Title
Research on Asset Prices

Author
Sheng, Yifei

Publication Date
2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

RESEARCH ON ASSET PRICES

A dissertation submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

ECONOMICS

by

Yifei Sheng

December 2020

The Dissertation of Yifei Sheng is approved:

Professor Weishi Steadmon, Chair

Professor Kenneth Kletzer

Professor Hikaru Saijo

Quentin Williams
Interim Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

Research on Asset Prices

by

Yifei Sheng

In this paper, I develop researches on three different financial assets: 1) institutional herding in China's security market, and its relation with stock prices; 2) news effects on WTI crude oil futures prices; and 3) the interplay between the monetary policy and fiscal policy on changing the supply of long-term treasury bonds and its further impact on long-term interest rate and other macroeconomic indicators.

The first chapter provides an examination of institutional herding in China's securities market, it addresses the following four research questions: (1) Does China's stock market exhibit herding behavior among institutional investors? (2) Does institutional investors' trade follow their own, or others' previous trade? (3) Are herding behaviors different among different types of institutional investors? and (4) What are possible explanations for institutional herding in China? Empirical results show that most types of institutional investors exhibit herding. Foreign institutional investors show some evidence on following domestic players lag trades when making sell decisions. Funds exhibit opposite herding behavior, i.e., they decrease in buy(sell) this quarter following an increase in their buy last quarter. Combining realization of stock returns, they prefer to sell securities to realize profits given previous higher returns and fail to buy more shares of securities which will have higher future returns.

In order to uncover the news impact on the price of WTI crude oil futures, the second chapter applies supervised and unsupervised machine learning algorithms to conduct news sentiment and topic analysis. With the assumption that the crude oil futures market is efficient enough to respond quickly to new information, this chapter obtains high-frequency price and news from the Bloomberg terminal. Using results from logistic regression and K-means clustering, this chapter defines the positive score and topic for each news article as inputs for the final logistic regression. The regression results show that the "World Crude Oil" news is more positively correlated with price increase than other topics. Moreover, the "WTI Crude Oil" news has the highest correlation with the price increase as the positive score increases.

Third chapter investigates the policy interplay between Federal Reserve LSAP program and long-term bonds supply policy by the Treasury department. Under zero lower bound, central bank reduces longer-term interest rates to stimulate aggregate demand by purchasing longer-term Treasury debt securities through asset purchase programs. Meanwhile, debt supply decision, which is made by the Treasury department, is found to increase the longer-term Treasury debt securities outstanding to the private sector during the same period of time. This supply behavior could deteriorate the effectiveness of LSAP. To study how the two main policies interact, this chapter develops a Dynamic Stochastic General Equilibrium Model (DSGE) by incorporating both the central bank's quantitative easing policy and the Treasury department's long-term bond supply policy.

To my beloved boy,

Nuoyi Xu.

Acknowledgments

As a first generation graduate student in my family, the decision of pursuing a Ph.D. in economics was greatly challenging and extraordinary difficult to me, especially when I was around 30 year old. During the past five years of my Ph.D. journal, I have left the program due to my personal medical condition for a while. However, with the enthusiasm for research, I decided again to chase my "dream" of pursuing the Ph.D. and returned back to the program. In China, when people are at the age of thirty, they should be able to determine the life goals and their own development direction. When they become forty, they would have no doubts about their lives. I have never regretted the decision I made in my thirties, even after experiencing all difficulties and struggles. Life is short and full of uncertainty. In the next stage of my life journal, I would never know exactly what my future is going to be, but confident in facing with any situation without fears and doubts.

I want to thank my advisor, Prof. Grace Gu. Her valuable guidance and patient support helps me to choose the right direction of research and successfully complete my dissertation. I want to thank my committee members, Prof. Kenneth Kletzer and Prof. Hikaru Saijo, for their time and efforts to give me suggestions on my dissertation. I also want to thank Prof. Carl Walsh, who was the chair of my oral qualification exam, for his great advice on my first chapter and strong support after I returned back to the program. I would also like to acknowledge our program coordinator Sandra Reebie, who provides unbelievable support, especially in helping me get through all the readmission process. Without her help, I would not have returned to the Ph.D. program.

In addition, I would like to thank my husband, Qiang Xu, who has always been with me. I could not have finished this program and completed my dissertation without the support of him. Finally, I want to thank my parents for everything they have gave me.

Chapter 1

Institutional herding in China

In January 2020, the government of United State of America and the government of People's Republic of China reached a consensus on the first phase of economic and trade agreement. In Chapter 4, Article 4.7 *Securities, Fund Management, Futures Services*, the two parties agree that foreign equity limits shall be eliminated and wholly U.S.-owned services suppliers shall be allowed to participate in securities, fund management, and futures sectors, no later than April 1, 2020. There are two major channels for foreign investors to participate in China's stock market: the Qualified Foreign Institutional Investors (QFIIs) and the Shanghai- and Shen Zhen-Hong Kong Connect program. China launched QFIIs in 2003. In January 2019, total investment quota for QFII was increased from \$150 billion to \$300 billion US dollars. The Shanghai and Shen Zhen-Hong Kong Connect was launched in November 2014 and December 2016. In August 2020, trading volumes have reached over 1 trillion RMB.

In addition to the openness of China's security market to foreign investors, domestic institutional investors also have developed, both in depth and variety, over the past decades. In Section 1.2.2, I provide an analysis on the evolution of institutional investment in Shanghai's Stock Exchange[1], and several key findings have been revealed as followed:

First, the China's stock market is dominated by individual investors, with their share holdings about 4 times as held by institutional investors. Nonetheless, there is an increasing trend both in number and participation of institutional investors. The number of institutional investors have increased by 8 times from 2008 to 2019. During the same time period, average number of institutional investors per stock also have increased by 4 times. Secondly, foreign equity investment is still relatively small in trading volume, comparing to domestic players. Third, domestic funds exhibit different trend from other types of institutional investors, in that the number of stocks held have increased while the average market value of share holding have increased only mildly or even decreased over time.

An equity market with the presence of institutional investors could potentially be associated with phenomenon of institutional herding. According to Sias(2004), institutional investors herd by following their own or others into and out of the same securities. It has been argued that stock market excess volatility and price deviation from fundamental values could be explained by herding among investors. With the

[1]There are two stock exchanges in mainland China: Shanghai Stock Exchange and Shen Zhen Stock Exchange. Firms listed on these two exchanges are mutually exclusive, i.e., they can only be listed on one of the two exchanges. In Shanghai Stock Exchange, firms are more likely to have large market capitalization. Small cap and growth companies usually choose to go public on Shen Zhen Stock Exchange.

growth of institutional investment in China's stock market, it is crucial to measure the existence of herding among institutional investors in China. Moreover, foreign investors, given their unfamiliarity with local market, may also commit to herding by following trading of domestic players. Thus, it is worthwhile to examine difference in herding, if exists, among domestic and foreign institutional investors.

In this chapter, four research questions have been examined: (1) Does China's stock market exhibit herding behavior among institutional investors? (2) Does institutional investors' trade follow their own, or others' previous trade? (3) Are herding behaviors different across different types of institutional investors? and (4) What are possible explanations for institutional herding in China's stock market?

To examine the existence of herding among institutional investors, following Sias(2004), I first report relation between institutional demand as a whole group in current time and its lag demand in prior period, and then extend the test by decomposing institutional investors into different types, and report corresponding results respectively. It is found that institutional investors in China, as a whole, do not exhibit the herding behavior defined by Sias(2004). Due to the fact that number of funds and securities they trade outweigh others in China, the result found in aggregate level is dominated by funds trading behavior. After decomposing institutional investors into different types, results show strong evidence of herding among institutional investors, except for funds.

Further analysis suggests that institutional investors more often follow their own type of lag trades. Even though, foreign institutional investors show some evidence on following domestic players' previous sell. Funds are different from others, in that the relation between their current and lag demand is statistically negative. Assuming that funds established by different ownership structure and specialize in different investment styles may focus on securities in particular industries/sectors, I classified securities into different sectors, and test funds herding across sectors. Results are robust that funds' current demand is negatively correlated with its lag demand in all sectors.

Last, to determine the source of herding, momentum trading, correlations between institutional demand and period returns(prior, current and subsequent), and herding by firm size are examined. Results demonstrate that there is no clear evidence of momentum trading and subsequent price reversal among institutional investors, and their herding behavior is stronger in large-cap securities, demonstrating investigative herding. Combining the realization of returns, I conclude that domestic funds are "conservative" in investment: they sell securities to realize profits given previous higher returns and fail to increase holding of securities which will have higher future returns. To my knowledge, this chapter is among the first to provide a thorough examination of herding among different types of institutional investors in China's stock market, in particular, taking into account the Shanghai-Hong Kong connect program.

1.1 Literature Review

Previous literature have studied the theoretical foundation and provide empirical evidence of institutional herding. There are three widely adopted methodologies of measuring herding. Lakonishok, Shleifer and Vishny (1992) design a measure of herding by testing cross-sectional relation of buying the same securities in the same time between money managers and other managers. Christie and Huang (1995) propose another methodology by calculating the standard deviation of individual returns as a proxy of equity return dispersion, which is predicted to be low when there exists herd behavior.

Sias(2004) provides a different approach by measuring the cross-sectional dependence of institutional buy in adjacent time period, directly. In each period, a cross-sectional regression model is applied, and the average coefficient from regressions in all time periods and the associated t-statistic are reported. The t-statistics are computed from Fama-MacBeth standard errors, which is the time-series standard error of cross-sectional averages. Fama-MacBeth procedure is widely used in finance literature, as it is designed to account for time effect in panel data.

Peterson (2009) compares different approaches to estimate standard errors in finance panel data and concludes that with only the presence of time effect, both Fama-MacBeth estimates and standard error clustered by time are unbiased. If the data is associated with both time and firm effect, the standard error clustered by both dimensions provides an unbiased estimate, giving sufficient number of clusters in each dimension. In this study, I adopt Sias(2004)'s methodology in measuring institutional herd. However, instead of using the Fama-MacBeth standard errors, I apply a panel regression with standard error clustered by both dimensions to test for herding behavior over adjacent quarters, taking account of potential firm effect in data set.

Some studies focus on explaining motivations for herding. Sias(2004) classifies them into five possible categories for institutional herding. Informational cascades results from investors trading with the herd as they infer information from other's trades. Investigative herding appears when institutional investors follow correlated signals of information. These two are driven by fundamental information, thus exhibit no subsequent price reversal. Reputational herding results from institutional investors preventing from acting differently from others due to a reputational cost. Fad is another reason that institutional investors face and result in herd. Last, specific characteristics of firms may attract investors to trade with herd. They are motivated by non-information based reasons, thus always associate with subsequent price reversal.

Many empirical work have studied evidence of herding among institutional investors in both developed and emerging financial markets. Nofsinger and Sias (1999) compare herding behavior between institutional and individual investors in the US market, and document that institutional investors trade more with herd than individual investors, and institutional herding is strongly positively correlated with lag stock returns. Choi and Sias (2009) investigate institutional industry herding in the US, by decomposing firms into different industries.

Choi and Sakiba (2015) examine institutional herding in 41 countries, and find that institutional herding occurs more often in markets with low levels of information asymmetry, and is more likely driven by investigative herding. Chang et al. (2000) investigate institutional herding in Asia financial markets, including Hong Kong, Japan, South Korea, and Taiwan. They document strong evidence of herding in South Korea and Taiwan. Chen et al. (2008) examine herding by QFIIs in Taiwan's market, and find that QFIIIs commit to herd in picking securities to trade in Taiwan and their herding is associated with high past return and more likely to occur in large-cap securities.

Furthermore, a growing body of literature has extended research on herding behavior in China's financial market. Li et al. (2017) compares differences in herding between individual and institutional investors by analyzing daily trading data for all the component stocks of the SSE[2] 180 Stock Index from 2002 to 2004. Li and Wang (2008), applying the same data set, document herding behavior among Chinese institutional investors, but is largely limited to large stocks, implying the presence of investigative herding. Some other literature focus on differences between domestic and foreign investors. Tan et al. (2008) examine herding behavior within and across the Shanghai and Shenzhen stock markets, and across A-share and B-share markets. Liu et al. (2014) compare investment preferences between Chinese domestic funds and QFIIs.

This chapter contributes to institutional herding literature on three issues. First, I investigate herding behavior within and across different types of institutional investors, including both domestic and foreign investors. In particular, I add trading behavior of investors from Shanghai-Hong Kong connect program. Second, Instead of using Fama-MacBetch approach, I use panel regression with standard error clustered by both time and firm effect for testing institutional herding, which provides unbiased estimate for data with presence of time and potential firm effect. Third, I evaluate institutional herding by extending the data set including all institutional traded securities listed on Shanghai Stock Exchange to 2019q3.

[2]SSE: Shanghai Stock Exchange.

The structure of this chapter is: Section 3 presents the data and descriptive statistics of institutional investors in Shanghai Stock Exchange. Section 4 first describe the definition of key variables for examining herding, and then tests for herding within and across institutions are reported. In Section 5, the potential source of herding is examined. Section 6 concludes the results.

1.2 Data and Descriptive Statistics

1.2.1 Data Source and Institutional Investors

The data consists of two components: institutional investors' share holdings and stock's financial information, including earning per share (EPS), price to earning (PE) ratio, return on equity (ROE), and asset to liability ratio. Stocks studied in this chapter are all listed on Shanghai Stock Exchange, which is one of the two stock exchanges in China (the other is Shenzhen Stock Exchange). The data set is taken from Choice Database, similar to Wind Database, collecting all the Chinese stocks' trading data, public firms' financial reports, and number of shares held by institutional investors. The China Security Regulatory Commission (CSRC) requires institutional investors to report their investment and financial data every quarterly, thus the frequency of institutional share holdings is quarter in my data set, ranging from 2008q1 to 2019q3, for a total of 47 quarters.

In this chapter, institutional investors include both domestic and foreign investors. They are classified into six types: domestic fund investors, domestic insurance investors, government fund investors, 'other' domestic institutional investors, Qualified Foreign Institutional Investors(QFIIs), and Shanghai-Hong Kong Connect (SH-HK) foreign investors. Qualified insurance companies and enterprise annuities are included in domestic insurance investors. The government fund consists of the National Social Security Fund, the China SAFE [3] Investment Limited, and the government pension fund. 'Other' domestic institutions contains domestic trusts, brokerages, banks and other financial companies.

1.2.2 Descriptive Statistics

In order to provide a better understanding of the evolution of institutional investment in China's stock market, this chapter reports basic statistics from two dimensions: firms and institutional investors' dimension. Due to the shortage of SH-HK institutional level data, only five types (excluding SH-HK) of institutional investors' statistics will be reported in 1.2.2.2.

1.2.2.1 Firms' Dimension

Table[1.1] first reports the annual average number of institutional investors per stock from 2008 to 2019. The second column shows the annual average number of all institutional investors, including all types. In 2019, for each stock, there are 60.3 institutional investors per stock on average, which is about 4 times as in 2008. This implies that there is an increasing trend of institutional investors in each stock in

[3]SAFE: State Administration of Foreign Exchange.

the China's stock market. Column 3-8 further decompose into the average number for different type's institutional investors. The average number of domestic funds and SH-HK are much higher than other institutional types. Among all types, average number of domestic funds increases the most, from 17.5 in 2008 to 47.6 in 2019. Even though 'other' investors' number increases by about 3 times, due to its magnitude, it is less crucial in explaining the primary increasing trend of all institutional investors.

Table[1.2] shows the average number of stocks held by institutional investors. The last column reports the number of stock-year observations with different amount and type of institutional investors in my data set. The top panel shows numbers for all types of institutional investors. The average number of stocks that held by at least one institutional investor is 1,238 in 2019, which is more than 2 times than it in 2008. The average number of stocks with more than 5 institutional investors also shows an increasing trend over time, so do those with more than 10 and 20 institutional investors. In recent three years (2017 - 2019), number of stocks with 5 institutional investors keeps about 80% of it with at least one institutional investor[4]. Although the total number of stocks listed on Shanghai Stock Exchange increases over time, the top panel still illustrates that institutional investors increase their diversification in investing in China's stock market, and more stocks are invested by more institutional investors over time.

The middle and bottom panel report average number of stocks with at least 1 and 5 institutional investors of each type, respectively. Numbers of stocks invested by

[4]The ratio is about 70% for those with more than 10 institutional investors and 50% for those with more than 20 institutional investors.

Table 1.1: Average Number of Institutional Investors per Stock

Year	ALL	FUNDS	INSUR.	GOV.	OTHER	QFII	SH-HK
2008	15.6	17.5	1.5	1.2	1.4	1.4	
2009	19.2	20.1	1.4	1.3	1.8	1.4	
2010	19.7	19.4	1.5	1.3	2.2	1.3	
2011	21.8	21.7	1.7	1.3	2.6	1.3	
2012	26.1	26.1	1.7	1.3	3.1	1.3	
2013	26.8	27.8	1.5	1.4	3.4	1.4	
2014	29.1	28.6	1.4	1.4	4.0	1.5	
2015	31.8	29.9	1.4	1.7	4.2	1.4	
2016	44.2	42.0	1.5	1.9	4.2	1.4	
2017	61.8	52.5	1.5	1.9	4.3	1.3	21.9
2018	64.0	52.1	1.5	1.8	4.2	1.3	22.8
2019	60.3	47.6	1.5	1.8	4.3	1.2	23.9
All period	38.5	34.9	1.5	1.7	3.6	1.4	22.9

Note: For each quarter, I calculate the average number of institutional investors over stocks that are held by institutional investors within that quarter. After that, I add each year's quarterly numbers and divide it by number of quarters to be the annual average number of institutional investors per stock reported in this table. For example, there are only 3 quarters in my data set in 2019, thus numbers in the row for year 2019 is calculated by averaging quarterly numbers over 3 quarters.

funds investors over time cover more than 80% as those by all institutional investors, implying that funds investors are well diversified in equity investment. SH-HK investors are also well diversified comparing to the remaining types of institutional investors.

Nonetheless, Table[1.3] reveals that Shanghai stock market is still dominated by individual investors, with more than 80% of shares, on average, held by individual investors. The first panel of Table[1.3] shows the average percentage of share holdings for all institutional investors. Results present that stocks with more institutional investors involved, higher percentage of its shares are held by those investors. However, comparing to individual investors, institutional investors' share holdings account only one fourth of individuals' share holdings[5]. The second and third panels decompose institutional investors into domestic and foreign institutions. Domestic institutional investors' average percentage of share holdings weights about 4 times as foreign institutional investors, which implies that domestic institutions involve deeper than foreign investors in China's stock market.

1.2.2.2 Institutional Investors' Dimension

Table[1.4] reports the aggregate number of institutional investors and market value of their share holdings over time. Due to the lack of trading data, statistics for SH-HK investors is not reported. Results indicate the number of all institutional investors in 2019 has increased by 8 times from 2008, and so do market value (by more than 4 times). All types of institutional investors also show this similar increasing trend.

[5]Without the data of individual investors accounts number and their trading data, it is imprecise to conclude that price movement in Shanghai Stock Exchange is also dominated by individual investors.

Table 1.2: Average number of Stocks with:

	2008	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**Total**
>= 1 **inst.**	579	642	743	789	789	786	854	1,004	1,084	1,237	1,239	1,238	42,686
>= 5 **inst.**	340	407	506	538	519	498	576	790	891	1,008	1,015	984	31,304
>= 10 **inst.**	209	253	320	353	365	365	420	586	694	824	878	864	23,652
>= 20 **inst.**	119	156	194	211	246	267	303	382	480	637	695	666	16,760
>= 1 **funds**	476	560	674	693	697	668	758	925	1,029	1,127	1,107	1,083	38,100
>= 1 **insur.**	119	149	221	286	265	200	180	230	285	267	206	173	10,142
>= 1 **gov.**	59	96	142	115	117	161	205	461	590	567	564	584	14,051
>= 1 **other**	255	319	428	539	533	535	607	728	699	797	627	560	25,945
>= 1 **qfii**	79	90	107	88	81	102	117	100	107	130	128	168	5,014
>= 1 **sh-hk**										736	766	807	7,693
>= 5 **funds**	304	358	431	441	431	422	490	658	778	836	786	680	25,776
>= 5 **insur.**	2	1	3	3	5	2	2	3	5	4	2	2	120
>= 5 **gov.**	-	1	1	1	-	1	2	6	13	15	6	9	195
>= 5 **other**	3	16	42	71	80	94	128	177	180	184	132	103	4,731
>= 5 **qfii**	1	2	1	1	2	3	4	2	2	2	1	1	61
>= 5 **sh-hk**										679	721	762	7,206

Note: The last column reports the number of samples under each condition. E.g., there are in total 42,686 stock-year observations with at least 1 institutional investor in my data set.

