Frank Zhang (2017). "The Characteristics That Provide Independent Information about Average U.S. Monthly Stock Returns". In: *The Review of Financial Studies* 30.12, pp. 4389–4436.

Greene, Jason T., Charles W. Hodges, and David A. Rakowski (2007). "Daily Mutual Fund Flows and Redemption Policies". In: *Journal of Banking & Finance* 31.12, pp. 3822–3842.

Greenwood, Robin (2005). "Short- and Long-Term Demand Curves for Stocks: Theory and Evidence on the Dynamics of Arbitrage". In: *Journal of Financial Economics* 75.3, pp. 607–649.

Greenwood, Robin and David Thesmar (2011). "Stock Price Fragility". In: *Journal of Financial Economics* 102.3, pp. 471–490.

Grullon, Gustavo and Roni Michaely (2004). "The Information Content of Share Repurchase Programs". In: *The Journal of Finance* 59.2, pp. 651–680.

Harford, Jarrad and Aditya Kaul (2005). "Correlated Order Flow: Pervasiveness, Sources, and Pricing Effects". In: *Journal of Financial and Quantitative Analysis* 40.1, pp. 29–55.

Harris, Lawrence and Eitan Gurel (1986). "Price and Volume Effects Associated with Changes in the S&P 500 List: New Evidence for the Existence of Price Pressures". In: *The Journal of Finance* 41.4, pp. 815–829.

Hartzmark, Samuel M. and David H. Solomon (2013). "The Dividend Month Premium". In: *Journal of Financial Economics* 109.3, pp. 640–660.

Hasbrouck, Joel and Duane J. Seppi (2001). "Common Factors in Prices, Order Flows, and Liquidity". In: *Journal of Financial Economics* 59.3, pp. 383–411.

Hasbrouck, Joel and George Sofianos (1993). "The Trades of Market Makers: An Empirical Analysis of NYSE Specialists". In: *The Journal of Finance* 48.5, pp. 1565–1593.

Hayashi, Fumio (1982). "Tobin's Marginal q and Average q: A Neoclassical Interpretation". In: *Econometrica* 50.1, pp. 213–224. JSTOR: 1912538.

He, Zhiguo, Bryan Kelly, and Asaf Manela (2017). "Intermediary Asset Pricing: New Evidence from Many Asset Classes". In: *Journal of Financial Economics* 126.1, pp. 1–35.

He, Zhiguo and Arvind Krishnamurthy (2013). "Intermediary Asset Pricing". In: *American Economic Review* 103.2, pp. 732–70.

Hendershott, Terrence, Dmitry Livdan, and Norman Schürhoff (2015). "Are Institutions Informed about News?" In: *Journal of Financial Economics* 117.2, pp. 249–287.

Hendershott, Terrence and Albert J. Menkveld (2014). "Price Pressures". In: *Journal of Financial Economics* 114.3, pp. 405–423.

Hendershott, Terrence and Mark S Seasholes (2007). "Market Maker Inventories and Stock Prices". In: *American Economic Review* 97.2, p. 9.

Hoberg, Gerard and Gordon Phillips (2016). "Text-Based Network Industries and Endogenous Product Differentiation". In: *Journal of Political Economy* 124.5, pp. 1423–1465.

Hou, Kewei, Chen Xue, and Lu Zhang (2015). "Digesting Anomalies: An Investment Approach". In: *The Review of Financial Studies* 28.3, pp. 650–705.

Jensen, Michael C. and Richard S. Ruback (1983). "The Market for Corporate Control: The Scientific Evidence". In: *Journal of Financial Economics* 11.1, pp. 5–50.

Jotikasthira, Chotibhak, Christian Lundblad, and Tarun Ramadorai (2012). "Asset Fire Sales and Purchases and the International Transmission of Funding Shocks". In: *The Journal of Finance* 67.6, pp. 2015–2050. JSTOR: 23324367.

Kaplan, Steven N. and Luigi Zingales (1997). "Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints?" In: *The Quarterly Journal of Economics* 112.1, pp. 169–215. JSTOR: 2951280.

Karolyi, G. Andrew, Kuan-Hui Lee, and Mathijs A. van Dijk (2012). "Understanding Commonality in Liquidity around the World". In: *Journal of Financial Economics* 105.1, pp. 82–112.

Kaul, Aditya, Vikas Mehrotra, and Randall Morck (2000). "Demand Curves for Stocks Do Slope Down: New Evidence from an Index Weights Adjustment". In: *The Journal of Finance* 55.2, pp. 893–912.