Table 1.3: Average Percentage of Share holdings

YEAR		2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
ALL	>= 1 inst.	18.3%	15.5%	16.4%	14.8%	13.8%	12.2%	11.4%	14.9%	15.9%	16.0%	14.8%	14.0%
	>= 5 inst.	29.3%	22.9%	22.5%	20.5%	19.8%	18.4%	15.9%	18.1%	18.7%	19.1%	17.9%	17.6%
	>= 12 inst.	39.1%	29.8%	30.2%	27.4%	24.6%	22.4%	19.9%	21.7%	22.1%	21.9%	20.5%	19.9%
Domestic	>= 1 inst.	17.9%	15.1%	16.0%	14.5%	13.5%	11.8%	10.9%	14.6%	15.6%	15.1%	13.3%	12.0%
	>= 5 inst.	28.6%	22.4%	22.0%	20.0%	19.4%	17.8%	15.3%	17.7%	18.3%	18.1%	16.1%	15.0%
	>= 12 inst.	38.3%	29.2%	29.5%	26.8%	24.0%	21.6%	19.1%	21.2%	21.7%	20.6%	18.3%	16.8%
Foreign	>= 1 inst.	3.3%	2.6%	2.5%	2.9%	3.0%	3.2%	3.2%	3.1%	2.9%	1.7%	2.3%	2.9%
	>= 5 inst.	3.7%	3.0%	2.8%	3.1%	3.3%	3.6%	3.6%	3.3%	3.1%	1.8%	2.4%	3.1%
	>= 12 inst.	4.2%	3.3%	3.0%	3.5%	3.7%	4.0%	3.9%	3.5%	3.5%	1.9%	2.5%	3.4%
Individual	>= 1 inst.	81.7%	84.5%	83.6%	85.2%	86.2%	87.8%	88.6%	85.1%	84.1%	84.0%	85.2%	86.0%
	>= 5 inst.	70.7%	77.1%	77.5%	79.5%	80.2%	81.6%	84.1%	81.9%	81.3%	80.9%	82.1%	82.4%
	>= 12 inst.	60.9%	70.2%	69.8%	72.6%	75.4%	77.6%	80.1%	78.3%	77.9%	78.1%	79.5%	80.1%

Note: $obs = \frac{Avg.\ share(\%)\ of\ type_i\ institutional\ investors}{Avg.\ share(\%)\ of\ all\ institutional\ investors + Avg.\ share(\%)\ of\ individual\ investors}$.

'ALL' contains all types of institutional investors listed in 1.2.1. 'Domestic' includes four domestic institutional investors: funds, gov, insurance, and other. 'Foreign' includes two foreign institutional investors: QFII and SH-HK investors. The sum of 'ALL' and 'Individual' is 100%.

Among others, the number of funds investors is the largest, and government funds held the highest market value of shares, with more than ten thousand billions of RMB in 2018.

Table[1.5] shows the average market value of institutional holdings per stock. Different from results in Table[1.4], funds' average market value per stock have dropped significantly over time, from 2,972.79 millions in 2008 to 441.82 millions in 2019, while numbers of other types have kept increasing mildly. Given the results that number of funds investors increases and number of stocks traded by funds also increases, funds' dropping in average market value implies that funds investment strategy in Shanghai stock market becomes more diversified, and might be different from other types of institutional investors.

1.2.3 Key Findings from Descriptive Statistics

Results from descriptive statistics demonstrate the following 4 facts: First, it affirms the consensus from previous literature that the China's stock market is dominated by individual investors, whose share holdings is about 4 times as held by institutional investors. Second, there is an increasing trend both in number and participation of institutional investors. In other word, institutional investors play more important and deeper role in China's stock market, with the number of institutional participants has increased by 8 times and average number of institutional investors per stock increased by 4 times. Third, foreign stock investment in China is still relatively small, which is about one fourth of the size of domestic players. Last, domestic funds investments are

Table 1.4: Number of Institutional Investors and Total Market Value of Share Holding

YEAR	ALL		FUNDS		GOV.		INSUR.		OTHER		QFII	
	cnt	**MV_SUM**	**cnt**	**MV_SUM**	**cnt**	**MV_SUM**	**cnt**	**MV_SUM**	**cnt**	**MV_SUM**	**cnt**	**MV_SUM**
2008	690	4,622.30	330	3,930.03	19	42.42	55	289.28	244	234.56	40	126.01
2009	883	8,226.78	441	5,166.48	25	2,141.36	73	321.05	298	430.35	45	167.54
2010	1,287	12,506.04	589	5,245.32	22	4,219.98	113	2,261.52	513	559.82	49	219.39
2011	1,670	12,268.96	719	4,781.18	27	4,555.55	131	2,015.36	741	682.48	51	234.39
2012	1,789	11,622.33	831	4,279.05	36	4,327.98	107	2,183.13	762	584.12	52	248.05
2013	2,105	12,126.67	981	4,220.46	43	5,052.25	111	1,940.17	901	607.81	68	305.99
2014	2,717	14,536.22	1,333	4,376.22	45	6,336.36	125	2,552.76	1,127	795.21	84	475.68
2015	3,963	23,482.59	2,120	6,818.31	47	10,414.89	147	4,353.48	1,563	1,426.32	85	469.59
2016	4,737	23,005.13	2,878	6,271.34	49	10,892.07	171	4,224.52	1,564	1,209.71	75	407.50
2017	6,039	27,266.72	3,584	6,755.59	50	13,090.54	179	5,063.28	2,147	1,908.24	78	449.07
2018	5,875	25,372.34	3,826	6,292.76	61	12,716.86	161	4,360.45	1,757	1,552.92	69	449.35
2019	5,566	20,265.07	4,055	5,375.17	80	9,644.24	148	3,958.79	1,225	943.04	57	343.83

Unit: Billions of RMB. Any institutional investor with its share holdings greater than zero in any quarter of a particular year is account into the number of institutional investors in that year. The total market value is calculated by adding market values of shares held by all institutional investors. The last row of year 2019 reports statistics for only three quarters.

Table 1.5: Average Market Value of share holdings per stock

YEAR	ALL	FUNDS	GOV.	INSUR.	OTHER	QFII
2008	1,672.93	2,972.79	550.93	1,303.08	239.59	773.04
2009	2,327.90	2,925.53	20,993.75	1,099.47	360.43	930.78
2010	2,428.83	2,226.37	47,415.56	5,003.37	272.42	1,113.65
2011	1,836.67	1,662.44	42,180.98	3,846.11	230.18	1,132.34
2012	1,623.91	1,286.93	29,643.71	5,100.77	191.58	1,186.84
2013	1,439.88	1,075.55	29,203.74	4,369.75	168.56	1,112.67
2014	1,337.53	820.28	35,007.49	5,085.19	176.28	1,403.18
2015	1,481.08	804.05	54,815.23	7,403.87	228.03	1,373.06
2016	1,213.93	544.77	54,734.01	6,176.19	193.37	1,358.33
2017	1,128.64	471.20	64,485.40	7,051.93	222.17	1,439.32
2018	1,079.58	411.16	51,905.57	6,770.89	220.87	1,622.20
2019	1,213.48	441.82	40,184.35	8,876.20	256.47	2,010.68

Unit: Millions of RMB. The average market value is calculated as dividing total market value by number of stocks held by different types of institutional investors.

more diversified, as the number of stocks held increases while the average market value of share holding per stock increases only mildly or even decreases.

1.3 Test for Herding Behaviors

In spirit of Sias (2004), the institutional herding behavior is defined as institutional investors following each other by increasing (or decreasing) their holding of the same securities over some period of time. Therefore, if institutional investors follow its own or other institutional investors' previous trades, the correlation between buying (selling) behaviors in the current and previous period will be positive.

This section will first define buying and selling behaviors of institutional investors, which will be used as key variables in the following regressions for evaluating herding behaviors. Then, relation between buying (selling) of a stock k on quarter t by particular institution type i with the aggregate institutional lagged buying (selling) will be estimated. Sections 1.3.2.1 and 1.3.2.2 further test relations between a particular institution type's buying (selling) with its own and other types' lagged buying (selling), respectively. Further tests will be conducted by decomposing stocks into different market caps and industry sectors. In the last part of this section, momentum trading will be evaluated.

1.3.1 Definition of Variables

For each stock k during each quarter t, an institutional investor is defined as a buyer/seller/holder if its holding of the stock increases/decreases/keeps unchanged.

The buying/selling fraction for stock k during quarter t ($B_{k,t}/S_{k,t}$) is calculated from the fraction of number of institutional investors that are buyers/sellers over the total number of institutional investors trading the stock k:

$$B_{k,t} = \frac{No.\ of\ institutional\ buyers_{k,t}}{No.\ of\ total\ institutional\ investors_{k,t}} \tag{1.1}$$

$$S_{k,t} = \frac{No.\ of\ institutional\ sellers_{k,t}}{No.\ of\ total\ institutional\ investors_{k,t}} \tag{1.2}$$

where $No.\ of\ total\ institutional\ investors_{k,t} = No.\ of\ buyers_{k,t} + No.\ of\ sellers_{k,t} + No.\ of\ holders_{k,t}$.

In Sias (2004), only buying fraction is calculated, and its denominator only consists of institutional buyers and sellers. In other word, he only includes institutional investors that change their positions in particular security when evaluating the buying fraction. However, there are institutions maintaining their positions unchanged during quarters in my data set. Without considering these institutions as holders, the fraction calculated will be overvalued. Thus, in this chapter, I add number of holders into the denominator for calculating buying fraction, and extend the test for herding in selling behaviors by calculating selling fraction ($S_{k,t}$).

To allow for direct comparison of estimated coefficients across different institutional investor types, capitalizations and industry sectors, I standardize the buying and selling fraction of stock k in quarter t as:

$$\triangle^{B}_{k,t} = \frac{B_{k,t} - \overline{B_t}}{\sigma(B_{k,t})} \tag{1.3}$$

$$\triangle_{k,t}^{S} = \frac{S_{k,t} - \overline{S_t}}{\sigma(S_{k,t})} \tag{1.4}$$

where $\overline{B_t}$ ($\overline{S_t}$) is the cross-sectional average buying (selling) fraction in quarter t (acros K stocks). $\sigma(B_{k,t})$ and $\sigma(S_{k,t})$ are the cross-sectional standard deviations (across K stocks) of the buying and selling fractions in quarter t.

1.3.2 Regression Results

I begin the analysis to answer my first research question:

Q1: Does China's stock market exhibit herding behavior among institutional investors?

In previous literature, time-series average coefficients from cross-sectional regressions are evaluated and t-statistics based on Fama-MacBeth standard errors are reported, given the presence of time effect in the data sets. Peterson (2009) compares different approaches to estimate standard errors in finance panel data sets, and finds that with only the presence of time effect, both Fama-MacBeth and clustered standard error estimates are unbiased, if there are sufficient number of clusters (e.g., quarters). However, it is unclear whether there presents fixed firm effect, and the precise form of dependence is also unknown, thus a model clustering by two dimensions (firm and time) produces less biased standard errors.

Therefore, a panel regression with time effect and standard error clustering by two dimensions is applied to evaluate the relation between the standardized buying(selling) fraction of stock k in quarter t by all institutional investors and its lag-standardized buying(selling) fraction of stock k in previous quarter $t-1$. The regression equations for buying and selling fractions are as followed:

Buy Side (with superscript B):

$$\triangle_{k,t}^{B,ALL} = \beta_t^B \triangle_{k,t-1}^{B,ALL} + \beta_t^{B,eps} EPS_{k,t} + \beta_t^{B,pe} PE_{k,t} + \beta_t^{B,roe} ROE_{k,t} + \beta_t^{B,a/l} A/L_{k,t} + Time + \epsilon_{k,t}^B \tag{1.5}$$

Sell Side (with superscript S):

$$\triangle_{k,t}^{S,ALL} = \beta_t^S \triangle_{k,t-1}^{S,ALL} + \beta_t^{S,eps} EPS_{k,t} + \beta_t^{S,pe} PE_{k,t} + \beta_t^{S,roe} ROE_{k,t} + \beta_t^{S,a/l} A/L_{k,t} + Time + \epsilon_{k,t}^S \tag{1.6}$$

where EPS (earning per share), PE (price to earning ratio), ROE (return on earning ratio) and A/L (asset to liability ratio) are control variables observed from Choice Database at the end of each quarter t, and standardized by the same definition in equation (1.3) and (1.4). β_t^B and β_t^S are coefficients of interest, which demonstrate herding in buying and selling if positive.

Table[1.6], [1.7] and [1.8], [1.9] report regression results from buy side and sell side, respectively. The top coefficient in the fist column of panel A in Table[1.6] equals -0.1724 and is statistically significant, which implies that an increase in the previous quarter buy fraction is followed by a decrease in buy fraction in current quarter. The sell side also exhibits same negative relation (-0.2598). Following Sias (2004) definition of herding, the estimation results do not demonstrate herding behavior by the aggregate group of institutional investors in Shanghai Stock Exchange. Results are consistent when applying to stocks that have more institutional investors involved in trading (see:

Panel B and C).

Given results in Section 1.2.2, number of funds and stocks they trade outweigh other institution types in Shanghai stock market, thus coefficients estimated under aggregate level must also be dominated by funds trading behavior. Columns (2) to (7) report results of regressions of lagged buy fraction by all institutional investors on each particular institution type's buying fraction. All coefficients, except the ones for funds, are statistically and significantly positive, implying that institution types other than funds do exhibit herding behavior in buying stocks. Among others, the SH-HK foreign investors show the strongest herding in buy behavior (with coefficient equals to 0.3686), while 'other' type investors showing the weakest relation with lagged aggregate buy.

Results are similar to the sell side: coefficients on funds are all negative and statistically significant; SH-HK exhibits the highest positive relation with lagged aggregate sell. Nonetheless, the remaining four types' results are different. All the coefficients are smaller in magnitude when comparing with the buy side, and they are less statistically significant when applying to all stocks in the data set. Moreover, 'other' type's coefficients are negative, indicating that there is no evidence of herding in sell among 'other' type of institutional investors.

1.3.2.1 Herding Within Own Type

Due to investment restrictions by regulations and purposes for chasing performance, different types of institutions always carry out diverse investment strategies.

Table 1.6: Tests for Herding (Buy Side) on All Institutional Investors

Dependent Variable	(1) ALL	(2) FUND	(3) QFII	(4) GOV.	(5) INSUR.	(6) OTHER	(7) SH-HK
Panel A: Securities with >= 1 Institutional Investors							
Lag Buy by ALL	-0.1724 ***	-0.2738 ***	0.0985 ***	0.1759 ***	0.1129 ***	0.0543 ***	0.3686 ***
	(0.0261)	(0.0191)	(0.0222)	(0.0244)	(0.0169)	(0.0150)	(0.0433)
EPS	0.0563 **	-0.0002	-0.0191 **	0.0083	-0.0007	0.0172	0.1021 ***
	(0.0228)	(0.0105)	(0.0094)	(0.0164)	(0.0083)	(0.0123)	(0.0361)
P/E ratio	-0.0041	-0.0016	-0.0051	-0.0168 ***	0.0145 *	-0.0007	-0.0196 **
	(0.0065)	(0.0135)	(0.0281)	(0.0063)	(0.0086)	(0.0046)	(0.0089)
ROE	0.0265 **	0.0052	0.0511	0.0954 ***	0.0981 **	0.0206 ***	0.1055 **
	(0.0134)	(0.0135)	(0.0312)	(0.0367)	(0.0427)	(0.0078)	(0.0343)
A/L Ratio	-0.0603 **	-0.0802 ***	-0.0238	-0.0187	-0.0892 ***	-0.0684 ***	0.0708 *
	(0.0159)	(0.0216)	(0.0405)	(0.0220)	(0.0266)	(0.0168)	(0.0365)
Observations	38,543	34,540	4,802	13,789	9,904	24,411	7,437
Time Effect	YES	YES	YES	YES	YES	YES	YES
Std. Err. Clustered	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH
R-Squared	0.0458	0.0744	0.0079	0.0220	0.0120	0.0069	0.1221
Panel B: Securities with >= 5 Institutional Investors							
Lag Buy by ALL	-0.1896 ***	-0.2860 ***	0.0758 ***	0.1861 ***	0.1100 ***	0.0181	0.3753 ***
	(0.0351)	(0.0200)	(0.0267)	(0.0247)	(0.0189)	(0.0165)	(0.0449)
EPS	0.0218	-0.0113	-0.0203 **	0.0016	-0.0061	0.0046	0.0998 ***
	(0.0138)	(0.0097)	(0.0098)	(0.0142)	(0.0086)	(0.0076)	(0.0351)
P/E ratio	-0.0032	-0.0005	0.0084	-0.0152 **	0.0160	0.0002	-0.0196 **
	(0.0069)	(0.0065)	(0.0228)	(0.0069)	(0.0117)	(0.0049)	(0.0332)
ROE	0.0212	0.0125	0.0594	0.1225 ***	0.1154 ***	0.0229 *	0.1038 ***
	(0.0179)	(0.0177)	(0.0403)	(0.0405)	(0.0431)	(0.0135)	(0.0332)
A/L Ratio	-0.0476 ***	-0.0718 ***	0.0012	-0.0170	-0.0963 ***	-0.0517 ***	0.0706 *
	(0.0160)	(0.0234)	(0.0464)	(0.0233)	(0.0275)	(0.0155)	(0.0368)
Observations	29,202	28,467	4,137	12,851	8,984	19,420	7,373
Time Effect	YES	YES	YES	YES	YES	YES	YES
Std. Err. Clustered	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH
R-Squared	0.0515	0.0944	0.0048	0.0230	0.0111	0.0022	0.1246

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

Table 1.7: Tests for Herding (Buy Side) on All Institutional Investors - Con't

Dependent Variable	(1) ALL		(2) FUND		(3) QFII		(4) GOV.		(5) INSUR.		(6) OTHER		(7) SH-HK	
Panel C: Securities with >= 12 Institutional Investors														
Lag Buy by ALL	-0.1350	***	-0.2637	***	0.0727	**	0.2101	***	0.1274	***	0.0104		0.3693	***
	(0.0393)		(0.0236)		(0.0339)		(0.0245)		(0.0223)		(0.0205)		(0.0460)	
EPS	0.0119		-0.0192	*	-0.0197	**	-0.0054		-0.0061		0.0008		0.0937	***
	(0.0139)		(0.0102)		(0.0093)		(0.0127)		(0.0073)		(0.0064)		(0.0333)	
P/E ratio	-0.0128		-0.0109		0.0105		-0.0150	**	0.0416	***	0.0072		-0.0205	**
	(0.0091)		(0.0086)		(0.0232)		(0.0067)		(0.0083)		(0.0048)		(0.0083)	
ROE	0.0692	***	0.0340		0.0735		0.1848	***	0.0926	**	0.0471	***	0.1007	***
	(0.0229)		(0.0236)		(0.0477)		(0.0495)		(0.0407)		(0.0170)		(0.0349)	
A/L Ratio	-0.0446	***	-0.0815	**	0.0104		-0.0173		-0.0961	***	-0.0566	***	0.0525	
	(0.0172)		(0.0276)		(0.0531)		(0.0256)		(0.0307)		(0.0159)		(0.0338)	
Observations	20,658		20,316		3,078		10,563		6,687		14,283		6,948	
Time Effect	YES		YES		YES		YES		YES		YES		YES	
Std. Err. Clustered	BOTH		BOTH		BOTH		BOTH		BOTH		BOTH		BOTH	
R-Squared	0.0343		0.0989		0.0050		0.0291		0.0133		0.0033		0.1174	

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

Table 1.8: Tests for Herding (Sell Side) on All Institutional Investors

Dependent Variable	(1) ALL		(2) FUND		(3) QFII		(4) GOV.		(5) INSUR.		(6) OTHER		(7) SH-HK	
Panel A: Securities with >= 1 institutional investor														
Lag Sell by ALL	-0.2598	***	-0.2882	***	0.0293		0.0705	***	0.0173		-0.0190		0.1984	***
	(0.0216)		(0.0162)		(0.0202)		(0.0208)		(0.0153)		(0.0139)		(0.0249)	
EPS	0.0323	***	0.0274	**	0.0257	***	0.0507	***	0.0197	**	0.0585	**	0.0731	***
	(0.0124)		(0.0121)		(0.0088)		(0.0169)		(0.0065)		(0.0182)		(0.0274)	
P/E ratio	0.0026		(0.0012)		-0.0087		0.0060		-0.0098		-0.0050		-0.0130	
	(0.0064)		(0.0089)		(0.0087)		(0.0083)		(0.0104)		(0.0056)		(0.0094)	
ROE	0.0191		0.0203		0.0459		0.0574	*	0.0417		0.0145	*	0.0724	***
	(0.0130)		(0.0139)		(0.0301)		(0.0301)		(0.0352)		(0.0083)		(0.0203)	
A/L Ratio	-0.0392	**	0.0220		-0.0579		(0.0307)		-0.0145		0.0153		0.0632	**
	(0.0166)		(0.0218)		(0.0402)		(0.0214)		(0.0254)		(0.0183)		(0.0318)	
Observations	38,543		34,540		4,802		13,789		9,904		24,411		7,437	
Time Effect	YES		YES		YES		YES		YES		YES		YES	
Std. Err. Clustered	BOTH		BOTH		BOTH		BOTH		BOTH		BOTH		BOTH	
R-Squared	0.0754		0.0770		0.0069		0.0106		0.0020		0.0071		0.0393	
Panel B: Securities with >= 5 institutional investor														
Lag Sell by ALL	-0.3176	***	-0.3096	***	0.0496	**	0.0745	***	0.0131		-0.0546	***	0.1965	***
	(0.0253)		(0.0198)		(0.0227)		(0.0218)		(0.0188)		(0.0190)		(0.0250)	
EPS	0.0475	***	0.0334	**	0.0205	***	0.0418	***	0.0178	***	0.0500	***	0.0709	***
	(0.0159)		(0.0140)		(0.0076)		(0.0150)		(0.0060)		(0.0172)		(0.0266)	
P/E ratio	-0.0076	*	-0.0037		(0.0054)		0.0078		-0.0121		-0.0061		-0.0129	
	(0.0044)		(0.0077)		(0.0115)		(0.0094)		(0.0125)		(0.0053)		(0.0087)	
ROE	0.0172		0.0144		0.0614	*	0.1008	***	0.0281		0.0213		0.0695	***
	(0.0224)		(0.0182)		(0.0361)		(0.0308)		(0.0358)		(0.0174)		(0.0191)	
A/L Ratio	-0.0002		0.0179		-0.0695		-0.0357		-0.0282		0.0082		0.0602	*
	(0.0182)		(0.0240)		(0.0447)		(0.0234)		(0.0268)		(0.0184)		(0.0315)	
Observations	29,202		28,467		4,137		12,851		8,984		19,420		7,373	
Time Effect	YES		YES		YES		YES		YES		YES		YES	
Std. Err. Clustered	BOTH		BOTH		BOTH		BOTH		BOTH		BOTH		BOTH	
R-Squared	0.1230		0.0996		0.0075		0.0118		0.0017		0.0077		0.0380	