Kayhan, Ayla and Sheridan Titman (2007). "Firms' Histories and Their Capital Structures". In: *Journal of Financial Economics* 83.1, pp. 1–32.

Kervel, Vincent Van and Albert J. Menkveld (2019). "High-Frequency Trading around Large Institutional Orders". In: *The Journal of Finance* 74.3, pp. 1091–1137.

Khan, Mozaffar, Leonid Kogan, and George Serafeim (2012). "Mutual Fund Trading Pressure: Firm-Level Stock Price Impact and Timing of SEOs". In: *The Journal of Finance* 67.4, pp. 1371–1395.

Kirilenko, Andrei, Albert S. Kyle, Mehrdad Samadi, and Tugkan Tuzun (2017). "The Flash Crash: High-Frequency Trading in an Electronic Market". In: *The Journal of Finance* 72.3, pp. 967–998.

Kleibergen, Frank and Richard Paap (2006). "Generalized Reduced Rank Tests Using the Singular Value Decomposition". In: *Journal of Econometrics* 133.1, pp. 97–126.

Koch, Andrew, Stefan Ruenzi, and Laura Starks (2016). "Commonality in Liquidity: A Demand-Side Explanation". In: *The Review of Financial Studies* 29.8, pp. 1943–1974.

Koijen, Ralph S. J. and Motohiro Yogo (2019). "A Demand System Approach to Asset Pricing". In: *Journal of Political Economy* 127.4, pp. 1475–1515.

Korajczyk, Robert A and Dermot Murphy (2019). "High-Frequency Market Making to Large Institutional Trades". In: *The Review of Financial Studies* 32.3, pp. 1034–1067.

Kvamvold, Joakim and Snorre Lindset (2018). "Do Dividend Flows Affect Stock Returns?" In: *Journal of Financial Research* 41.1, pp. 149–174.

Leary, Mark T. and Michael R. Roberts (2005). "Do Firms Rebalance Their Capital Structures?" In: *The Journal of Finance* 60.6, pp. 2575–2619.

Lee, Charles M.C. and Eric C. So (2017). "Uncovering Expected Returns: Information in Analyst Coverage Proxies". In: *Journal of Financial Economics* 124.2, pp. 331–348.

Li, Jiacui (2019). "Slow-Moving Capital and Flow-Driven Common Factors in Stock Returns". In: *Unpublished Working Paper*.

Lou, Dong (2012). "A Flow-Based Explanation for Return Predictability". In: *The Review of Financial Studies* 25.12, pp. 3457–3489.

Lou, Xiaoxia and Albert Y. Wang (2018). "Flow-Induced Trading Pressure and Corporate Investment". In: *Journal of Financial and Quantitative Analysis* 53.1, pp. 171–201.

Madhavan, Ananth and Seymour Smidt (1991). "A Bayesian Model of Intraday Specialist Pricing". In: *Journal of Financial Economics* 30.1, pp. 99–134.

Madhavan, Ananth and Seymour Smidt (1993). "An Analysis of Changes in Specialist Inventories and Quotations". In: *The Journal of Finance* 48.5, pp. 1595–1628.

Madhavan, Ananth and George Sofianos (1998). "An Empirical Analysis of NYSE Specialist Trading". In: *Journal of Financial Economics* 48.2, pp. 189–210.

Malceniece, Laura, Kārlis Malcenieks, and Tālis J. Putniņš (2019). "High Frequency Trading and Comovement in Financial Markets". In: *Journal of Financial Economics* 134.2, pp. 381–399.

Marshall, Ben R., Nhut H. Nguyen, and Nuttawat Visaltanachoti (2012). "Commodity Liquidity Measurement and Transaction Costs". In: *The Review of Financial Studies* 25.2, pp. 599–638.

Meling, Tom Grimstvedt (2021). "Anonymous Trading in Equities". In: *The Journal of Finance* 76.2, pp. 707–754.

Menkveld, Albert J. (2016). "The Economics of High-Frequency Trading: Taking Stock". In: *Annual Review of Financial Economics* 8.1, pp. 1–24.

Menkveld, Albert J. and Ting Wang (2013). "How Do Designated Market Makers Create Value for Small-Caps?" In: *Journal of Financial Markets* 16.3, pp. 571–603.

Mitchell, Mark, Todd Pulvino, and Erik Stafford (2004). "Price Pressure around Mergers". In: *The Journal of Finance* 59.1, pp. 31–63.