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

Table 1.9: Tests for Herding (Sell Side) on All Institutional Investors - Con't

Dependent	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable	ALL	FUND	QFII	GOV.	INSUR.	OTHER	SH-HK
Panel C: Securities with >= 12 institutional investor							
Lag Sell by ALL	-0.2934 ***	-0.2970 ***	0.0770 ***	0.0691 **	0.0132	-0.0681 ***	0.1750 ***
	(0.0263)	(0.0234)	(0.0282)	(0.0270)	(0.0235)	(0.0257)	(0.0219)
EPS	0.0434 ***	0.0364 **	0.0146 **	0.0391 **	0.0142 ***	0.0397 ***	0.0628 ***
	(0.0164)	(0.0167)	(0.0064)	(0.0155)	(0.0033)	(0.0152)	(0.0241)
P/E ratio	-0.0067	0.0049	-0.0033	0.0071	-0.0293 **	-0.0149 **	-0.0089
	(0.0053)	(0.0107)	(0.0148)	(0.0108)	(0.0139)	(0.0054)	(0.0076)
ROE	0.0523 *	0.0118	0.0562	0.1353 ***	0.0251	0.0437 **	0.0563 ***
	(0.0306)	(0.0239)	(0.0445)	(0.0333)	(0.0354)	(0.0223)	(0.0169)
A/L Ratio	0.0163	0.0244	-0.1019 **	-0.0318	-0.0493	0.0080	0.0494
	(0.0218)	(0.0301)	(0.0487)	(0.0255)	(0.0320)	(0.0215)	(0.0328)
Observations	20,658	20,316	3,078	10,563	6,687	14,283	6,948
Time Effect	YES	YES	YES	YES	YES	YES	YES
Std. Err. Clustered	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH
R-Squared	0.1244	0.1053	0.0097	0.0124	0.0024	0.0082	0.0319

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

Moreover, foreign investors, unfamiliar with local stock market, should also exhibit different strategies in trading from domestic institutional investors. The analysis of estimating relation of particular type buy/sell fraction with its own-type lagged buy/sell fraction provides an insight of differences among different types of institutional investors, and an answer to the second research question: Are herding behaviors different across different types of institutional investors? The regression equations are similar to Eq. (1.5) and (1.6), but the independent variables of the buy/sell fraction are under the same institution type as the dependent variables:

Buy Side (with superscript B):

$$\triangle_{k,t}^{B,i} = \beta_t^B \triangle_{k,t-1}^{B,i} + \beta_t^{B,eps} EPS_{k,t} + \beta_t^{B,pe} PE_{k,t} + \beta_t^{B,roe} ROE_{k,t} + \beta_t^{B,a/l} A/L_{k,t} + Time + \epsilon_{k,t}^B \tag{1.7}$$

Sell Side (with superscript S):

$$\triangle_{k,t}^{S,i} = \beta_t^S \triangle_{k,t-1}^{S,i} + \beta_t^{S,eps} EPS_{k,t} + \beta_t^{S,pe} PE_{k,t} + \beta_t^{S,roe} ROE_{k,t} + \beta_t^{S,a/l} A/L_{k,t} + Time + \epsilon_{k,t}^S \tag{1.8}$$

where $i \in [FUNDS, QFII, GOV., INSUR., OTHER, SH-HK]$. Results are reported in Table [1.10], [1.11]and [1.12],[1.13] for buy and sell side respectively.

Consistent with previous findings, coefficients associated with funds are all negative, and they become more negative when estimating stocks with more institutional investors, both from buy and sell side. The results also reveal stronger evidence of herding for the two types of foreign investors (QFII and SH-HK). In sell side, coefficients associated with these two types are larger than those with domestic types. To sum up,

Table 1.10: Tests for Herding (Buy Side) on Own Type

Dependent Variable	(1) FUND	(2) QFII	(3) GOV.	(4) INSUR.	(5) OTHER	(6) SH-HK
Panel A: Securities with >= 1 Institutional Investor						
Lag Buy by Own Type	-0.2538 ***	0.1803 ***	0.2184 ***	0.1782 ***	0.1021 ***	0.5163 ***
	(0.0231)	(0.0209)	(0.0187)	(0.0130)	(0.0125)	(0.0340)
EPS	0.0128	0.0124	0.0253	0.0174 *	0.0443 **	0.0779 **
	(0.0134)	(0.0094)	(0.0204)	(0.0100)	(0.0173)	(0.0355)
P/E ratio	-0.0032	-0.0802 ***	-0.0077	0.0161 *	-0.0015	-0.0135
	(0.0079)	(0.0259)	(0.0056)	(0.0087)	(0.0053)	(0.0103)
ROE	0.0236	0.0974 **	0.1087 ***	0.1267 ***	0.0182 **	0.1178 ***
	(0.0170)	(0.0426)	(0.0307)	(0.0328)	(0.0074)	(0.0409)
A/L Ratio	-0.0602 **	-0.0534	0.0009	-0.0450 **	-0.0212	0.0244
	(0.0244)	(0.0365)	(0.0197)	(0.0184)	(0.0138)	(0.0279)
Observations	31,842	3,282	11,563	7,555	19,832	6,599
Time Effect	YES	YES	YES	YES	YES	YES
Std. Err. Clustered	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH
R-Squared	0.0810	0.0508	0.0693	0.0491	0.0238	0.3176
Panel B: Securities with >= 5 Institutional Investors						
Lag Buy by Own Type	-0.2843 ***	0.1790 ***	0.2149 ***	0.1696 ***	0.0676 ***	0.5125 ***
	(0.0181)	(0.0212)	(0.0200)	(0.0140)	(0.0141)	(0.0333)
EPS	-0.0064	0.0066	0.0180	0.0092	0.0278 **	0.0776 **
	(0.0090)	(0.0066)	(0.0179)	(0.0072)	(0.0123)	(0.0353)
P/E ratio	-0.0020	-0.0949 ***	-0.0061	0.0196 *	0.0022	-0.0137
	(0.0067)	(0.0268)	(0.0062)	(0.0113)	(0.0057)	(0.0102)
ROE	0.0173	0.1156 ***	0.1218 ***	0.1429 ***	0.0344 **	0.1191 ***
	(0.0200)	(0.0427)	(0.0357)	(0.0276)	(0.0137)	(0.0412)
A/L Ratio	-0.0618 **	-0.0416	0.0005	-0.0544 ***	-0.0156	0.0260
	(0.0274)	(0.0403)	(0.0213)	(0.0196)	(0.0150)	(0.0292)
Observations	26,576	2,939	10,942	6,915	15,958	6,539
Time Effect	YES	YES	YES	YES	YES	YES
Std. Err. Clustered	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH
R-Squared	0.1100	0.0464	0.0655	0.0443	0.0118	0.3160

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

Table 1.11: Tests for Herding (Buy Side) on Own Type - Con't

Dependent Variable	(1) FUND		(2) QFII		(3) GOV.		(4) INSUR.		(5) OTHER		(6) SH-HK	
Panel C: Securities with >= 12 Institutional Investors												
Lag Buy by Own Type	-0.2872	***	0.1725	***	0.2227	***	0.1822	***	0.0496	***	0.4804	***
	(0.0204)		(0.0229)		(0.0216)		(0.0139)		(0.0165)		(0.0372)	
EPS	-0.0192	**	0.0031		0.0068		0.0042		0.0205	*	0.0758	**
	(0.0088)		(0.0052)		(0.0142)		(0.0047)		(0.0111)		(0.0348)	
P/E ratio	-0.0114		-0.0668	**	-0.0090		0.0346	***	0.0103		-0.0169	*
	(0.0083)		(0.0316)		(0.0059)		(0.0087)		(0.0065)		(0.0102)	
ROE	0.0345		0.1401	***	0.1711	***	0.1264	***	0.0630	***	0.1181	***
	(0.0239)		(0.0436)		(0.0433)		(0.0352)		(0.0160)		(0.0438)	
A/L Ratio	-0.0787	**	-0.0401		-0.0056		-0.0537	**	-0.0208		0.0173	
	(0.0321)		(0.0452)		(0.0240)		(0.0234)		(0.0148)		(0.0273)	
Observations	19,254		2,225		9,321		5,241		11,899		6,122	
Time Effect	YES		YES		YES		YES		YES		YES	
Std. Err. Clustered	BOTH		BOTH		BOTH		BOTH		BOTH		BOTH	
R-Squared	0.1304		0.0429		0.0712		0.0487		0.0102		0.2823	

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

Table 1.12: Tests for Herding (Sell Side) on Own Type

Dependent Variable	(1) FUND		(2) QFII		(3) GOV.		(4) INSUR.		(5) OTHER		(6) SH-HK	
Panel A: Securities with >= 1 Institutional Investor												
Lag Sell by Own Type	-0.3226	***	0.2218	***	0.1503	***	0.1136	***	0.0409	**	0.3313	***
	(0.0229)		(0.0231)		(0.0257)		(0.0179)		(0.0165)		(0.0291)	
EPS	0.0072		0.0041		0.0399	**	0.0069		0.0412	**	0.0660	**
	(0.0081)		(0.0063)		(0.0151)		(0.0077)		(0.0172)		(0.0264)	
P/E ratio	0.0015		-0.0256		0.0016		-0.0092		-0.0050		-0.0113	*
	(0.0090)		(0.0375)		(0.0097)		(0.0121)		(0.0066)		(0.0075)	
ROE	0.0093		0.0425		0.0585	*	0.0576		0.0194	***	0.0781	***
	(0.0150)		(0.0367)		(0.0327)		(0.0383)		(0.0072)		(0.0148)	
A/L Ratio	0.0018		-0.0574		-0.0400	*	-0.0435		(0.0255)		0.0304	
	(0.0232)		(0.0528)		(0.0237)		(0.0287)		(0.0171)		(0.0276)	
Observations	31,842		3,282		11,563		7,555		19,832		6,599	
Time Effect	YES		YES		YES		YES		YES		YES	
Std. Err. Clustered	BOTH		BOTH		BOTH		BOTH		BOTH		BOTH	
R-Squared	0.1147		0.0421		0.0267		0.0148		0.0090		0.1328	
Panel B: Securities with >= 5 Institutional Investors												
Lag Sell by Own Type	-0.3462	***	0.2294	***	0.1481	***	0.0996	***	0.0159		0.3233	***
	(0.0191)		(0.0254)		(0.0255)		(0.0184)		(0.0182)		(0.0284)	
EPS	0.0193	**	0.0003		0.0341	**	0.0062		0.0323	**	0.0648	**
	(0.0089)		(0.0063)		(0.0137)		(0.0074)		(0.0152)		(0.0260)	
P/E ratio	-0.0008		0.0380		0.0034		-0.0128		-0.0075		-0.0115	*
	(0.0081)		(0.0537)		(0.0111)		(0.0140)		(0.0050)		(0.0074)	
ROE	0.0123		0.0710		0.0944	***	0.0391		0.0266	*	0.0751	***
	(0.0189)		(0.0444)		(0.0316)		(0.0372)		(0.0151)		(0.0136)	
A/L Ratio	0.0035		-0.0521		-0.0448	*	-0.0581	*	-0.0249		0.0280	
	(0.0257)		(0.0579)		(0.0255)		(0.0307)		(0.0179)		(0.0275)	
Observations	26,576		2,939		10,942		6,915		15,958		6,539	
Time Effect	YES		YES		YES		YES		YES		YES	
Std. Err. Clustered	BOTH		BOTH		BOTH		BOTH		BOTH		BOTH	
R-Squared	0.1468		0.0459		0.0277		0.0130		0.0089		0.1263	

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

Table 1.13: Tests for Herding (Sell Side) on Own Type - Con't

Dependent Variable	(1) FUND		(2) QFII		(3) GOV.		(4) INSUR.		(5) OTHER		(6) SH-HK	
Panel C: Securities with >= 12 Institutional Investors												
Lag Sell by Own Type	-0.3590	***	0.2340	***	0.1505	***	0.1217	***	0.0052		0.2759	***
	(0.0222)		(0.0303)		(0.0236)		(0.0215)		(0.0209)		(0.0196)	
EPS	0.0274	**	-0.0026		0.0345	**	0.0053		0.0224	*	0.0609	**
	(0.0118)		(0.0067)		(0.0151)		(0.0049)		(0.0124)		(0.0249)	
P/E ratio	0.0064		0.0592		0.0034		-0.0195		-0.0176	***	-0.0077	
	(0.0110)		(0.0719)		(0.0111)		(0.0140)		(0.0038)		(0.0075)	
ROE	0.0060		0.0558		0.1258	***	0.0190		0.0553	***	0.0634	***
	(0.0208)		(0.0476)		(0.0359)		(0.0413)		(0.0178)		(0.0137)	
A/L Ratio	0.0137		-0.0792		-0.0391		-0.0811	**	-0.0266		0.0228	
	(0.0326)		(0.0643)		(0.0272)		(0.0354)		(0.0197)		(0.0299)	
Observations	19,254		2,225		9,321		5,241		11,899		6,122	
Time Effect	YES		YES		YES		YES		YES		YES	
Std. Err. Clustered	BOTH		BOTH		BOTH		BOTH		BOTH		BOTH	
R-Squared	0.1784		0.0521		0.0306		0.0198		0.0114		0.0973	

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

the analysis provides two key findings: funds herding behaviors are completely different from others; foreign investors trading strategies show stronger herding than domestic players. I will show further analysis on funds to provide evidence on funds' difference in herding.

Table 1.14: Tests for Herding (Buy Side) on One Other Type

	(1) FUND		(2) QFII		(3) GOV.		(4) INSUR.		(5) OTHER		(6) SH-HK	
Lag FUND	-0.2538	***	-0.0120		0.0395		-0.0909	***	-0.0136		0.0002	
	(0.0231)	/31,842	(0.0429)	/1,140	(0.0267)	/4,778	(0.0262)	/2,552	(0.0182)	/5,898	(0.0282)	/2,471
Lag QFII	-0.0149		0.1803	***	-0.0271		0.0410		0.0100		-0.0345	
	(0.0095)	/1,466	(0.0209)	/3,282	(0.0358)	/796	(0.0405)	/491	(0.0188)	/1,121	(0.0369)	/387
Lag GOV.	0.0287	***	0.0044		0.2184	***	0.0216		0.0166	*	0.0486	***
	(0.0096)	/5,066	(0.0349)	/675	(0.0187)	/11,563	(0.0233)	/1,555	(0.0100)	/3,714	(0.0176)	/2,196
Lag INSUR.	-0.0015		-0.0229		0.0711	***	0.1782	***	-0.0228	*	-0.0017	
	(0.0124)	/3,144	(0.0423)	/474	(0.0190)	/1,771	(0.0130)	/7,555	(0.0119)	/2,472	(0.0335)	/755
Lag OTHER	0.0503	***	0.0568	*	-0.0207		0.0019		0.1021	***	0.0384	
	(0.0096)	/6,576	(0.0328)	/974	(0.0200)	/3,897	(0.0229)	/2,205	(0.0125)	/19,832	(0.0242)	/1,953
Lag SH-HK	0.0575	***	0.0538		0.0397		0.0228		0.0534	**	0.5163	***
	(0.0218)	/2,455	(0.0799)	/316	(0.0342)	/2,163	(0.0471)	/610	(0.0246)	/1,719	(0.0340)	/6,599

Each column represents regression on each particular type of buy fraction at quarter t. Each row reports the coefficient on each type's lag buy fraction of particular regression, with standard error shown in parenthesis, and numbers of observations reported after left slashes.

,, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

Table 1.15: Tests for Herding (Sell Side) on One Other Type

	(1)		(2)		(3)		(4)		(5)		(6)	
	FUND		QFII		GOV.		INSUR.		OTHER		SH-HK	
Lag FUND	-0.3226	***	-0.0115		0.0417		-0.0786	***	-0.0027		0.0093	
	(0.0229)	/31,842	(0.0677)	/1,140	(0.0318)	/4,778	(0.0301)	/2,552	(0.0262)	/5,898	(0.0237)	/2,471
Lag QFII	0.0010		0.2218	***	0.0601	*	0.0424		0.0314	*	0.0153	
	(0.0090)	/1,466	(0.0231)	/3,282	(0.0347)	/796	(0.0462)	/491	(0.0178)	/1,121	(0.0249)	/387
Lag GOV.	0.0315	**	0.0886	**	0.1503	***	0.0315		0.0330	**	0.0241	
	(0.0128)	/5,066	(0.0436)	/675	(0.0257)	/11,563	(0.0284)	/1,551	(0.0145)	/3714	(0.0177)	/2,196
Lag INSUR.	-0.0163	**	0.0049		0.0126		0.1136	***	-0.0062		-0.0353	***
	(0.0081)	/3,144	(0.0547)	/474	(0.0234)	/1,771	(0.0179)	/7,555	(0.0121)	/2,472	(0.0126)	/755
Lag OTHER	0.0803	***	0.0051		0.0146		0.0006		0.0409	**	0.0599	***
	(0.0115)	/6,576	(0.0376)	/974	(0.0171)	/3,897	(0.0262)	/2,205	(0.0165)	/19,832	(0.0175)	/1,953
Lag SH-HK	0.0518	*	0.0463		0.0273		0.0087		0.0238		0.3313	***
	(0.0294)	/2,455	(0.0621)	/316	(0.0227)	/2,163	(0.0552)	/610	(0.0394)	/1,719	(0.0291)	/6,599

Each column represents regression on each particular type of sell fraction at quarter t. Each row reports the coefficient on each type's lag sell fraction of particular regression, with standard error shown in parenthesis, and numbers of observations reported after left slashes.

,, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

1.3.2.2 Herding Across Different Types

Studies have found that institutional investors follow themselves as well as other types into and out of the same stocks in other countries. In this subsection, other types' lagged trading behavior will be added to estimate institutional herding that following other types' previous trades. I begin by adding one other type a time to the regression equations(1.7) and (1.8). Table[1.14] and [1.15] report coefficients on different types' lagged buy/sell fraction in each particular regression listed in columns. Numbers in parenthesis and after left slashes are standard errors and numbers of observations, respectively.

From the buy side, funds show positive relations with lagged buy fraction of SH-HK, government funds, and 'other' institutions, but the magnitudes are smaller (about one fifth the size of coefficient on its own type (-0.2538)). Both funds and 'other' show evidence of herding on following SH-HK's previous buy, where 'other' exhibits stronger in herding with the coefficient (0.0534) to be half of the coefficient on its own type (0.1021). However, foreign investors do not show significant herding on other type's previous buy. Results are similar from the sell side for domestic investors, but foreign investors show some evidence on herding: QFII exhibits herding with lagged government funds sell, and SH-HK follows 'other' type previous sell[6].

[6] I report results for regressions that include all different types' lagged buy/sell fractions in the appendix. Even though regressions that include lagged SH-HK buy/sell fraction have much smaller size of observations, given the fact that SH-HK connect program starts in 2017, results from regressions include SH-HK are similar to those exclude SH-HK.

1.3.3 Herding Across Sectors

In the previous analysis, funds buy/sell behaviors are completely different from other types of institutions, the coefficients associate with its own lagged buy/sell fractions are all negative and statistically significant. Over the last decades, Chinese funds have developed to be diverse in investment style (e.g. active or passive, pure equity or commingled fund) and ownership structure (e.g. state-owned or privately own or joint venture). Given this diversification, funds trading decisions could be different across securities. Without the availability of all funds' basic information, I am not able to decompose funds into different categories and estimate herding behaviors across them. Nonetheless, established with different purposes or under different regulations, funds' portfolios might specialize in different sectors, e.g., funds that chasing high yields are specialized in high-tech firms. Thus, estimating herding across sectors is useful for shedding some light on diverse investment decisions across different funds.

Stocks listed on Shanghai Stock Exchange are classified into 18 industries. For comparison, I combine similar industries and merge those with number of stocks smaller than others, which yields in total 8 sectors. Table[1.16] reports number of securities, and industries included of each sector in 2019q3. Number of stocks in Manuf & Const. weight over half of the total number of stocks held by institutional investors in 2019q3, while Agric & Energy accounts for 10.37% , and each of the remaining sectors takes about 5%.

To estimate funds' herding across sectors, I run regressions equations (1.7) and

Table 1.16: Sector Analysis

Sector	# Securities(2019q3)	Industry Included
Agric&Energy	119	Agriculture, Forestry, Animal Husbandry and Fishery; Water Conservancy, Environment and Public Facilities Management; Electricity, Heat, Gas and Water Production and Supply Industry; Mining Industry;
Finance	65	Finance Industry;
Manuf&Const.	656	Manufacturing Industry; Construction Industry;
RealEstate	58	Real Estate Industry;
Technology	71	Information Transmission, Software and Information Technology Services; Scientific research and technology services
Transport&Com	58	Transportation, Warehousing and Postal Services;
Wholesale&Retail	68	Wholesale and Retail Industry;
OTHER	52	Accommodation and Catering; Health and Social Work; Culture, Sports and Entertainment; Leasing and Business Services; Comprehensive Operation;

(1.8) for funds under each sector respectively. Columns (1)-(8) in Table[1.17] and [1.18] report the regression results. All the coefficients on funds own lag buy/sell fraction are statistically significantly less than 0, implying that funds exhibit 'opposite' herding, i.e., they will follow themselves in trading the same securities in an opposite direction.

Recall from Table[1.10], funds' coefficient on its lag buy over all sectors is -0.2538[7]. It can be shown in Table[1.17] that funds exhibit more negative relation of its current and lag buy in Finance (-0.4806), Agriculture & Energy (-0.3187), and Wholesale & Retail (-0.2912) sectors. When making buy decisions in these sectors, fewer funds would choose to buy given there was an increase in funds holding the securities in last quarter, comparing to other sectors. In Real Estate, Technology and Other sector, the coefficients are less negative, implying that decrease in fraction of funds buying securities in these sectors is smaller than in other sectors.

From the sell side, comparing with the coefficient (-0.3226) on funds own lag sell fraction over all section from Table [1.12], coefficients from Finance (-0.5152), Agriculture & Energy (-0.3470), and Wholesale & Retail (-0.3635) sectors are again more negative, suggesting that fewer funds choose to decrease their share holding in these sectors, given funds selling these securities in last quarter. In Technology sector, funds show more negative relation of its current and lag sell (-0.3380) than the average level, which is different from the buy side (coefficient is less negative than the average). When making trade decisions in Technology sector, funds exhibit more 'opposite' herding in selling than in buying the securities.

[7] All the variables in the regression equations are standardized, thus I am able to compare coefficients across different regressions.

Table 1.17: Tests for Herding Across Sections – Buy Side

Dependent Variable	(1) Agric&Energy	(2) Finance	(3) Manuf&Const	(4) Real Estate	(5) Technology	(6) Transport&Com	(7) Wholesale&Retail	(8) Other
Lag Buy	-0.3187 ***	-0.4806 ***	-0.2390 ***	-0.1800 ***	-0.2173 ***	-0.2349 ***	-0.2912 ***	-0.1504 ***
	(0.0392)	(0.0534)	(0.0245)	(0.0583)	(0.0585)	(0.0553)	(0.0424)	(0.0484)
EPS	0.0185	-0.0659 ***	0.0105	0.0782	0.0795	0.0592	0.0425 *	0.0250
	(0.0258)	(0.0252)	(0.0153)	(0.0567)	(0.0646)	(0.0472)	(0.0231)	(0.0362)
P/E ratio	0.0115	-0.0027	-0.0067	0.0218	-0.0639	0.0673	0.0145 **	-0.0059
	(0.0116)	(0.0032)	(0.0109)	(0.0404)	(0.0492)	(0.0424)	(0.0074)	(0.0026)
ROE	0.0379	-0.0855	0.0292	0.0432	-0.0287	0.0135	-0.0510 ***	0.1098
	(0.0275)	(0.0723)	(0.0200)	(0.0838)	(0.0781)	(0.0681)	(0.0159)	(0.0957)
A/L Ratio	-0.0469 *	-0.0638	-0.0751 **	0.0148	-0.1370	0.0040	-0.0283	0.0549
	(0.0279)	(0.0791)	(0.0367)	(0.0841)	(0.1073)	(0.0398)	(0.0356)	(0.0419)
Observations	3,920	1,729	17,035	2,040	1,191	1,895	2,501	1,350
Time Effect	YES	YES	YES	YES	YES	YES	YES	YES
Std. Err. Clustered	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH
R-Squared	0.1125	0.3003	0.0754	0.0419	0.0784	0.0707	0.1052	0.0414

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

Table 1.18: Tests for Herding Across Sections – Sell Side

Dependent Variable	(1) Agric&Energy	(2) Finance	(3) Manuf&Const	(4) Real Estate	(5) Technology	(6) Transport&Com	(7) Wholesale&Retail	(8) Other
Lag Sell	-0.3470 ***	-0.5152 ***	-0.3066 ***	-0.3129 ***	-0.3380 ***	-0.3029 ***	-0.3635 ***	-0.1953 ***
	(0.0329)	(0.0619)	(0.0244)	(0.0363)	(0.0665)	(0.0451)	(0.0438)	(0.0460)
EPS	0.0284	0.0571 ***	0.0021	-0.0357	-0.0762	-0.0454	-0.0353	0.0071
	(0.0228)	(0.0233)	(0.0075)	(0.0453)	(0.0623)	(0.0396)	(0.0233)	(0.0350)
P/E ratio	-0.0150	-0.0092 ***	0.0056	-0.0297	0.0680	-0.0676	-0.0135	0.0066 **
	(0.0142)	(0.0033)	(0.0145)	(0.0365)	(0.0598)	(0.0434)	(0.0072)	(0.0029)
ROE	-0.0472	0.0670	0.0061	0.0239	0.1488	0.0033	0.0482 ***	0.0195
	(0.0433)	(0.0580)	(0.0178)	(0.0850)	(0.1201)	(0.0494)	(0.0176)	(0.0783)
A/L Ratio	-0.0751	0.1454 **	-0.0264	-0.0062	0.0368	0.0426	0.0075	-0.0438
	(0.0479)	(0.0738)	(0.0342)	(0.0480)	(0.0874)	(0.0313)	(0.0334)	(0.0389)
Observations	3,920	1,729	17,035	2,040	1,191	1,895	2,501	1,350
Time Effect	YES	YES	YES	YES	YES	YES	YES	YES
Std. Err. Clustered	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH
R-Squared	0.1260	0.3202	0.1061	0.1016	0.1279	0.1012	0.1456	0.0557

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

1.4 Source of Herding

1.4.1 Momentum Trading

Momentum trading, suggested by Sias (2004), is a form of characteristic herding, which implies that investors herd to (away from) stocks with high (low) past returns. According to Sias (2004), I evaluate momentum trading by adding lag standardized return as an independent variable in regression equation(1.5) and (1.6) as followed:

Buy Side (with superscript B):

$$\triangle_{k,t}^{B,ALL} = \beta_t^B \triangle_{k,t-1}^{B,ALL} + \beta_t^{B,re} R_{k,t-1} + \beta_t^{B,X} X_{k,t} + Time + \epsilon_{k,t}^B \qquad (1.9)$$

Sell Side (with superscript S):

$$\triangle_{k,t}^{S,ALL} = \beta_t^S \triangle_{k,t-1}^{S,ALL} + \beta_t^{S,re} R_{k,t-1} + \beta_t^{S,X} X_{k,t} + Time + \epsilon_{k,t}^S \qquad (1.10)$$

Where $X_{k,t}$ are control variables in Equation(1.5) and (1.6). From Table[1.19], coefficients on stock lag return is much smaller than coefficients on lag buy/sell, implying that institutional investors do not engage in momentum trading. Furthermore, comparing the coefficients on lag buy/sell in panel A of Table[1.6], [1.8] and [1.19], the number for all institutional investors does not change significantly (from -0.1724 to -0.1788 for buy side, and -0.2598 to -0.2606 for sell side), this suggests that momentum trading is not crucial in explaining institutional herding. In other word, institutional buy/sell relates much stronger to lag buy/sell than to lag returns.

Sias(2004) also argues that there exists weakly positive relation of institutional demand(buy) with future returns, in particular, the fads, reputational herding,

Table 1.19: Test for Momentum Trading

Dependent Variable	(1) ALL	(2) FUND	(3) QFII	(4) GOV.	(5) INSUR.	(6) OTHER	(7) SH-HK
Panel A: Institutional Buy							
Lag Buy by ALL	-0.1788***	-0.2888***	0.0838***	0.1656***	0.0928***	0.0556***	0.3978***
	(0.0260)	(0.0191)	(0.0223)	(0.0259)	(0.0168)	(0.0153)	(0.0367)
Stock Lag Return	0.0007**	0.0014***	0.0014***	0.0014*	0.0025***	-0.0002	-0.0034***
	(0.0004)	(0.0003)	(0.0003)	(0.0008)	(0.0003)	(0.0004)	(0.0011)
EPS	0.0557**	-0.0014	-0.0192**	0.0072	-0.0009	0.0173	0.1057***
	(0.0227)	(0.0102)	(0.0094)	(0.0162)	(0.0082)	(0.0124)	(0.0375)
P/E ratio	-0.0042	-0.0018	-0.0046	-0.0172***	0.0137	-0.0007	-0.0175**
	(0.0065)	(0.0077)	(0.0280)	(0.0063)	(0.0087)	(0.0046)	(0.0080)
ROE	0.0263**	0.0042	0.0490	0.0944***	0.0894**	0.0207***	0.1114***
	(0.0134)	(0.0134)	(0.0305)	(0.0356)	(0.0417)	(0.0078)	(0.0363)
A/L Ratio	-0.0586***	-0.0771***	-0.0218	-0.0177	-0.0868***	-0.0688***	0.0702*
	(0.0161)	(0.0221)	(0.0406)	(0.0218)	(0.0265)	(0.0166)	(0.0369)
Observations	38,543	34,540	4,802	13,789	9,904	24,411	7,437
Time Effect	YES	YES	YES	YES	YES	YES	YES
Std. Err. Clustered	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH
R-Squared	0.0467	0.0770	0.0102	0.0230	0.0156	0.0069	0.1281
Panel B: Institutional Sell							
Lag Sell by ALL	-0.2606***	-0.2961***	0.0282	0.0821***	0.0114	-0.0124	0.1938***
	(0.0223)	(0.0175)	(0.0203)	(0.0216)	(0.0152)	(0.0140)	(0.0278)
Stock Lag Return	-0.0001	-0.0009**	-0.0001	0.0019***	-0.0008**	0.0011**	-0.0007
	(0.0004)	(0.0004)	(0.0005)	(0.0007)	(0.0003)	(0.0004)	(0.0008)
EPS	0.0325***	0.0285**	0.0257***	0.0484***	0.0199***	0.0572***	0.0747***
	(0.0124)	(0.0122)	(0.0088)	(0.0160)	(0.0065)	(0.0177)	(0.0281)
P/E ratio	0.0026	-0.0011	-0.0088	0.0057	-0.0094	-0.0050	-0.0127
	(0.0064)	(0.0089)	(0.0087)	(0.0082)	(0.0105)	(0.0056)	(0.0097)
ROE	0.0192	0.0212	0.0462	0.0548*	0.0457	0.0138*	0.0745***
	(0.0130)	(0.0140)	(0.0300)	(0.0285)	(0.0357)	(0.0081)	(0.0199)
A/L Ratio	-0.0395**	0.0196	-0.0582	-0.0290	-0.0152	0.0180	0.0630**
	(0.0164)	(0.0217)	(0.0403)	(0.0214)	(0.0253)	(0.0182)	(0.0315)
Observations	38,543	34,540	4,802	12,789	9,904	24,411	7,437
Time Effect	YES	YES	YES	YES	YES	YES	YES
Std. Err. Clustered	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH
R-Squared	0.0754	0.0779	0.0069	0.0125	0.0025	0.0082	0.0396

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

and characteristic herding models. Correlation between buy/sell fraction and returns in previous quarter, same quarter, and following quarter are reported in Table[1.20]. Institutional buy is found to be positively correlated with current quarter returns and weakly correlated with prior quarter returns. The correlation between institutional buy and following quarter return is weakly positive, which is consistent with the argument that institutional herding driven by fads, reputational and characteristic herding would exhibit price reversal in subsequent periods[8].

From the sell side, institutional sell fraction in aggregate (first row in panel B) is negatively correlated with same quarter returns, but weakly positively correlated with prior quarter returns, which suggests that institutional investors choose to sell by observing a positive previous return. Following returns are negatively correlated with institutional sell, but in a weakly manner. Looking at different types of institutions, government funds, insurance and SH-HK show positive relation between same quarter return and sell fraction, implying that securities current prices do not depends on their selling decisions (stock prices do not drop with their selling of shares). Relations between institutional sell and following quarter returns are different for the two foreign investors from those domestic players. Furthermore, combining results from Table[1.20] and findings that funds exhibit negative relation between current and lag own buy/sell, it is found that given a previous quarter higher return, lag buy(sell) fraction was higher(lower), and current buy(sell) tends to be lower(higher). Intuitively, funds prefer to realize profits given previous higher return by selling securities, and be

[8]Sias(2004) reports a positive correlation between institutional buy and following quarter returns in the US stock market.

conservative in chasing higher returns by decreasing in buy.

Table 1.20: Correlation between Institutional Buy/Sell and Returns

InstitutionType	Panel A: Institutional Buy						Panel B: Institutional Sell					
	Previous quarter return		Same quarter return		Following quarter return		Previous quarter return		Same quarter return		Following quarter return	
ALL	-0.0002		0.1895	***	0.0025	***	0.0011	***	-0.1557	***	-0.0025	***
FUNDS	-0.0005	*	0.1853	***	0.0029	***	0.0008	***	-0.1734	***	-0.0027	***
GOV	0.0017	***	0.1221	***	0.0028	***	0.0011	**	0.0079		-0.0009	**
INSUR	0.0023	***	0.0405	**	0.0002		-0.0005	**	0.0147		-0.0003	
OTHER	0.0001		0.1220	***	0.0016	***	0.0010	***	-0.0822	***	-0.0016	***
QFII	0.0014	***	0.0622	***	-0.0008		-0.0001		-0.0132		0.0008	
SH-HK	0.0006		0.0371		0.0050	*	-0.0010		0.0712	*	0.0029	**

For each period return, I regress it on institutional buy/sell under different types this quarter respectively. Coefficients are the correlations, as the data are standardized and the regression equation only include one single explanatory variable. ***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

1.4.2 Herding by Firm Size

According to Wermer (1999), information cascades arises when institutional investors ignore their own information, and trade with herd because they infer information from each other's trades. This type of herding is more likely in small-cap stocks. On the other hand, Sias (2004) proposes the investigative herding which is stronger in large-cap stocks, because investors follow correlated signals of information. In this section, I decompose my data set into different sizes of firms, and test which type of herding exists among institutional investors. I sort all securities in my data set from smallest to largest by capitalization and group them into five groups, then estimate herding behavior for each type of institutional investors in each group, using equation(1.5) and (1.6)[9]. Results show that herding behaviors occur more likely in large capitalization securities, thus suggest that both domestic and foreign institutional investors commit to investigative herding.

1.5 Conclusion

In this chapter, I study the descriptive statistics for institutional investment and examine institutional herding in Shanghai Stock Exchange. Four research questions are tested and results conclude the following key findings: There is an increasing trend both in number and participation of institutional investors. Most types of institutional investors exhibit herding, especially when buying. Institutional investors more often follow their own type of lag trades. Foreign institutional investors show some evidence on following domestic players' previous sell. There is no clear evidence of momentum trading among institutional investors, and their herding behavior is stronger in large-cap securities, demonstrating investigative herding. Funds exhibit statistically significant opposite herding behavior. Increase in buy(sell) last quarter is followed by a decrease

[9]Results, using equation(1.7) and (1.8) for herding on own lag type in each capitalization group are attached in Appendix.

Table 1.21: Tests for Herding by Firm Size – Buy Side

Dependent Variable	(1) ALL	(2) FUND	(3) QFII	(4) GOV.	(5) INSUR.	(6) OTHER	(7) SH-HK
Panel A: Firm size <= 20% percentile							
Lag Buy by ALL	-0.0115	-0.1499 ***	0.0702 *	0.1572 ***	0.1717 ***	0.1191 ***	-0.0175
	(0.0205)	(0.0248)	(0.0396)	(0.0324)	(0.0379)	(0.0216)	(0.0596)
Observations	6,416	4,695	590	760	851	3,579	415
R-Squared	0.0279	0.0302	0.0251	0.0535	0.0393	0.0515	0.0972
Panel B: Firm size <= 40% percentile							
Lag Buy by ALL	-0.1026 ***	-0.2154 ***	0.1812 ***	0.1732 ***	0.1822 ***	0.0746 ***	0.1893 ***
	(0.0267)	(0.0258)	(0.0409)	(0.0382)	(0.0282)	(0.0201)	(0.0668)
Observations	7,478	6,377	749	1,485	1,384	4,108	788
R-Squared	0.0206	0.0416	0.0444	0.0394	0.0395	0.0127	0.0954
Panel C: Firm size <= 60% percentile							
Lag Buy by ALL	-0.1650 ***	-0.2356 ***	0.0690 *	0.2493 ***	0.1445 ***	0.0308	0.3504 ***
	(0.0267)	(0.0239)	(0.0374)	(0.0300)	(0.0303)	(0.0212)	(0.0556)
Observations	8,005	7,272	952	2,401	2,005	4,603	1,395
R-Squared	0.0388	0.0515	0.0085	0.0413	0.0135	0.0066	0.1025
Panel D: Firm size <= 80% percentile							
Lag Buy by ALL	-0.1947 ***	-0.2537 ***	0.1573 **	0.2267 ***	0.0689 **	0.0269	0.3714 ***
	(0.0313)	(0.0225)	(0.0631)	(0.0543)	(0.0361)	(0.0200)	(0.0395)
Observations	8,260	7,890	1,022	3,859	2,533	5,407	2,073
R-Squared	0.0480	0.0749	0.0091	0.0244	0.0043	0.0013	0.0798
Panel E: Firm size > 80% percentile							
Lag Buy by ALL	-0.3874 ***	-0.5290 ***	0.0232	0.1128 ***	0.0225	-0.0397	0.2904 ***
	(0.0544)	(0.0684)	(0.0548)	(0.0368)	(0.0399)	(0.0315)	(0.0548)
Observations	8,384	8,306	1,489	5,284	3,131	6,714	2,766
R-Squared	0.1597	0.2194	0.0115	0.0125	0.0093	0.0027	0.0813

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

Table 1.22: Tests for Herding by Firm Size – Sell Side

Dependent Variable	(1) ALL		(2) FUND		(3) QFII		(4) GOV.		(5) INSUR.		(6) OTHER		(7) SH-HK	
Panel A: Firm size <= 20% percentile														
Lag Sell by ALL	-0.0833	***	-0.1820	***	0.0082		0.0453		0.0040		0.0286		0.0635	*
	(0.0244)		(0.0233)		(0.0383)		(0.0427)		(0.0326)		(0.0179)		(0.0378)	
Observations	6,416		4,695		590		760		851		3,579		415	
R-Squared	0.0147		0.0306		0.0040		0.0108		0.0099		0.0060		0.0165	
Panel B: Firm size <= 40% percentile														
Lag Sell by ALL	-0.1563	***	-0.2007	***	-0.0003		0.0109		0.0855	***	-0.0115		0.1870	***
	(0.0244)		(0.0198)		(0.0337)		(0.0271)		(0.0202)		(0.0177)		(0.0470)	
Observations	7,478		6,377		749		1,485		1,384		4,108		788	
R-Squared	0.0285		0.0355		0.0074		0.0119		0.0068		0.0011		0.0281	
Panel C: Firm size <= 60% percentile														
Lag Sell by ALL	-0.2465	***	-0.2380	***	-0.0013		0.0722	**	0.0215		-0.0681	***	0.1309	**
	(0.0274)		(0.0277)		(0.0339)		(0.0368)		(0.0307)		(0.0210)		(0.0639)	
Observations	8,005		7,272		952		2,401		2,005		4,603		1,395	
R-Squared	0.0640		0.0522		0.0048		0.0090		0.0050		0.0046		0.0154	
Panel D: Firm size <= 80% percentile														
Lag Sell by ALL	-0.3018	***	-0.2475	***	0.0625		0.0922	***	-0.0518		-0.0406		0.1624	***
	(0.0248)		(0.0237)		(0.0516)		(0.0344)		(0.0330)		(0.0300)		(0.0442)	
Observations	8,260		7,890		1,022		3,859		2,533		5,407		2,073	
R-Squared	0.1061		0.0680		0.0073		0.0115		0.0014		0.0057		0.0152	
Panel E: Firm size > 80% percentile														
Lag Sell by ALL	-0.4652	***	-0.5331	***	0.0384		0.0725	**	-0.0269		-0.0588		0.0723	*
	(0.0565)		(0.0615)		(0.0615)		(0.0344)		(0.0293)		(0.0462)		(0.0391)	
Observations	8,384		8,306		1,489		5,284		3,131		6,714		2,766	
R-Squared	0.2254		0.2308		0.0115		0.0133		0.0065		0.0104		0.0210	

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

in buy(sell) this quarter. Combining the realization of returns, they are conservative in selling securities to realize profits given previous higher returns and decreasing in buy for chasing higher future returns.

Chapter 2

News sentiment and topic analysis on crude oil futures prices

[1]Crude oil is a commodity, and as such, it tends to have large fluctuations at price than more stable investments such as stocks and bonds. Crude oil prices are influenced by a variety of factors. As Baumeister and Kilian (2014) conclude that the explanatory power of these factors vary over time and that different factors are important at different time horizons. Though the factors to explain the change of oil prices are inconclusive, Brandt and Gao (2019) mention that news information can provide a way to quantify macroeconomic and other events that could affect crude oil prices. In this chapter, we analyze the contents of news articles to study how information about crude oil related news affects crude oil futures price.