Mullainathan, Sendhil and Jann Spiess (2017). "Machine Learning: An Applied Econometric Approach". In: *Journal of Economic Perspectives* 31.2, pp. 87–106.

Nakamura, Emi and Jón Steinsson (2018). "High-Frequency Identification of Monetary Non-Neutrality: The Information Effect". In: *The Quarterly Journal of Economics* 133.3, pp. 1283–1330.

Nimalendran, M. and Giovanni Petrella (2003). "Do Thinly-Traded Stocks Benefit from Specialist Intervention?" In: *Journal of Banking & Finance*. The Future of Stock Exchanges in a Globalizing World 27.9, pp. 1823–1854.

Norli, Oyvind, Charlotte Ostergaard, and Ibolya Schindele (2014). "Liquidity and Shareholder Activism". In: *The Review of Financial Studies* 28.2, pp. 486–520.

Ogden, Joseph P. (1994). "A Dividend Payment Effect in Stock Returns". In: *Financial Review* 29.3, pp. 345–369.

Pasquariello, Paolo and Clara Vega (2015). "Strategic Cross-Trading in the U.S. Stock Market*". In: *Review of Finance* 19.1, pp. 229–282.

Perotti, Pietro and Barbara Rindi (2010). "Market Makers as Information Providers: The Natural Experiment of STAR". In: *Journal of Empirical Finance* 17.5, pp. 895–917.

Phillips, Gordon M. and Alexei Zhdanov (2013). "R&D and the Incentives from Merger and Acquisition Activity". In: *The Review of Financial Studies* 26.1, pp. 34–78.

Polk, Christopher and Paola Sapienza (2009). "The Stock Market and Corporate Investment: A Test of Catering Theory". In: *Review of Financial Studies* 22.1, pp. 187–217.

Rakowski, David (2010). "Fund Flow Volatility and Performance". In: *Journal of Financial and Quantitative Analysis* 45.1, pp. 223–237.

Rakowski, David and Xiaoxin Wang (2009). "The Dynamics of Short-term Mutual Fund Flows and Returns: A Time-series and Cross-sectional Investigation". In: *Journal of Banking & Finance*. Financial Globalisation, Risk Analysis and Risk Management 33.11, pp. 2102–2109.

Raman, Vikas, Michel A. Robe, and Pradeep K. Yadav (2021). *Man vs. Machine: Liquidity Provision and Market Fragility*. SSRN Scholarly Paper ID 3757848. Rochester, NY: Social Science Research Network.

Schmickler, Simon (2020). "Identifying the Price Impact of Fire Sales Using High-Frequency Surprise Mutual Fund Flows". In: *Unpublished Working Paper*.

Schmidheiny, Kurt and Sebastian Siegloch (2020). "On Event Studies and Distributed-Lags in Two-Way Fixed Effects Models: Identification, Equivalence, and Generalization". In: *SSRN Electronic Journal*.

Shleifer, Andrei (1986). "Do Demand Curves for Stocks Slope Down?" In: *The Journal of Finance* 41.3, pp. 579–590.

Skjeltorp, Johannes Atle and Bernt Arne Odegaard (2015). "When Do Listed Firms Pay for Market Making in Their Own Stock?" In: *Financial Management* 44.2, pp. 241–266.

Stoll, Hans R. (1978). "The Supply of Dealer Services in Securities Markets". In: *The Journal of Finance* 33.4, pp. 1133–1151.

Titman, Sheridan, K. C. John Wei, and Feixue Xie (2004). "Capital Investments and Stock Returns". In: *Journal of Financial and Quantitative Analysis* 39.4, pp. 677–700.

Tookes, Heather E. (2008). "Information, Trading, and Product Market Interactions: Cross-sectional Implications of Informed Trading". In: *The Journal of Finance* 63.1, pp. 379–413.

Venkataraman, Kumar and Andrew C. Waisburd (2007). "The Value of the Designated Market Maker". In: *The Journal of Financial and Quantitative Analysis* 42.3, pp. 735–758.

Vermaelen, Theo (1981). "Common Stock Repurchases and Market Signalling: An Empirical Study". In: *Journal of Financial Economics* 9.2, pp. 139–183.

Wardlaw, Malcolm (2018). *Measuring Mutual Fund Flow Pressure As Shock to Stock Returns*. SSRN Scholarly Paper ID 3248750. Rochester, NY: Social Science Research Network.

Wardlaw, Malcolm (2020). "Measuring Mutual Fund Flow Pressure As Shock to Stock Returns". In: *Journal of Finance* Forthcoming.