News information could affect crude oil prices in different ways. Shiller (2015)

[1]This chapter is coauthored with Yunxiao Zhang and Zijing Zhu.

argues that the news media plays an important role in setting the stage for market moves and provoking them. On one hand, news could convey information on the current significant variables that affect the price, reflecting the current market confidence. On the other hand, news could serve as an update of the changes in these significant variables, reflecting the expectations about future oil supply and demand conditions, which could affect the current crude oil prices.

In this chapter, we consider broad web news from various sources. We analyze these web crude oil related news to uncover how this information affects the crude oil futures price using the intra-day high-frequency data. To do so, we use both supervised and unsupervised machine learning algorithms to study the impact of news sentiment and news topics on crude oil futures price increase. The important assumption in our analysis is that the crude oil futures market is nearly perfectly efficient, which means that the price can adjust quickly to any new information released to the public.

A key contribution of our chapter is to demonstrate finer and more objective classifications of news effects on crude oil price change. First, we use unsupervised machine learning algorithms to group crude oil related news articles into different topics without providing prior knowledge on how each topic links to a particular set of words. Second, we conduct both news sentiment analysis and news topics analysis of each news article using intra-day high- frequency data, these results providing us a new index data indicating crude oil price increase or decrease for future studies based on textual analysis.

In this chapter, first, we conduct news sentiment analysis with logistic regression to see how each word in an article can affect the increase or decrease of the crude oil price from September 2019 to December 2019. We find that among all the news words, 152 words have coefficients smaller than -0.5, which are collected as the most negative words. Also, 159 words have coefficient over 0.5, we defined these words as the most positive words. Based on the news sentiment analysis results, we also construct a novel index indicating the sentiment score of a news article given the most positive (negative) words. Second, we categorize news articles discussing crude oil over the entire year of 2019 into topics using unsupervised machine learning algorithms K-means. The K-means algorithm generates 4 clusters of news topics. We rename and interpret each of them to be “World Crude Oil” topic, “WTI Crude Oil” topic, “Financial Analysis” topic, and “Editorial Opinion” topic. Each of the news articles would be assigned to be one of the topics. Finally, we estimate how the news topic and sentiment score would affect the crude oil future price using logistic regression in 5 minutes window. The results suggest that on average, ”World Crude Oil” news has the highest correlation with a crude oil price increase. Moreover, the more positive news is under the topic ”WTI Crude Oil,” the higher probability that WTI crude oil futures price will increase within 5 minutes.

The remainder of this chapter is organized as follows: Section 2.1 presents the related literature. Section 2.2 provides detailed data information. Section 2.3 shows the news sentiment analysis methodology and how the positive score is calculated for every news article. Section 2.4 details the construction of topic analysis. Section 2.5 presents our main empirical analysis. Section 2.6 concludes.

2.1 Literature Review

This chapter mainly relates to three streams of research: (1) studies on general crude oil price; (2) studies on news effects of crude oil price; (3) studies on news textual analysis on general equity market. Broad literature has studied the explaining factor of crude oil price. Hamilton (2009) suggests that the real price of oil follows a random walk without drift. Rapaport (2013) distinguish between demand and supply driven component of crude oil returns by examining its correlation with the equity market. Baumeister and Kilian (2014) discuss an exhaustive set of oil pricing factors from the literature, conclude that the explanatory power of these factors varies over time and that different factors are important at different time horizons. These findings crucially depend on the underlying model structure and assumptions. This chapter corroborates these effects from the assumption that the crude oil futures market is nearly perfectly efficient, updated news information is a direct way to reflect the market sentiment.

Our chapter is part of a growing body of research using textual analysis to examine how news affects economic and financial variables. Most work in this literature deals with the general equity market and aggregate news about equities and the economy. For example, Garcia (2013) studies the effect of sentiment on asset prices using the New York Times between 1905 and 2005. They find that after controlling for other well-known time-series patterns, the predictability of stock returns using news' content is concentrated in recessions. Soo (2015) develops a measure of sentiment across local housing markets by quantifying the positive and negative tone of housing news in local newschapter articles. Bi and Traum (2020) examines how newschapter reporting affects government bond prices during the U.S. state default of the 1840s. Our chapter is different from the aforementioned studies in that we consider a much broader set of news resources and news categories as inputs, we can capture both micro-level supply and demand factors and macroeconomic related factors, even geopolitical developments.

This chapter also contributes to the literature about news effects on the crude oil prices. Most of the literature work with the effects of regularly scheduled macroeconomic releases on crude oil price. For example, Kilian and Vega (2011) propose a formal test of the identifying assumption that energy prices are predetermined for U.S. macroeconomic aggregates using daily energy prices on daily news from the U.S. macroeconomic data releases. In this aspect, this chapter is different from the existing literature focusing on responses of crude oil prices to scheduled macro news announcement from U.S. Another difference from chapters on the schedule announcement is that our news indices are at high frequency than the scheduled releases with a fixed frequency. A recent chapter by Brandt and Gao (2019) uses sentiment scores for global news from RavenPack global macro package to see how news about macroeconomic fundamentals and geopolitical events affect crude oil markets. Our chapter is different from this chapter since we use machine learning algorithms to group news into different topics, which avoid prior knowledge and generate more objective classifications of news effects on crude oil price change.

2.2 Data

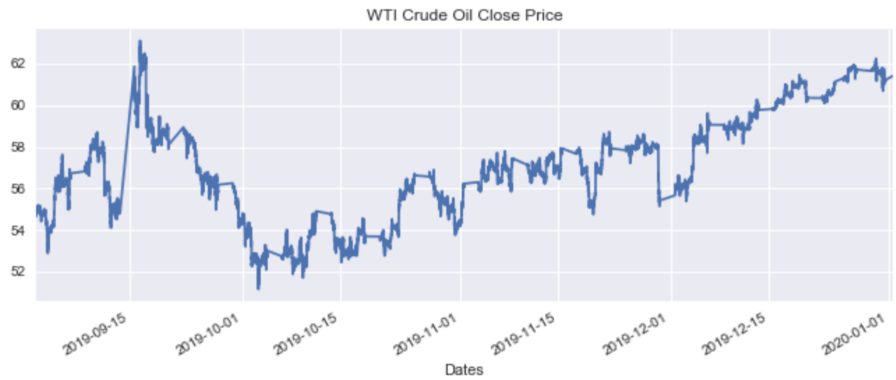
2.2.1 Data Source and Description

The data used in this chapter are obtained through Bloomberg terminal, including the west texas intermediate (WTI) crude oil futures prices, and crude oil-related news articles. This chapter focuses on WTI crude oil. WTI refers to oil extracted from wells in the US and sent via pipeline to Cushing, Oklahoma. This chapter chooses WTI crude oil over Brent crude oil, which counts two-thirds of all crude contracts around the world, and Dubai crude oil, which is the major supply for the Asian market. The reason is that WTI crude oil has been the main benchmark for crude oil consumed in the United States. Thus its price is more related to the supply and demand conditions in the US market than in other markets worldwide.

The intra-day trading data is provided by Bloomberg, with price updating at high frequencies. However, the trading will be suspended every day at 14:00 to 15:00, as well as every Saturday. In order to match the frequency of news, this chapter uses high-frequency intra-day trading futures prices for WTI crude oil. The price is updating every five minutes, which matches the first quantile of news frequency distributions[2]. This chapter takes the sample period to be the last quarter of 2019. Figure 2.1 shows the WTI crude oil close price throughout the period. From Figure 2.1, it is clear that the close price has high volatility but no obvious trend during the last quarter of 2019, which makes it a perfect sample for this chapter's analysis.

[2]Shown in Figure 2.4

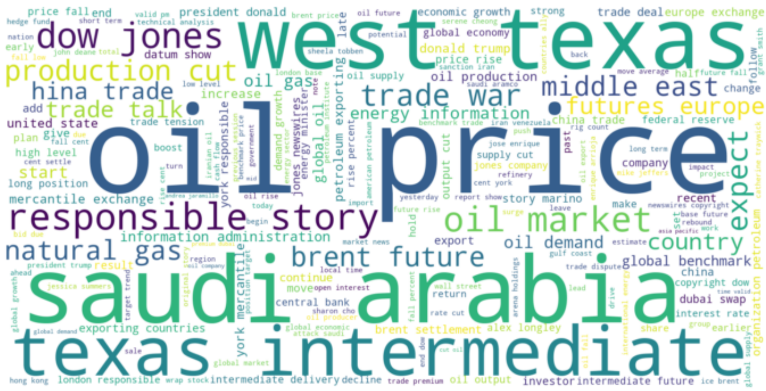
Figure 2.1: Close Price Fluctuation in Sample Period



News articles are the major sources for text analysis. This chapter analyzes contents of each news article to find the news effect on WTI crude oil futures price. The news articles are collected through the Bloomberg terminal's news section. By specifying the news topics to be 'Crude Oil', including the keyword, 'WTI', Bloomberg terminal gives editorial recommended Bloomberg and web news from various sources. Figure 2.2 is a word cloud showing the most frequent words in the news articles collected by this chapter. In 2019, news articles under the topic of WTI crude oil discussed the 'oil price', 'trade war', 'saudi arabia', 'texas intermediate', etc. More detail analysis in news contents will be presented in Section 2.3.

Figure 2.3 shows the major news sources for this chapter's dataset. In total, there are 75 news sources. Figure 2.3 plots the news sources with total news article released counts over ten. All news are global news written in English, reporting crude

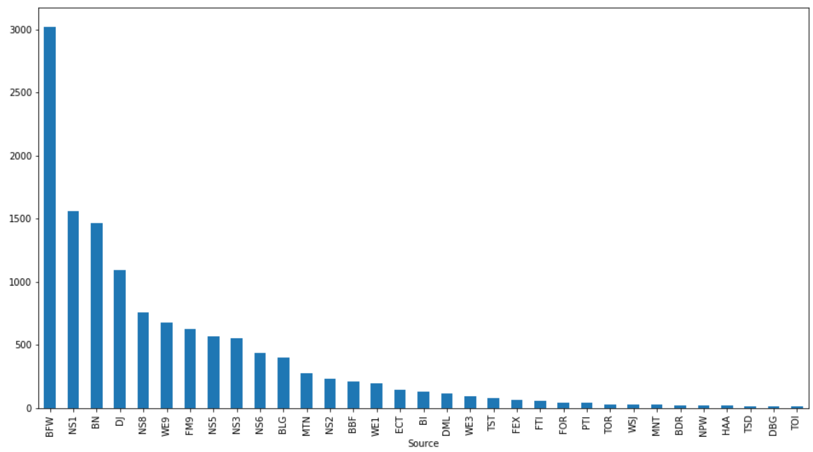
Figure 2.2: 2019 WTI Crude Oil Related News Frequent Words



oil-related topics worldwide. Around 6000 news articles are web news, Bloomberg provides the news subject and the link of the websites. Web news content is collected through web scraping. Two-thirds of the web news is successfully obtained through web-scraping, with the rest replaced by its subject. Throughout 2019, there are 13183 news articles, and 4615 of them are released after September. Since this chapter only has price data from September 2019 to December 2019, only news articles released after September will be included in news sentiment analysis and the final regressions, but all news articles in 2019 will be included in news topic analysis.

As for the frequencies of news articles, Figure 2.4 presents that, on average, around every thirty minutes, there is one news article about WTI crude released. The news frequency distribution gives guidance on what should be price frequency to analyze the news effect. Attempt to uncover the news effect, this chapter assumes the market is efficient, which means the market reacts fast enough to new information. This chapter

Figure 2.3: Major News Source

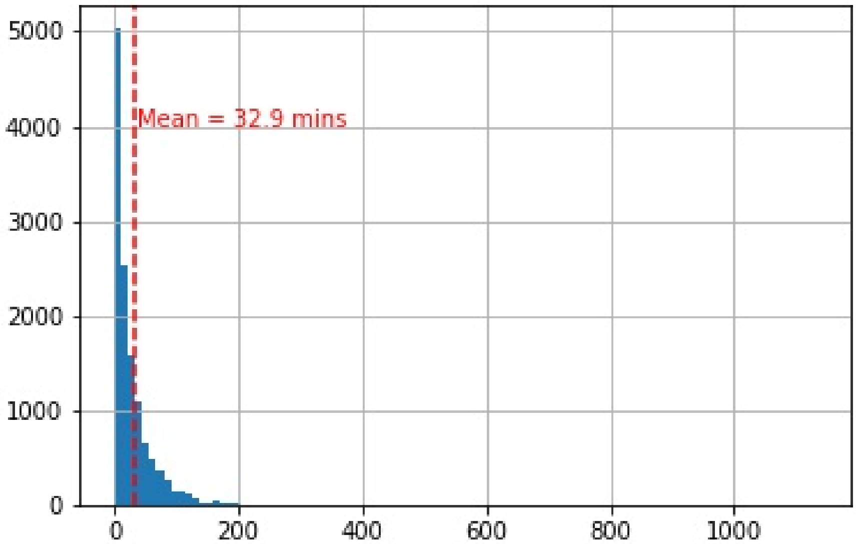


Note: News sources are extracted from each news body, provided by Bloomberg

thus defines any news article's effect on WTI crude oil future price is reflected by how the price was changed within five minutes after this news has been released.

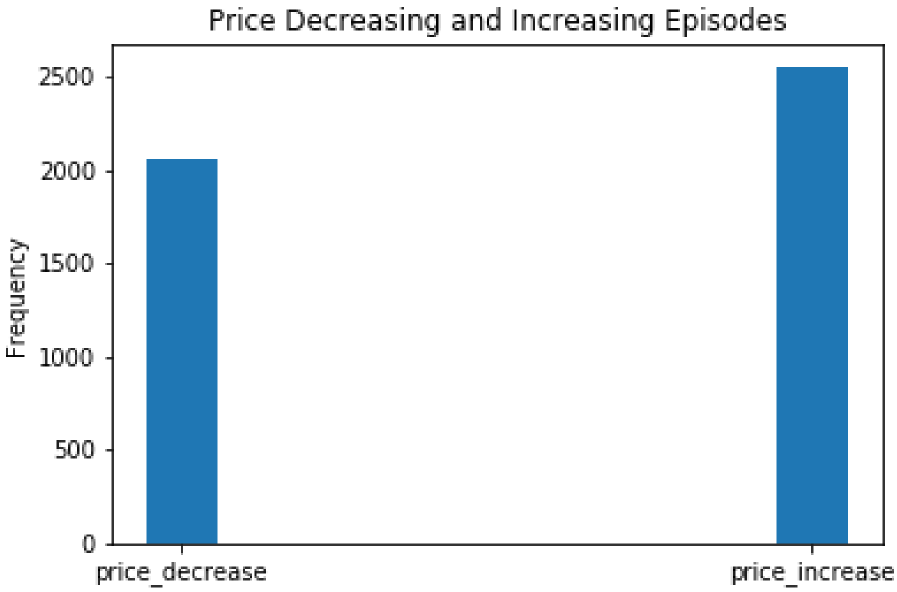
To analyze how the price was changed, this chapter establishes a dummy variable based on WTI crude oil future price, indicating a price increase or decrease episode. If price increases within five minutes after a news article has been released, the price dummy will be one. Otherwise, if price decreases or doesn't change, the price dummy will be zero. There is nearly no case for the price to remain the same within the five-minute slot throughout the dataset. Thus, when price dummy equals zero, it means the price has decreased within five minutes after this news release. For each news article, by looking for the price change within five minutes after release, this chapter matches news

Figure 2.4: News Frequency Distribution



Note: This graph shows the distribution of how frequently a news has been released in the dataset. The x-axis is in the unit of minute, and the y-axis is the count. On average, around every 33 minutes, there release one news article about WTI crude oil in 2019.

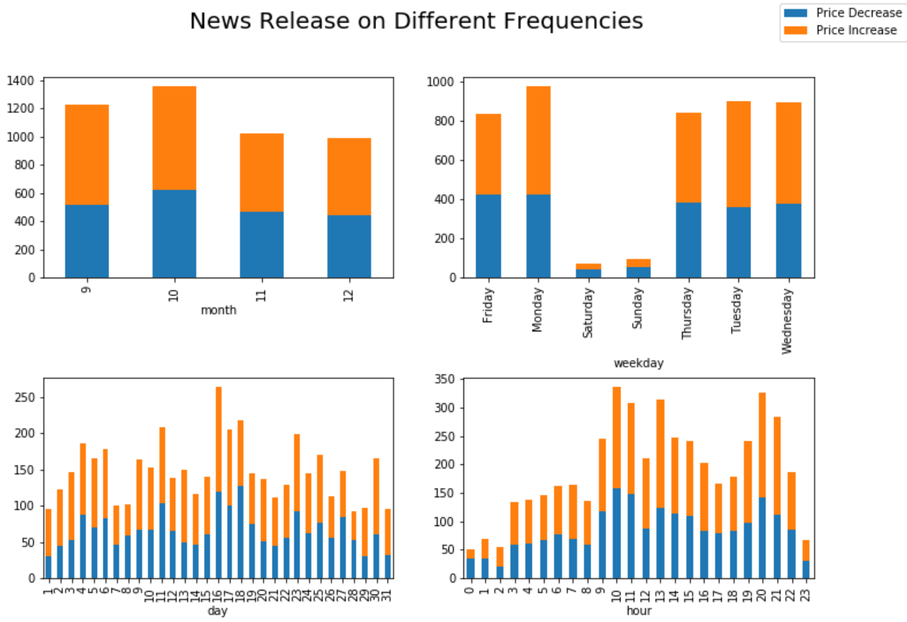
Figure 2.5: Price Dummy Distribution



and the price dummy. Figure 2.5 examines whether the data is balanced by comparing the number of price increase episodes with price decrease episodes. Roughly speaking, the data is balanced, with increasing price episodes slightly larger.

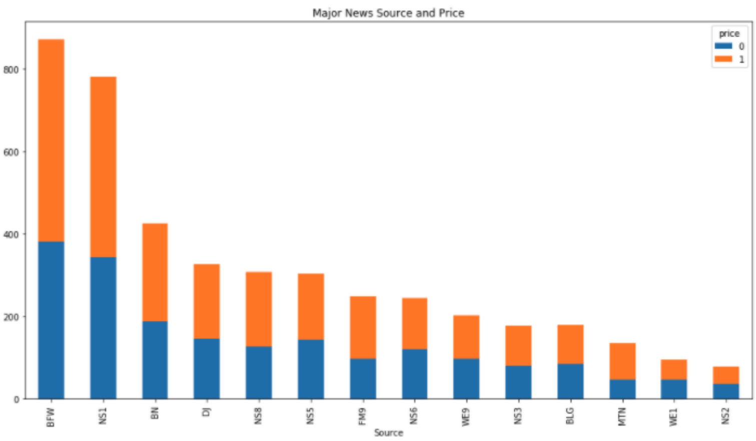
Except examining the balance of episodes across all datasets, this chapter also presents the price sensitivity in different release times of the news articles. In other words, whether news released at certain times, like in the morning or at the end of the month are more likely to have a biased impact on price. Figure 2.6 shows the price decrease and increase episodes distributions across different months, days of the week, date of the month, and hours of the day. The figure indicates that over one thousand

Figure 2.6: Matching News and Price



news on the topic of crude oil news were released every month. Moreover, most of the news are released on weekdays rather than on the weekends, and most of them are released during the daytime rather than night. However, there are no obvious patterns for the date of the news release. Comparing the number of news following price increase and price decrease episodes, Figure 2.6 presents that there are no significant differences in these two episodes at any time frequencies. Similarly, Figure 2.7 attempts to uncover the relationship between price and news sources. By examining the major news sources, it is clear that all these new sources have almost even numbers of the price increase and decrease episodes across the dataset. Thus, it is essential to analyze the news contents, rather than the time of the news release, or the news sources, to understand the positive

Figure 2.7: Matching News and Price



or negative news effect on crude oil futures prices.

2.2.2 Analyze Text Data

In order to analyze the news contents, this chapter preprocesses news before further analysis with regressions. Text data are unique data types that need transformation before fitting into a regression model. This chapter follows the standard text mining procedures to extract the useful features from the news contents, including tokenization, removing stopwords and lemmatization.

The First step of preprocessing text data is to break every sentence into individual words, which is called tokenization. Taking individual words rather than sentences breaks down the connections between words. However, it is a common method to use to analyze large sets of text data. It is efficient and convenient for computers to analyze the text data by examines what words appear in an article and how many times these words appear, and this analysis is sufficient enough to give insightful results.

After tokenization, each news article will transform into a list of words, symbols, digits, and punctuation. The next step is to remove useless information. For this analysis, symbols, digits, and punctuation are not very useful, so that this chapter removes them. Furthermore, this chapter removes stopwords. Stopwords are words that frequently appear in many articles, but without significant meanings. Examples of stopwords are 'I', 'the', 'a', 'of'. These are the words will not intervene in the understanding of articles if removed. Besides using the standard English stopwords provided by the NLTK library[3], this chapter also includes other lists of stopwords, provided by Loughran and McDonald(2016). These lists of stopwords are widely used in economic analysis, including dates and time, more general words that are not economic meaningful[4].

Removing stopwords, along with symbols, digits, and punctuation, each news article will transform into a list of meaningful words. However, in order to count the appearance of each word, it is essential to remove grammar tense and transform each word into its original form. For example, if we want to calculate how many times the word 'open' appears in a news article, we need to count the appearances of 'open', 'opens', 'opened'. Thus, lemmatization is an essential step for text transformation. Lemmatization is taking each word into its original lemma. Another way of converting words is called stemming, which is taking the linguistic root of a word. The reason why this chapter chooses lemmatization over stemming is that after stemming, some words be-

[3]NLTK is a python package for text analysis. It contains a list of English stopwords.

[4]Loughran and McDonald's list can be found at https://sraf.nd.edu/textual-analysis/resources/StopWords.

Figure 2.8: Original News Article

```
'     Total Refinery, Alpiq, Trade Woes: European Energy Pre-Market  2019-08-26 06:12:07.16 GMT      By John Viljoen
(Bloomberg) -- The following may affect European energy  shares today:  * Note, U.K. markets closed due to holiday
News  * Watch These European Stocks as U.S.-China Trade Woes Escalate  * ALPH SW: Alpiq First Half Adjusted Ebitda CH
F55 Mln  * FP FP: Total Reduces Some Gonfreville Refinery Unit Rates Amid  Strike  * LSNG RM: Lenenergo Second Quarte
r Net Income 3.91 Bln Rubles,  +36% Y/y  * ORSTED DC: Ørsted, Eversource Submit Massachusetts Wind Farm  Proposal
Commodities  * WTI Crude: -1% to $53.64/bbl  * Brent Crude: -0.9% to $58.83/bbl  * Natgas: +1.1% to $2.175/Mmbtu    A
genda  * N.A. Energy Weekly Agenda  * Oil daybook Europe  * Earnings:  ** Other  *** Alpiq Holding AG (ALPH SW)  ***
Maha Energy AB (MAHAA SS)    For more energy wraps in Europe, click here.  For more energy sector wraps in the U.S.,
click here.    To contact the reporter on this story:  John Viljoen in Cape Town at jviljoen@bloomberg.net  To contac
t the editor responsible for this story:  Blaise Robinson at brobinson58@bloomberg.net    '
```

Figure 2.9: News Article after Text Preprocessing

```
'total refinery alpiq trade woe european energy pre market john viljoen bloomberg follow affect european energy share
today note market close due holiday news watch european stocks hina trade woes escalate alph alpiq half adjust ebitda
mln total reduces gonfreville refinery unit rates amid strike lsng lenenergo net bln rubles orste rste eversource sub
mit massachusetts wind farm proposal commodity crude brent crude natgas mmbtu agenda energy agenda oil daybook europe
earning alpiq holding maha energy energy wrap europe click energy sector wrap click contact reporter story john viljo
en cape town contact responsible story blaise robinson'
```

come hard to read. For interpretation purposes, the lemma is better than the linguistic root. After lemmatization, each news article will transform into a list of words that are all in their original forms.

Figure 2.8 and 2.9 shows an example of news article before and after text preprocessing. After tokenization, removing unnecessary words and lemmatization, the original news articles only contains informative words that are ready for further transformation, which will be discussed in Section 2.3.

2.3 News Sentiment Analysis

This section will discuss the news sentiment analysis methodology and how the positive score is calculated for every news article. First, this chapter analyze the sentiment of each unique word using a logistic regression. The estimated logistic regression coefficient for each unique word represents its sentiment. This chapter defines the effect's direction of each unique word by the sign of its coefficient, and the size of the effect by the absolute value of its coefficient. Moreover, by selecting the words with the highest absolute value in coefficients, this chapter defines the most positive and negative words indicating a price increase or decrease episode. Lastly, this chapter calculates the positive score for each news article based on how many positive and negative words this news article contains.

2.3.1 News Sentiment Analysis on Unique Words

News sentiment analysis is the analysis that uncovers the predicting power of each unique word in indicating a price increase or decrease episode. In order to do so, this chapter uses a supervised machine learning algorithm, which is called logistic regression. Logistic regression is a classification algorithm that deals with binary classification problems. Binary classification has exactly two classes to choose between. In this chapter, there are positive and negative classes indicating price increase or price decrease. Logistic regression is a linear classifier, it is a transformation from a linear function:

$$f(\boldsymbol{x}) = b_0 + b_1 * x_1 + ... + b_n * x_n$$

where b_0, $b_1 ... b_n$ are the estimators of the regression coefficients for a set of independent variable $\boldsymbol{x} = (x_1, x_2 ... x_n)$. The logistic regression function $p(\boldsymbol{x})$ is the sigmoid function of $f(\boldsymbol{x})$:

$$p(\boldsymbol{x}) = \sigma(f(\boldsymbol{x})) = \frac{1}{1 + exp^{-f(\boldsymbol{x})}}$$

After transformation, $p(\boldsymbol{x})$ will be in the range of $[0, 1]$, which can be interpreted as probability. Generally, $p(\boldsymbol{x})$ is interpreted as the predicted probability that $f(\boldsymbol{x})$ given $\boldsymbol{x}$ is equal to one, and $1 - p(\boldsymbol{x})$ is the probability that $f(\boldsymbol{x})$ is zero. In this chapter, $p(\boldsymbol{x})$ is defined as the probability that WTI crude oil futures price increases within five minutes after news article $\boldsymbol{x}_i$'s release.

Applying logistic regression to conduct news sentiment analysis, this chapter treats each news article as a observation, and the contents in news article as the features, and estimates β_{w0}, β_{w1}, ... β_{wj} from the following equation:

$$\begin{pmatrix} Y_0 \\ Y_1 \\ Y_2 \\ ... \\ Y_i \end{pmatrix} = \begin{pmatrix} X_{0,w0} & X_{0,w1} & ... & X_{0,wj} \\ X_{1,w0} & X_{1,w1} & ... & X_{1,wj} \\ X_{2,w0} & X_{2,w1} & ... & X_{2,wj} \\ ... & & & \\ X_{i,w0} & X_{i,w1} & ... & X_{i,wj} \end{pmatrix} * \begin{pmatrix} \beta_{w0} \\ \beta_{w1} \\ ... \\ \beta_{w_j} \end{pmatrix} \tag{2.1}$$

where i stands for each news article as a new observation, and wj is the jth unique word in all news articles. On the left hand side, Y_i is the price change dummy described in the previous section. Specifically, the value of Y is decided by the following conditions:

$$Y_i = \begin{cases} 1, & \text{if } price_{t+5} - price_t > 0 \\ 0, & \text{otherwise} \end{cases}$$

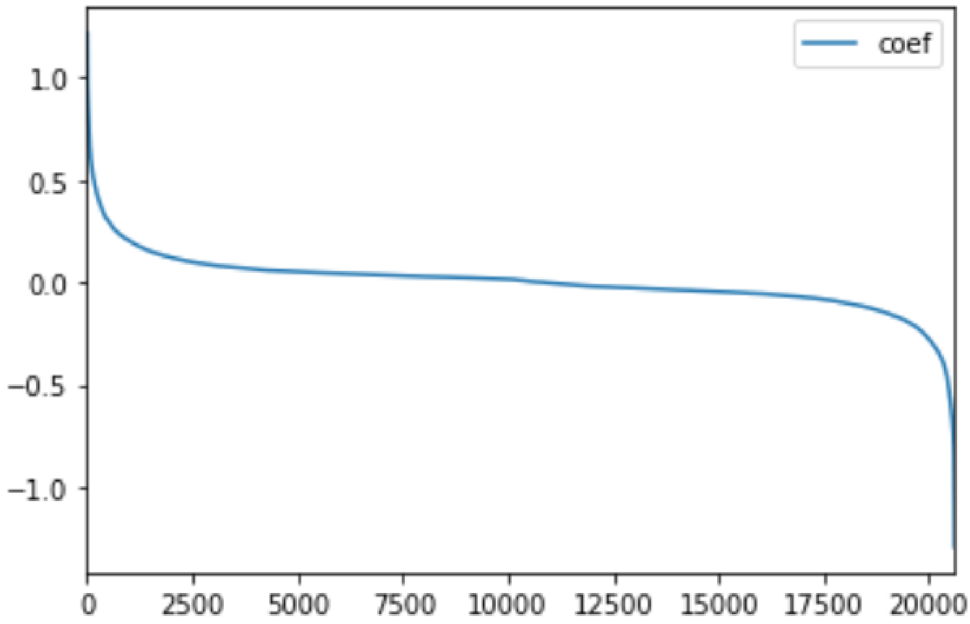
On the right hand side, the first term is a sparse matrix, with each row stands for each news article and each column stands for each unique word. There are over 20,606 unique words that has ever shown in 4616 news articles, which indicates the shape of the sparse matrix. Each value $X_{i,wj}$ of the sparse matrix is denoted as the tfidf value for each unique word wj in each news article i. Tfidf is short for term frequency, inverse document frequency. It is a common feature engineering method for text analysis and is widely used in literature. For example Bi and Traum(2020), Fraiberger(2019), Shapiro(2018) have used tfidf to extract text features for different analysis. Specifically, tfidf is calculated by:

$$X_{i,wj} = \frac{1 + log(t_{i,wj})}{1 + log(\sum_i^N t_{i,wj})} * log(\frac{N}{\sum_{wj} t_{i,wj}})$$

where $t_{i,wj}$ is the frequency of word wj appears in news article i. By examining the equation, it is clear that the first term is the calculating the term frequency and the second term is calculating the inverse document frequency. The first term is evaluating how many time the word wj appear in news article i, normalized by the length of news article i. The higher term frequency indicating a higher tfidf value, presenting the fact that the word wj plays a very important role in news article i by appearing significant times. However, the effect of wj will be weaken if wj also appears in many other news articles besides i, which means it is a common word for this topic. This process is captured by the second term, which is the inverse of how many news articles wj appears divided by the total number of news articles. In this case, N equals to 4616. Combining two effects, a word wj with high tfidf values in news article i means that wj appears many times in news article i, and only appears in few other news articles.

After transforming all preprocessed news articles into the sparse matrix, all

Figure 2.10: The Effect of Words



Note: Each point in the x axis stands for a unique word, collecting from all news articles

data are ready for regression. Figure 2.10 shows the estimation results from Equation 2.1 fitting a logistic regression model. Each point in the x-axis stands for a unique word collecting from all news articles, and there are 20,606 of them. The y-axis stands for the sign and the size of the coefficient for each word. Figure 2.10 indicates that most of the unique word by itself has very limited effect in affecting price, with coefficient very close to zero. However, there are some words that have coefficients with absolute value over 0.5, which are defined as the most positive and negative words by this chapter.

For the most positive and most negative words, based on the logistic regression

Figure 2.11: The Word Clouds

results, 152 words have coefficients smaller than -0.5, and they are collected as the most negative words. On the other side, 159 words have coefficient over 0.5, and are defined as the most positive words. In total, positive and negative words account for around 1.5% of all unique words. Figure 2.11 shows the word clouds of the most positive and most negative words. The font size of each word in the word clouds indicates the size of its coefficient in absolute value. Figure 2.11 shows that words like 'analyst','investing' are showing the greatest effect in indicating a crude oil futures price increase episode, and words like 'partner','hike' contribute the most to crude oil futures price decrease in a news article.

2.3.2 Positive Score for Each News Article

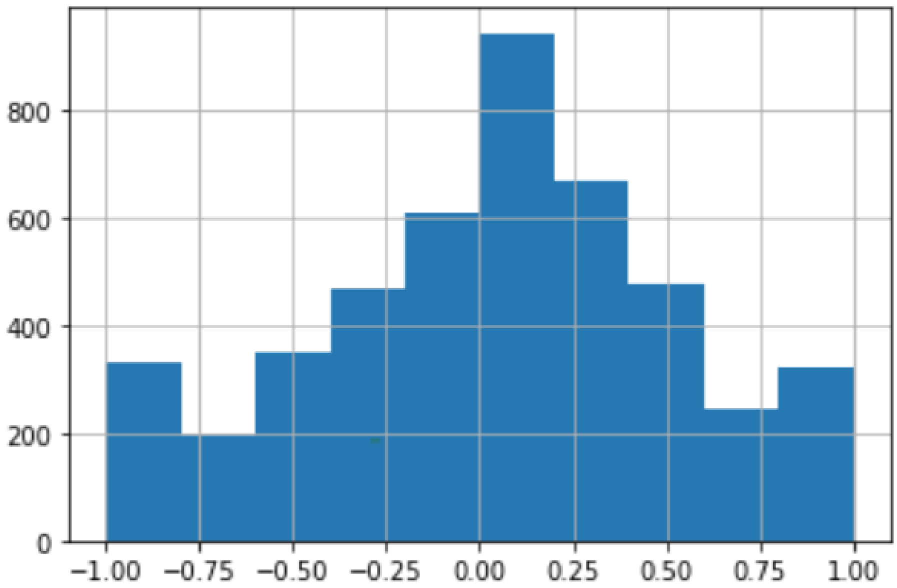
After defining the most positive and negative words in terms of contributing to price increase or decrease episodes, this section will discuss how to calculate the sentiment score for each news article. This chapter defines the sentiment score of a news article based on how many positive and negative word it contains. Specifically, this chapter calculates the positive score for each news article by comparing how many times positive words appear with the appearance of the negative words. Thus, the positive score of a news article i is defines as follow:

$$Score_i = \frac{N_{pos} * F_{pos} - N_{neg} * F_{neg}}{N_{pos} * F_{pos} + N_{neg} * F_{neg}} \tag{2.2}$$

where N_{pos} is the total counts of positive words in news article i, and F_{pos} is the total frequency of each positive/negative word in news article i. For example, if news article i has two positive words, and one appears 10 times and another one appears 15 times, N_{pos} will be two and F_{pos} will be 25. To calculate the positive score, this chapter finds the total occurance of positive words and negative words in each news article, takes the difference, and normalizes it by their summation. Figure 2.12 shows the positive score distribution across all news articles.

The positive score of a news article takes a value between $[-1, 1]$. A positive value means this news article is positive, while a negative value means this article is negative. As the absolute value of positive score increases, it becomes more positive or more negative depending on the sign. Figure 2.12 shows that most of the news article

Figure 2.12: The Positive Score Distribution



has a positive score close to zero, indicating the distribution is roughly a normal distribution with mean at zero, if ignore the two ends. If the positive score for a news article is one, it only has positive words and vice versa. Among all 4616 news articles, only around 600 articles are in such cases. Thus, this chapter still assume the monotonic positive correlation between the positiveness of a news article and its positive score.

In summary, this section uncovers the effect of each unique word collection from all news articles in the last quarter of 2019, attempts to estimate the coefficients for each unique word on how it affects the crude oil futures prices. In order to do so, this chapter uses a supervised machine learning algorithm called logistic regression to solve this binary classification problem. Logistic regression estimates the coefficients for all unique words and selects the most positive and negative words based on the coefficients' sign and size. The most positive and negative words help identify the sentiment score for each news article. By calculating the total occurrence of positive and negative words in a news article, this chapter defines the positive score for each news article, which is useful for the regression in Section 2.5.

2.4 Topic Analysis

Instead of reading each article and manually separating our news sample into different topics, we use an unsupervised machine learning algorithm, K-means, to detect common patterns in news articles and group them into clusters, i.e., topics.

2.4.1 K-means Algorithm

K-means is one of the most commonly used clustering algorithms in machine learning. Unlike supervised learning, which first defines a list of keywords in each topic and classify each article that includes those keywords into a specific topic, researchers first need to determine the number of clusters (topics), then the K-means algorithm will assign each sample to the cluster where its distance from the centroid of the cluster is minimized. The algorithm of K-means is described as followed:

1. Specify the number of K clusters. [5]

[5] In theory, the Elbow method helps to determine the optimal K values, by plotting K on the horizontal axis and sum of minimized distances of each cluster on the vertical axis. The optimal K is found when the slope of Elbow curve flattens, i.e., when the y value converges. However, in practice, the

2. Initialize the centroid point $\mu_k (k \in K)$ of each cluster with a random value.

3. Calculate the squared Euclidean distance of each sample x_i to the centroid point of each cluster.

$$||x_i - \mu_k||^2$$

4. Assign the sample x_i to the closest cluster k where the squared Euclidean distance is minimized.

$$\min_k ||x_i - \mu_k||^2$$

5. Update μ_k by taking the mean of sample points assigned to cluster k.

6. Repeat 3-5 until the sum of the squared distances overall K clusters is minimized.

$$\min \sum_{k=1}^{K} \sum_{x_i \in c_k} ||x_i - \mu_k||^2$$

where c_k denotes sample points in cluster k.

Following Bi and Traum(2019), we use the tfidf vector of each article (i.e., row vector of the tfidf matrix discussed in the previous section) as a sample point for training the algorithm, so that each article will be assigned uniquely to a cluster. The training set covers the whole sample period (20190101 - 20191231), which includes a total of 13,183 news.

When determining the value of K, the Elbow method doesn't provide us an optimal choice of K (the Elbow curve doesn't converge). Therefore, we run the algorithm several times with different K values and thus choose $K = 4$ based on the results

Elbow curve does not converge so that users need to determine the value of K based on their purpose. Our result of Elbow method can be provided upon request.

that provide us with the most meaningful and best interpretability of news topics.

2.4.2 Clustering Results

The K-means algorithm generates 4 clusters of news topics. Figure 2.13 plots the word clouds of top 50 important words[6] for each topic, with the more important ones shown in a bigger font. Given our specification of four news topics, we are able to interpret each topic to be: (a) - “World Crude Oil”, (b) - “WTI Crude Oil”, (c) - “Financial Analysis” and (d) - “Editorial Opinion”. The ”World Crude Oil” cluster has keywords related to the global oil market, such as market, trade, energy, China, USD, etc. ”WTI Crude Oil” cluster includes keywords like price, supply, demand, production, which reflects more about the information of WTI crude oil in the North American region[7]. ”Financial Analysis” contains keywords position, spread, call, put, option, relating to the price analysis, and use of derivatives for crude oil. The last cluster is ”Editorial Opinion,” with keywords story, contact, reporter, etc.

Table 2.1 shows the number of news in each topic. More than 80% of news throughout year 2019 are assigned to the “World Crude Oil” and “WTI Crude Oil” topics. 13.86% news are related to “Editorial Opinion”, which is more likely to be a summary or report of already released new information. There are only 563 news

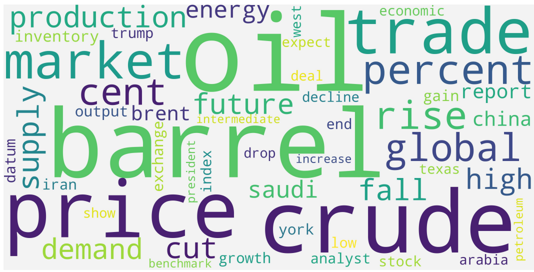
[6]For each cluster, we sum each word’s tfidf value over all the articles within this cluster, rank them from the highest to the lowest, and then pick the top 50 words for word cloud plot.

[7]There are some words shown in both of these two topics, which interprets these two topics to be unclear and need to be improved in future work.

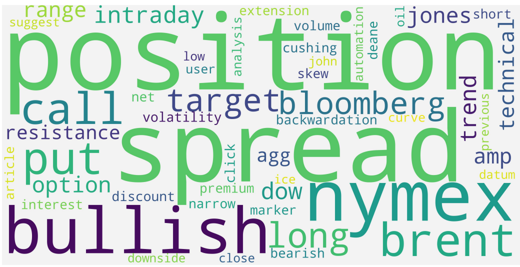
Figure 2.13: Word clouds of Clustering Results

(a) World Crude Oil

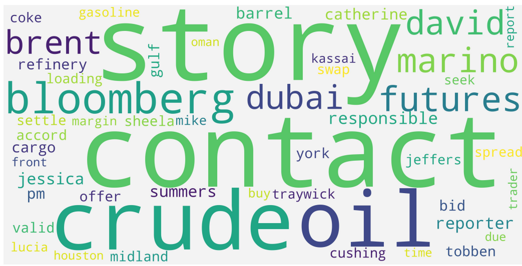
(b) WTI Crude Oil



(c) Financial Analysis



(d) Editorial Opinion



classified into "Financial Analysis" topic, which includes both analysis of price change and analysts' forecast of future price movement. As we can see in the next section, the last two topics are less important indicators of price changes.

Table 2.1: Number of News in Each Topic

Topic Index	**Topic Name**	**news_cnt**	**%**
a	World Crude Oil	6,386	48.44%
b	WTI Crude Oil	4,407	33.43%
c	Financial Analysis	563	4.27%
d	Editorial Opinion	1,827	13.86%
total		13,183	100%

2.5 Test News Topic and Sentiment Score on Futures Price Change

In this section, we estimate the relation of news topic and sentiment score with futures price change using logistic regression. By assuming that the WTI crude oil futures market is nearly perfectly efficient, investors responding quickly to any new information released to the market, futures price should adjust quickly in line with its intrinsic value. For example, a good news article about the discovery of new oil wells should push down the futures price immediately, as investors are competing with each other to find this price decrease opportunity. Since the market is nearly efficient, the futures price can reflect this new information very quickly. Given this assumption, we can test the relation between crude oil news released and its futures price change.

We first define our dependent variable to be one if futures price after 5 minutes is higher than the price when news released; and zero otherwise[8], as stated in Section 2.2. In future work, we will use different time spam, calculating price change for robustness checks. To test how crude oil news relates to price changes, we use the logistic regression, which estimates the influence of $\boldsymbol{x}$ on Y's probability for each news article. The regression equation we estimate follows:

$$Prob(Y_i = 1|\boldsymbol{x}) = \sigma(\beta_0 + \beta_1 Score_i + \sum_k^3 \beta_k Topic_k + \sum_k^3 \beta_{k,i} Topic_k * Score_i) \tag{2.3}$$

where

$$Y_i = \begin{cases} 1, & if \quad price_{t+5} > price_t \\ 0, & \quad otherwise \end{cases} \tag{2.4}$$

and $Topic$, a categorical variable, is {0: "World Crude Oil"; 1: "WTI Crude Oil"; 2: "Financial Analysis"; 4: "Editorial Opinion"}. In the regression, three dummy variables are generated indicating which topic this news article is in, and topic 0('World Crude Oil') is chosen as the benchmark topic. As described in section 2.3, *score* evaluates the degree of news sentiment. The higher the *Score* is, the more positive the news is.

The LHS of equation (2.3) denotes the probability of price increase ($Y = 1$) given all $\boldsymbol{x}$. As discussed in Section 2.3.1, on the RHS, $\sigma(\cdot)$ is a Sigmoid function. In the parenthesis of $\sigma(\cdot)$ is the linear combination of independent variables that we are

[8]The price data we use is high-frequency data, with 5 minutes interval. In our final regression sample (2019Q4), there doesn't exist price unchanged within each 5 minutes interval.

interested.

Our final logistic regression sample covers only the last quarter of 2019, with 4,616 observations due to the availability of high-frequency futures price data. Table 2.2 reports the regression results. The coefficient on constant is interpreted as the impact of news in topic 0 ("World Crude Oil") on the probability of price increase. On average, news in the "World Crude Oil" topic is more likely to be correlated with price increases. If the news is related to "WTI Crude Oil," the coefficient becomes 0.0159 (0.3267 minuses 0.3108), meaning that "WTI Crude Oil" news on average doesn't have a significant impact on a price increase. For topic 2 & 3 ("Financial Analysis" and "Editorial Opinion"), news has relatively less positive impacts on price increases, as their coefficients (-0.2821 and -0.2581) are both statistically significant and smaller than the coefficient of topic 0 (0.3267).

After considering the impact of sentiment score for each news under different topics, the coefficient on $Score$ can be interpreted as a correlation between the degree of topic 0 ("World Crude Oil") news sentiment and the probability of price increase. The result is statistically significantly positive (1.2642), meaning that the more positive "World Crude Oil" news is, the higher probability that price will increase within 5 minutes. The coefficient on the interaction term $Topic1 * Score$ (0.6034), is also positive and statistically significant, which implies that the sentiment score of news under topic 1 ("WTI Crude Oil") is more correlated with the probability of price increase than news under topic 0 ("World Crude Oil"). Coefficients of the interaction terms $Topic2 * Score$ and $Topic3 * Score$ are negative but not statistically significant. Thus we reject the hypothesis that news sentiment under topic 2 and 3 is statistically different from those under topic 0.

To sum up, on average, "World Crude Oil" news has the highest correlation with a price increase. Nonetheless, the more positive news is under the topic "WTI Crude Oil," the higher probability that WTI crude oil futures price will increase within 5 minutes.

2.6 Conclusion

With the development of machine learning algorithm and its use in economics literature, it is worthwhile to apply machine learning algorithm to understand news impact on financial asset price movement. In this project, we use both supervised and unsupervised machine learning algorithms to learn impact of news sentiment and news topics on crude oil futures price increase. By assuming that crude oil futures market are nearly perfectly efficient (price adjusts quickly to any new information released to the public), we use high frequency data for estimating the news impact. The results show that "World Crude Oil" news on average is positively correlated with price increase, and the more positive "WTI Crude Oil" news is, the higher probability that crude oil futures price will increase within five minutes. The implication of our project is that using the coefficients in our last regression results, we are able to construct a news index, which can be used further in estimating the magnitude of price change, together with other macro and micro economic control variables.

There are a lot of work can be done in the future to improve our results. For example, we will modify our filter of stop word in order to improve our sentiment results to be better consistent with human cognition. We can also try Neural Network algorithm to learn news topic, which as a supervised machine learning algorithm might provide more precise learning than the K-means algorithm that we used in the project.

Table 2.2: Test for News Effect on Price Change

$$Prob(Y = 1|X) = \sigma(\beta_0 + \beta_1 Score + \beta_2 Topic + \beta_3 Topic * Score)$$

Dependent Variable	**Coef.**	
Score	1.2642	***
	(0.082)	
Topic1 ”WTI Crude Oil”	-0.3108	***
	(0.073)	
Topic2 ”Financial Analysis”	-0.2821	*
	(0.161)	
Topic3 ”Editorial Opinion”	-0.2581	**
	(0.105)	
Topic1 * Score	0.6034	***
	(0.201)	
Topic2 * Score	-0.1910	
	(0.429)	
Topic3 * Score	-0.0868	
	(0.205)	
Const.	0.3267	***
	(0.044)	
No. Observations	4,616	
Pseudo R-squ.	0.071	
Robust Std. Err.	YES	

*Notes: The number in the parenthesis reports the t-statistics. *** indicates* $P < 0.01$*. ** indicates* $P < 0.05$*. * indicates* $P < 0.1$.

Chapter 3

What role did the Treasury play in Fed's large-scale asset purchase programs (LSAP)?

After the recent 2007-2008 financial crisis, several rounds of large-scale asset purchase programs (LSAP), also known as quantitative easing (QE), have been conducted by major central banks around the world. In particular, from the second round of QE, the Federal Reserve purchased longer-term Treasury securities by raising reserves and selling short-term Treasury bills they held. The purpose of the second and third round of QE is to reduce longer-term interest rates to stimulate aggregate demand when conventional monetary policy is not available as the short-term interest rate is around zero. This approach is also known as the central bank's balance sheet policy.

The mechanism of LSAP is to reduce the amount of longer-term Treasury bonds to the private sector, which increases bond prices for longer maturities and reduces longer-term interest rates while the short-term interest rate is kept at zero. In order for this mechanism to work, we need to assume short-term and long-term bonds are imperfect substitutes; otherwise, any difference in short-term and long-term interest rates can be arbitraged away by investors with full arbitrage opportunities. A broad literature finds that the LSAP did reduce long-term interest rates[1]. Meanwhile, some literature find that the LSAP is not quite as effective in stimulating aggregate demand. For example, Chen *et al.*(2012) calibrate the size of the second round of LSAP. They find that the posterior median effect on GDP growth is an increase of 0.13%, the posterior median inflation increase is 3bp. These findings imply that a calibrated $600 billion purchase by the Fed only had a modest effect in supporting economic activities.

The smaller than expected effect leads me to suspect that some potential mechanism or channel might be ignored when evaluating the effectiveness of LSAP. Greenwood *et al.*(2014) document the fact that when the Fed purchased longer-term Treasury securities from the private sector, the Treasury department, on the other hand, issued more longer-term Treasury securities at the same period of time. They adjust the outstanding Treasury debt into 10-year duration equivalents and find that the debt grew from $ 2 trillion in December 2007 to $ 6.3 trillion in July 2014, where the increase equaled to 25 percent of GDP. This expansionary debt issuance decision from the Treasury was in the opposite direction to influence long-term interest rates as the LSAP, so that its impact could deteriorate the effectiveness of LSAP. To fully understand the

[1]Hamilton and Wu (2010)[52] estimates a 13 basis point (bp) decrease in 10-year Treasury yield from the impact of LSAPs.

impact of LSAP on macroeconomic activities, it is essential to taking into account the debt issuance policy contemporaneously.

In the descriptive analysis part, I study the relation between longer-term Treasury debt outstanding and Fed's holding of these assets. My dataset includes two series, the first one is the US treasury securities held by the Fed, extracted from the Federal Reserve Economic Data (FRED); and the another is the sum of marketable Treasury Notes and Bonds held by the public from the monthly statement of public debt of the US, released by the Treasury Bulletin. Both of the two series are monthly data from January 2003 to March 2019. I find that both the correlations in level and growth rate are higher during the QE period relative to the periods before and after QE. This positive and higher correlation during the QE period may suggest that the estimated impact of LSAP on lowering longer-term interest rates is biased without considering the Treasury debt supply policy.

In the modeling part, I build a DSGE model that focuses on estimating the offsetting impact of an expansionary long-term debt issuance on the effectiveness of LSAP. The model introduces limit to households' arbitrage ability and bond market segmentation. I assume that households have heterogeneous preferences for different maturities of bonds. There is a fraction of households with specific preference that only invest in long-term bonds. For those who can invest in both short-term and long-term bonds, they pay certain transaction cost for purchasing long-term bonds, leading short-term and long-term bonds to be imperfect substitutes. These assumptions provide a channel for LSAP to be effective in influencing real activity.

When modeling the debt issuance policy, I establish a Taylor-rule type of debt supply policy. I first assume that the Treasury department would adjust long-term bonds issuance in response to output growth. Thus the offsetting impact of debt issuance on the effectiveness of LSAP can be evaluated by estimating different degrees of fiscal responsiveness to output. Furthermore, I develop fiscal policies for government expenditure and transfers respectively, and then assume that the issuance of long-term bonds adjust in response to these two fiscal instruments, which better characterizes the government financing need in each period.

In my model, I build in both central bank's QE policy and Treasury department's bond supply policy. Therefore, the long-term bond supply, which actively adjusts to the dynamic of government budget deficit, provides a channel for understanding and quantitatively evaluating Treasury's offsetting impact on LSAPs over financial and macroeconomic variables.

My project can further contribute to two implications for the interaction between the central bank and Treasury department. Similar to the "Operation Twist" program in 1961, the Federal Reserve and the Treasury department can cooperate with each other. When the short-term interest rate is fixed at some level, in order to lower longer-term interest rate for stimulating the economy, the Federal Reserve could choose to purchase longer-term Treasury bonds from the private sector. A cooperative Treasury department, at the same time, could actively adjust the issuance of short-term bonds while keeping long-term bonds outstanding stable. It can be found in my model that the QE policy is more effective under cooperation. Another implication is that even without cooperation, by observing the issuance behavior by the Treasury department, the Federal Reserve can decide the optimal level of asset purchases, given some target amount in lowering longer-term interest rate, or increasing in output.

3.1 Literature Review

There exists a large literature investigating the impact of QE, especially after the 2008 financial crisis and when the U.S. Federal Funds rate reached around zero. Kuttner (2018) provides an overview of those works studying the unconventional monetary policy, including the QE policy, where central bank purchases longer-term securities by increasing reserves or selling short-term securities held.

Event studies find that the LSAPs are effective in lowering 10-year Treasury yield. Gagnon *et al.*(2011) estimate the response of several financial variables to eight announcements within a one-day window, and they find a 91 basis points (bps) decline in ten-year Treasury yield due to QE1. Swanson (2011) studies yields response to the Operation Twist in 1961, and presents a 15bps drop in the 10-year Treasury yield. Swanson (2017) revisits the effect of QE2 by using a high-frequency event study, and concludes that both forward guidance and LSAPs are statistically significant in affecting ten-year yield. Greenlaw *et al.* (2018) use a different method to select events, which are called "Reuters Fed News" days, and conclude a smaller effects of QE1 on yields.

In addition to investigating the overall impact of QE on Treasury yields, Krishnamurthy and Vissing-Jorgensen (2011) and Bauer and Rudebusch (2013) find evidence that QE also works through the signaling channel, which conveys information on policy expectations. Bauer and Rudebusch (2013) decompose Treasury yields into term premium and the risk-neutral rate, which is the average of short-term interest rates over the maturity of bond. So that reducing in the risk-neutral rate component of long-term rates can be interpreted as a signaling channel impact of LSAPs. Their results suggest that about half of the decrease in ten-year yield is due to changing policy expectations via this signaling channel in QE1.

All these event studies restrict their event window to be one- or two-day (Swanson (2017) look at an even tighter window of 30 minutes), in the purpose of excluding other events that might have significant impact on yields, i.e., they are able to control for supply behavior conducted by the Treasury. My work is different from them in that my model can be used to investigate the offsetting impact of the Treasury's active bond supply on LSAPs, and estimate changes in aggregate output and other macroeconomic variables due to LSAPs.

There are also a large theoretical literature on modeling the impact of LSAPs. A majority of them study the impact of LSAPs via portfolio balance channel, with some others focus on the signaling channel effect, such as Bhattarai *et al.*(2019).

Some literature studies the portfolio balance effect of QE using the preferred habitat model of Vayanos and Vila (2009). Hamilton and Wu (2012) imply an affine term structure model with preferred habitat investors to estimate the impact of maturity structure of Treasury debt on the term structure of interest rates. Their results show that ten-year rate drops by 13 bps due to a $400 maturity swap at the zero lower bound.

Ray (2019) incorporates a preferred habitat term structure with a New Keynesian three-equation model to estimate the impact of QE. They assume that the "effective" nominal rate depends on the entire yield curve, so that it creates a channel for the LSAPs, which is considered to be an exogenous shock to the bond demand of preferred habitat investors, to have impact on output and inflation.

Chen *et al.*(2012) embed a preferred habitat household setup to a median-scale sticky price and wage DSGE model. In their model, a fraction of households (restricted), with preferred habitat, only invest in bonds with longer term maturities, and the rest of households (unrestricted) are indifferent between long-term and short-term bonds. In addition, they introduce asset imperfect substitution by assuming that there exists transaction costs of purchasing long-term bonds for unrestricted households. My modeling setup follows Chen*et al.*(2012) in heterogeneous households sector and sticky price and wage setup. I differs from them in modeling the QE and long-term bond supply policy separately, which provides a unique channel for Treasury debt management to impact the efficacy of LSAPs.

Another important set of chapters by Gertler and Kiyotaki (2011), Gertler and Karadi (2010) and Carlstrom *et al.*(2017) study the balance sheet of banks with incentive constraints. The market segmentation setup in their model is introduced by assuming households can access to long-term bonds indirectly through holding banks' deposit. Because of this financial intermediary's setup and the assumption of incentive constraint, their model can be used to study the impact of LSAP both in alleviating credit spread in private sector and in lowering longer-term interest rate. Sims and

Wu (2019a), incorporating this financial intermediary setup, develop a DSGE model to embed all unconventional policy tools: QE, forward guidance and negative interest policy together. They compare implementations of these unconventional policy tools with the conventional monetary policy, and find that QE is relatively effective, but with costs associated with quantitative tightening (QT).

This class of model finds larger effect of QE1 when there was considerable credit market dysfunction, but smaller effect of QE2 and QE3. In contrast to these literature, this project focuses on change in maturity structure of Treasury securities, which occurred in QE2 and QE3. Therefore, I choose to follow the preferred-habitat model structure in this project.

My research also relates to debt management literature. Greenwood *et al.*(2014) document that 35 percent of duration supply impact of QE is canceled by the Treasury maturity extension policy by looking at the change in ten-year duration equivalents. Inspired from this evidence, they solve for the optimal debt maturity for a trade-off model between minimizing issuance cost and managing fiscal risk. They argue that it is optimal for the Treasury to issue more short-term debt instead of tilting to long-term debt.

Belton *et al.*(2018) develop a model that solves for the optimal maturity of debt, which responds to the dynamic condition of budget deficits and interest rates. They conclude that if deficit rises, debt managers will increase the issuance of intermediate-term bonds by reducing short-term bonds. In contract to these chapters, I use a reduced-form of long-term debt supply policy, which is assumed to be actively managed by the

Treasury to meet government's primary financing needs, together with a central bank's QE policy to examine the interaction between these two policies on the supply of long-term bonds, and its impact on long-term rate and other macroeconomic variables in interest.

In the literature of fiscal policy, Leeper et al. (2010)[64] estimates response policies of four fiscal instruments to the US data, which are labor and consumption taxes, government expenditure and transfers. They find that having all four instruments adjust to stabilize both output and debt fits the US time series best, but only weakly preferred to having only transfers adjust. Thus, when constructing the fiscal policy in my model, I only include transfers and government expenditure as fiscal instruments for the government, and set their feedback rule to be responsive to output and debt level.

Kliem and Kriwoluzky[57] assume in their paper that fiscal policies are Taylor-rule type of feedback rules. They estimate fiscal policies for labor and capital tax, and find that, instead of responding to output, these two taxes are better described to respond to hours worked and investment, respectively. Following their idea of forming the fiscal policy, the long-term bond supply policy in my model is assumed to be also a Taylor-rule type of feedback rule, which responds to output growth at first, and later adjusted to respond to government financing need, which is described as the sum of government expenditure and transfers.

The structure of this chapter is: Section 3 presents the empirical exercise, Section 4 and 5 describe the model and results. Section 6 and 7 conclude and discuss future works.

3.2 Stylized Facts, Motives

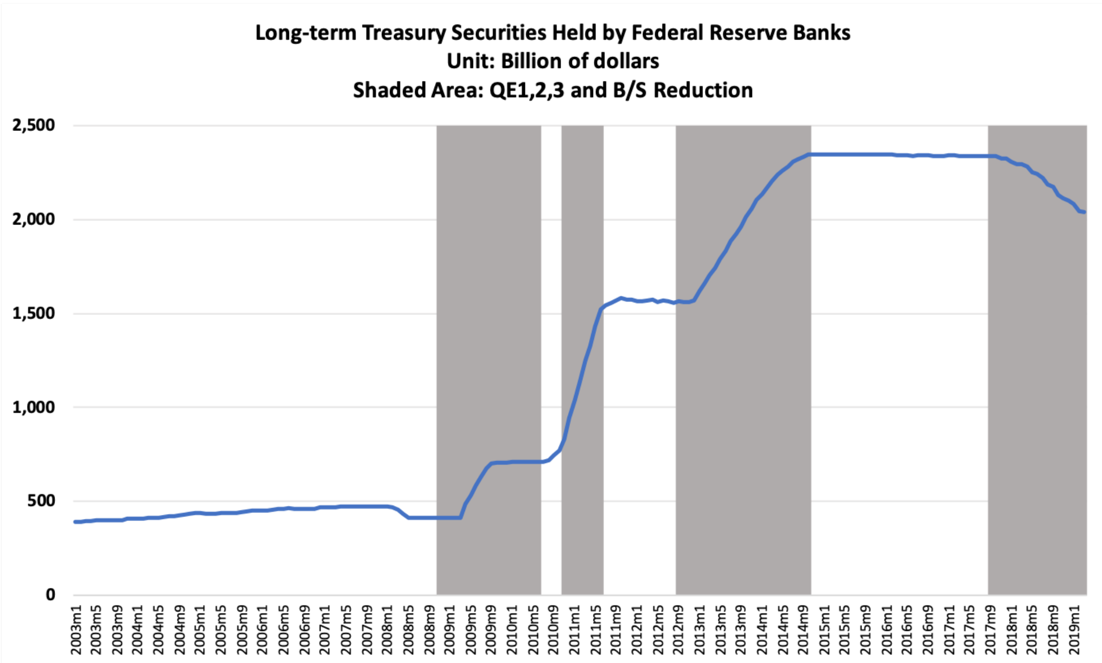
3.2.1 Relation Between Longer-term Bonds Outstanding and Fed Holding

Fact shows that the Federal Reserve, before the recession, held between $700 billion and $800 billion of Treasury notes on its balance sheet. During the first round of QE, it accumulated its holding of bank debt, mortgage-back securities (MBS), and Treasury notes to reach $2.1 trillion in June 2010. In QE2, an announcement of purchasing $600 billion of longer-term Treasury securities by the end of the second quarter of 2011 was released by FOMC. The last round of QE contained a $40 billion per month, open-ended bond purchasing program of agency MBS. This amount was raised to $85 billion per month, and later adjusted back to $65 billion per month. By 29 October 2014, the Fed accumulated $4.5 trillion assets purchased on its balance sheet.

Figure 3.1 shows the Federal Reserve Banks holding of the US Treasury notes and bonds from January 2003 to March 2019. The first three shaded areas denotes the three rounds of QE, and the last shaded area is the period when the Fed decreases their balance sheet size by rolling over the amount of principle payment matured that exceeds some level each month.

Meanwhile, the Treasury department increased its supply of Treasury notes and bonds sharply during the QE period. The total outstanding Treasury notes and bonds have increased from $ 780.9 billion in January 2007 to $2830.1 billion in September 2016. Figure 3.2 presents the evolution of Treasury notes and bonds outstanding for

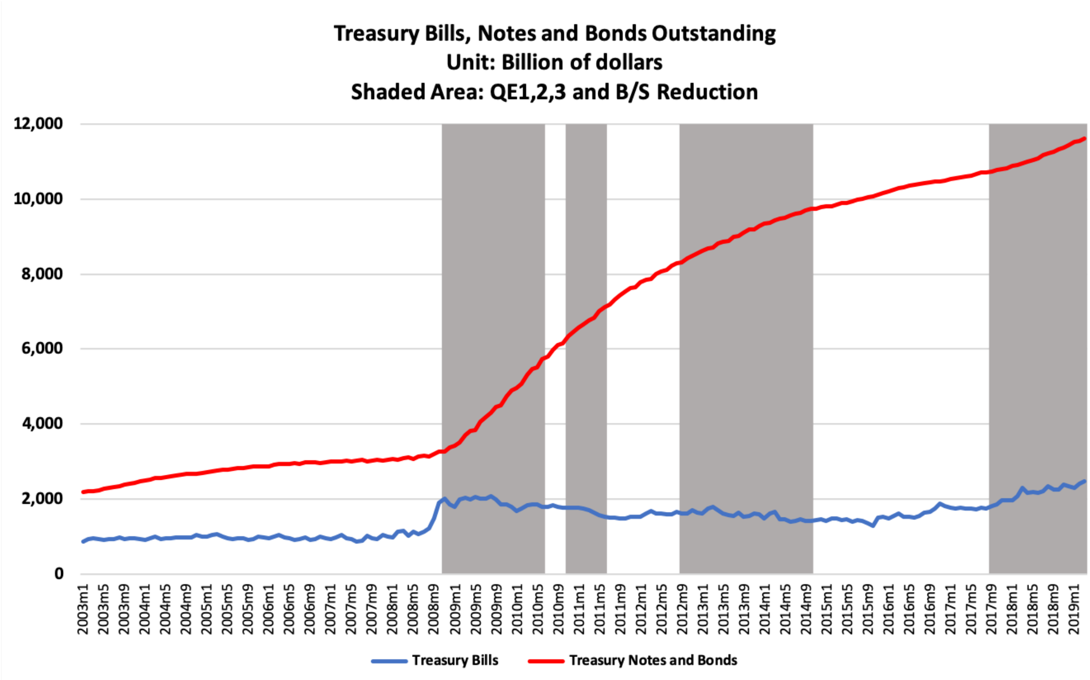
Figure 3.1: Long-term Treasury Securities Held by Federal Reserve Banks



Note: (i)The shaded areas denotes for the three rounds of Quantitative Easing and the Balance Sheet reduction period. (ii) The blue line denotes the sum of Treasury notes and bonds held by the Fed.

the same sample period.

Figure 3.2: Treasury Bills, Notes and Bonds Outstanding



Note: The shaded areas denotes for the three rounds of Quantitative Easing and the Balance Sheet reduction period.

To examine the relation between the Fed's hold of Treasury long-term (LT) securities [2] and the total outstanding LT securities, I analyzed two series: the first one is the Fed holding of LT Treasury securities as of last Wednesday of each month, extracted from the Federal Reserve weekly balance sheet released by Federal Reserve Statistical Release, the another is the outstanding marketable Treasury notes and bonds held by the public from monthly statement of public debt of the US, released by the Treasury

[2] I define long-term securities as those with maturities longer than 2 years, so they include Treasury notes and bonds.

Bulletin. Both of the two series are monthly data from January 2003 to March 2019.

Table 3.1 and Table 3.2 show the descriptive statistics of these two series in level and growth rate respectively for different sample periods: (i) all sample period; (ii) before QE (200301 – 200810); (iii) within QE (200811 – 201412); (iv) after QE (201501 – 201708) and (v) B/S reduction (201709 – 201903).The level correlation before QE is 0.7269, which raises to 0.9687 within the QE period, and drops to -0.9119 and -0.9939 after QE and during the B/S reduction period. The path in growth rate correlation is not as obvious as comparing to level correlation. The growth rate correlation before and within QE period are very close (0.2288 and 0.2281 respectively). After the QE, there is no apparent correlation between the growth rates of Fed's holding of LT debts and LT notes and bonds outstanding. The reason why growth rate correlation before QE is as of similar size to which between QE is because both the standard deviations of these two series are smaller than in the period within QE.

Table 3.1: Descriptive Statistics in Level

Sample Period	Longer-term Securities		Fed Holding		Correlation
	mean	std	mean	std	
All Sample(200301-201903)	6,499.03	3,342.21	1,281.80	823.97	0.9793
Before QE(200301-200810)	2,791.92	281.76	436.73	27.01	0.7269
Within QE(200811-201412)	7,175.89	2,004.43	1,382.28	620.35	0.9687
After QE(201501-201708)	10,288.97	284.00	2,343.24	3.37	-0.9119
BS Reduction(201709-201903)	11,137.59	282.59	2,216.13	102.99	-0.9939

Table 3.2: Descriptive Statistics in Growth Rate

Sample Period	Longer-term Securities		Fed Holding		Correlation
	mean	std	mean	std	
All Sample(200301-201903)	0.8636%	0.0105	0.8536%	0.0253	0.3882
Before QE(200301-200810)	0.0059%	0.0073	0.0781%	0.0108	0.2281
Within QE(200811-201412)	1.4851%	0.0131	2.3551%	0.0346	0.2228
After QE(201501-201708)	0.2807%	0.0015	-0.0122%	0.0005	0.1581
B/S Reduction(201709-201903)	0.4225%	0.0016	-0.7191%	0.0059	-0.2949

Figure A.1 presents the scatter plots for the growth rats of Fed's holding and LT debts outstanding exhibiting positive correlations both before and within QE period. The slope coefficients are 0.588 and 0.337 respectively, and both are statistically significant at 90% confidence level (as shown in Table A.8). The slope coefficient of 0.588 can be interpreted as when the Treasury Department increases growth rate of outstanding LT debt by 1 percentage point, the Federal Reserve will increase the growth rate of Fed's holding of LT debt by 0.588 percentage points within the period of QE, which is larger than 0.337 percentage points before QE.

The coefficient of growth rate of LT debt after QE is 0.053, and is not statistically significant, which implies that after QE, there is no significant positive comovement between changes in Fed's holding and LT debt outstanding. For the period of B/S reduction, the slope turns to negative (-1.114), and is not statistically significant.

Because in this time period, the Fed's holding is decreasing over time, while the total LT debt outstanding keeps increasing.

Given these evidences, I conclude that Fed's purchasing of Treasury long-term securities is positively correlated with the issuance of long-term debts, so that the net reduction in LT Treasury bonds supplied to the private sector is smaller in the case when we assume LT Treasury bonds supply is fixed. Thus, when studying the effect of Fed's LASP program, without considering the Treasury debt issuance policy would cause the estimated impact of LSAP on long-term interest rate to be biased upward.

3.3 Model

There are two types of households, unrestricted and restricted households, in the economy. Both types of households supply differentiated labor inputs and are owners of firms and receive dividends from firms. Competitive labor agencies combine the differentiated labor supply into a homogeneous composite of labor. Monopolistic competitive firms hire the labor composite to produce intermediate goods. Competitive final goods firms produce homogeneous goods by purchasing the intermediate goods as inputs. The central bank sets monetary policy, the government sets fiscal policy, and the Treasury department sets bonds supply policy.

3.3.1 Households

3.3.1.1 Individual Household Sector

A fraction ω of unrestricted households invest in both short-term and long-term government bonds. The remaining fraction $1 - \omega$ of restricted households, who have specific preference in longer maturity, only invest in long-term government bonds. This assumption captures the bond market segmentation in the model, and I will show in section 3.3.6 why it helps to ensure that the LSAP can impact real economic activities in the model.

Moreover, the restricted households are assumed to be more experienced in investing long-term bonds, so their transaction costs for purchasing long-term bonds are minimal. However, unrestricted households, who invest in both short-term and long-term bonds for the purpose of diversification, pay a transaction cost ξ_t for each unit purchase of long-term bonds. Thus, by this transaction cost setup, unrestricted households generally pay higher costs for purchasing long-term bonds than restricted households. Moreover, the short-term and long-term bonds are imperfect substitutes for unrestricted households. If there exists a difference between long-term and short-term bond yields, they can arbitrage away this difference, up to some transaction costs. However, on the other hand, restricted households do not have the opportunity for arbitrage.

The life-time utility function for household j is

$$E_t \sum_{t=0}^{\infty} \beta_j^t \left\{ \frac{(C_t^j)^{1-\sigma} - 1}{1-\sigma} - \varphi \frac{L_t^j(i)^{1+\eta}}{1+\eta} \right\} \tag{3.1}$$

where the superscript $j = ur, r$ denotes for the unrestricted and restricted, $\beta_j \in (0,1)$ is the discount factor, $\sigma > 0$ is the coefficient of relative risk aversion.

There are two types of bonds issued by Treasury department, which are both assumed to be non-defaultable. Short-term bonds B_t^S are one-period sovereign bonds purchased at time t and promise to pay a nominal return of R_t at time $t+1$. Long-term bonds B_t^L are perpetuity contracts that cost P_t^L at time t, and promise to pay an exponentially decaying coupon κ^s at time $t+s+1$, for $\kappa \in (0,1)$. By this setup, P_t^L can be written as $P_t^L = \frac{1}{R_t^L - \kappa}$, where R_t^L is the gross yield to maturity at time t on the long-term bond.

The budget constraints are different for different types of households. For unrestricted households, they can invest in both short-term and long-term government bonds, their budget constraint is:

$$C_t^{ur} + B_t^{s,ur} + (1+\xi_t) P_{L,t} B_t^{L,ur} \leq R_{t-1} B_{t-1}^{s,ur} + \sum_{s=1}^{\infty} \kappa^{s-1} B_{t-s}^{L,ur} + W_t^{ur}(i) L_t^{ur}(i) + \mathcal{P}_t + \mathcal{P}_t^{cp} + \mathcal{P}_t^{fi} \tag{3.2}$$

For restricted households, they only invest in long-term bonds but with no transaction cost, their budget constraint is:

$$C_t^r + P_{L,t} B_t^{L,r} \leq \sum_{s=1}^{\infty} \kappa^{s-1} B_{t-s}^{L,r} + W_t^r(i) L_t^r(i) + \mathcal{P}_t + \mathcal{P}_t^{cp} + \mathcal{P}_t^{fi} \tag{3.3}$$

W_t is the wage set by a household i who supply labor, $\mathcal{P}$, $\mathcal{P}_t^{cp}$,are the dividends distributed by intermediate goods producers and capital goods producers respectively. The transaction costs are paid by the unrestricted households to some financial institute, which is also owned by households, thus its profit $\mathcal{P}_t^{fi}$ is distributed back to both the unrestricted and restricted households.

3.3.1.2 Labor Agencies

Households supply differentiated labor inputs that are purchased by labor agencies that aggregate labor into a labor index L_t given by

$$L_t = \left[\int_0^1 L_t(i)^{\frac{\theta_n - 1}{\theta_n}} di \right]^{\frac{\theta_n}{\theta_n - 1}} \tag{3.4}$$

where $\frac{\theta_n}{\theta_n - 1}$ is the wage markup. This labor index is used by intermediate producing firms in production.

The optimization problem for labor agencies yields the demand for labor i as

$$L_t(i) = \left[\frac{W_t(i)}{W_t} \right]^{-\theta_n} L_t \tag{3.5}$$

where

$$W_t = \left[\int_0^1 W_t(i)^{1-\theta_n} di \right]^{\frac{1}{1-\theta_n}}$$

In each period, a fraction of $1 - \omega_w$ of households are randomly selected and allowed to adjust their wages. The remaining fraction ω_w of households set their wage equal to the optimal wage set in the previous period with a growth rate of steady-state inflation rate,

$$W_{t+s}^j(i) = (\pi)^s \tilde{W}_t^j(i) \tag{3.6}$$

If a household of type i is allowed to adjust wage at time t, she chooses $\tilde{W}_t^j(i)$ to maximize

$$E_t \sum_{s=0}^{\infty} (\beta_j \omega_w)^s \left\{ \lambda_{t+s}^j (\pi)^s \tilde{W}_t^j(i) L_{t+s}^j(i) - \varphi \frac{(L_{t+s}^j(i))^{1+\eta}}{1+\eta} \right\}$$

subject to (7) and (8), where $j = ur, r$. λ_{t+s} is the marginal utility of consumption at time $t+s$. The first-order condition for this problem gives us the optimal wage

$$(\tilde{W}_t^j(i))^{\theta_n \eta + 1} = \frac{\theta_n}{\theta_n - 1} \frac{E_t \sum_{s=0}^{\infty} (\beta_j \omega_w)^s \varphi \pi^{-s\theta_n(1+\eta)} (W_{t+s})^{\theta_n(1+\eta)} L_{t+s}^{1+\eta}}{E_t \sum_{s=0}^{\infty} (\beta_j \omega_w)^s \lambda_{t+s}^j \pi^{s(1-\theta_n)} (W_{t+S})^{\theta_n} L_{t+s}}.$$

and the aggregate wage

$$W_t^{1-\theta_n} = (1-\omega_w)\left[(\omega \tilde{W}_t^{ur})^{1-\theta_n} + ((1-\omega)\tilde{W}_t^{r})^{1-\theta_n}\right] + \omega_w (\pi W_{t-1})^{1-\theta_n}.$$

3.3.2 Firms

3.3.2.1 Capital Goods Producers

Competitive capital producers make investment decisions, and rent capital to intermediate goods producers. They discount future profits at the weighted average of shareholder's marginal utility, by assuming that dividends are equally distributed among all households.

$$\gamma^{t+s} = \omega \beta^{ur} \lambda_{t+s}^{ur} + (1-\omega) \beta^r \lambda_{t+s}^r.$$

where γ is the capital producers discount, $\lambda^j (j = (ur, r))$ is the marginal utility for households. Capital producers maximize the expected discounted sum of dividends to their shareholds

$$E_t \sum_{s=0}^{\infty} \gamma^{t+s} \{ R^k_{t+s} K_{t+s-1} - I_{t+s} \}.$$

subject to law of motion of capital follows:

$$K_t = (1-\delta) K_{t-1} + \left[1 - S(\frac{I_t}{I_{t-1}}) \right] I_t. \tag{3.7}$$

where δ is the depreciation rate of capital. I_t is investment in each period, $S(\cdot)$ is the investment adjustment cost with $S'(\cdot) \geq 0$ and $S'' > 0$.

3.3.2.2 Final Goods Producers

Perfectly competitive final goods firms combine differentiated intermediate goods $Y_t(f)$ into a homogeneous good Y_t given by

$$Y_t = \left[\int_0^1 Y_t(f)^{\frac{\theta_f - 1}{\theta_f}} df \right]^{\frac{\theta_f}{\theta_f - 1}} \tag{3.8}$$

where $\frac{\theta_f}{\theta_f - 1}$ is the price markup.

The demand for intermediate goods can be derived from solving the optimization problem of final goods producing firms.

$$Y_t(f) = \left[\frac{P_t(f)}{P_t} \right]^{-\theta_f} Y_t \tag{3.9}$$

where

$$P_t = \left[\int_0^1 P_t(f)^{1-\theta_f} df \right]^{\frac{1}{1-\theta_f}}.$$

3.3.2.3 Intermediate Goods Producers

Monopolistic competitive firms hire labor to produce intermediate goods. The production function is

$$Y_t = Z_t K_t^{\alpha} L_t^{1-\alpha}. \tag{3.10}$$

where Z_t is aggregate productivity which follows a Markov process.

For the sticky price setup, in each period t, $(1-\omega_f)$ fraction of firms can adjust prices. ω_f fraction of firms cannot adjust their prices, and set prices equal to $P_{t+s}(f) = \tilde{P}_t(f)\pi^s$. For those who can choose price, they pick $\tilde{P}_{t+s}(f)$ that maximize

$$E_t \sum_{s=0}^{\infty} \omega_f^s \gamma^{t+s} [\tilde{P}_{t+s}(f) Y_{t+s}(f) - \mu_{t+s} Y_{t+s}(f)].$$

where μ_t is the firm's real marginal cost.

The first-order condition for this problem gives us the optimal price

$$\tilde{P}_t(f) = \frac{\theta_f}{\theta_f - 1} \frac{E_t \sum_{s=0}^{\infty} (\omega_f)^s \triangle_{t+s} \mu_{t+s} \pi^{-s\theta_f} P_{t+s}^{\theta_f} Y_{t+s}}{E_t \sum_{s=0}^{\infty} (\omega_f)^s \triangle_{t+s} \pi^{s(1-\theta_f)} P_{t+s}^{\theta_f} Y_{t+s}}.$$

and the average price: $P_t^{1-\theta_f} = (1-\omega_f)(\tilde{P}_t)^{1-\theta_f} + \omega_f (\pi P_{t-1})^{1-\theta_f}$.

3.3.3 (Un)conventional Monetary Policy

The central bank sets a conventional monetary policy that follows Taylor rule, and also takes into account of the deviation of inflation rate from steady state value and the growth rate of output.

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\rho_m} \left[\left(\frac{\pi_t}{\pi}\right)^{\phi_\pi} \left(\frac{Y_t}{Y_{t-4}}\right)^{\phi_y}\right]^{1-\rho_m}. \tag{3.11}$$

where $\pi_t \equiv P_t/P_{t-1}$ is the inflation rate, $\rho_m \in (0,1)$, $\phi_\pi > 1$, $\phi_y \geq 0$.

In addition to conventional monetary policy, the central bank also conducts QE policy, which is purchasing long-term bonds by exchanging short-term bonds they held or raising the interest-bearing reserves on their liabilities. I assume that return on short-term bonds and interest on reserves are the same, thus making short-term bonds and reserves to be perfect substitutes, and the central bank's balance sheet always holds. Following Sims and Wu (2019a), the central bank's QE policy can be described either exogenously, or endogenously. Here I will first consider an exogenous QE policy.[3] The QE policy is assumed to follow an exogenous AR(1) process:

$$\frac{B_t^{L,CB}}{P_t} = \left(\frac{B_{t-1}^{L,CB}}{P_{t-1}}\right)^{\rho_{qe}} e^{\epsilon_{qe,t}} \tag{3.12}$$

The central bank conducts QE policy both in normal time and in the case of ZLB. The LSAP is considered here to be a positive shock to this QE policy. It can be interpreted as central bank increases demand for the quantity of long-term bonds, which increases price and decreases long-term interest rate.

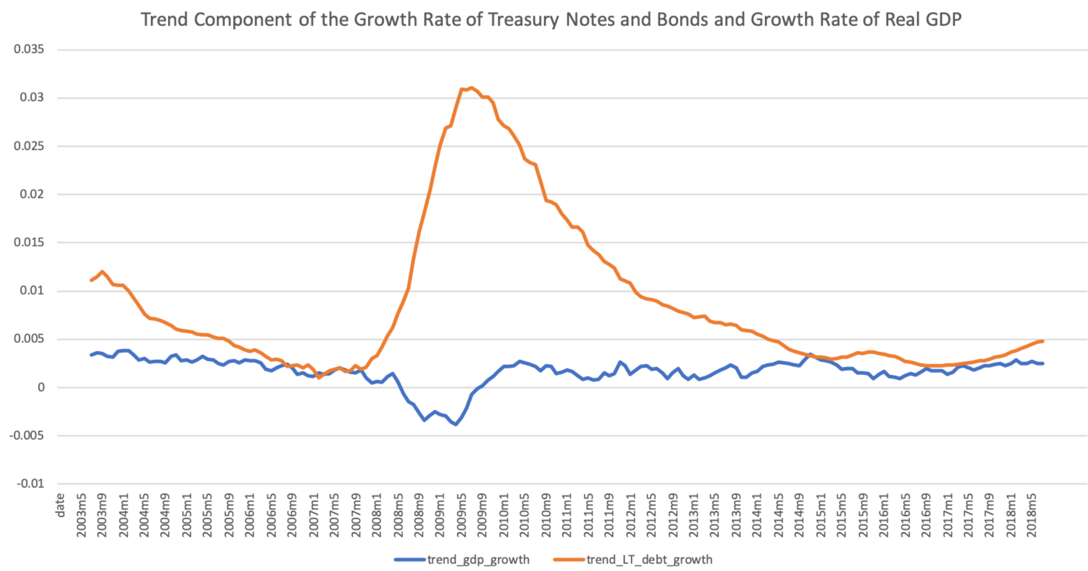
3.3.4 Long-term Bond Supply Policy

Figure3.3 shows the trend components of the growth rate of Treasury notes and bonds and the growth rate of real GDP. During the QE period, there exhibits some

[3]In future work, I will also develop an endogenous QE policy.

negative correlations between these two series. The correlation between the two growth rates series is -0.09973 over the whole sample period (200301–201903), -0.1874 within the QE period, and are very close to zero before and after QE period. This result confirms that when the economy is in a bad state, Treasury department issues more long-term notes and bonds, which pushes up longer-term interest rates, and deteriorates the impact of LSAP.

Figure 3.3: Trend Components of Growth Rate of Treasury Notes and Bonds and Growth Rate of Real GDP



Note: Monthly Real GDP is obtained from IHS Markit, and the growth rate is calculate by taking the first-order log-difference.

To formalize the long-term bond supply policy, the first case I consider follows from the objectives of Treasury debt management, which includes managing aggregate demand and promoting financial stability. When making issuance decision, the Treasury department is supposed to monitor the economic conditions. I also assume that the Treasury department only manage the supply of long-term bonds actively, and let the short-term bonds to balance the remaining government financing need. They would issue more long-term bonds to meet higher government financing need, which is used to stimulate the aggregate demand when the economy is in a bad state. Thus, the long-term bond supply policy is assumed to be responsive negatively to output growth. In addition, for the purpose of issuing debt in a regular and predictable pattern, the supply of long-term bonds also depends on its previous issuance level.

$$\frac{P_t^L B_t^L}{P_t} = \left(\frac{P_{t-1}^L B_{t-1}^L}{P_{t-1}}\right)^{\rho_B} \left[\left(\frac{Y_t}{Y_{t-4}}\right)^{\phi_{B,y}}\right]^{1-\rho_B}. \tag{3.13}$$

where $\phi_{B,y} < 0$ captures the response of long-term bonds issuance to the output growth. The total amount of short-term and long-term bonds issued should meet all the government financing need, including the repayment of principle for debt matured, coupon payments for the outstanding debts, and additional financing need decided by the government. The government budget constraint is

$$B_t^S + P_t^L B_t^L = R_{t-1} B_{t-1}^S + (1 + \kappa P_t^L) B_{t-1}^L + G_t \tag{3.14}$$

The left-hand side of (15) is the amount received by the government from issuing new short-term and long-term bonds at time t. The right-hand side includes the amount needed to repay for the outstanding debts, plus G_t the government expenditure, which is assumed to be given exogenously.

3.3.5 Equilibrium and Bond Market Clearing

In equilibrium, households maximize their discounted expected life time utility subject to their budget constraints, firms maximize their profits. The aggregate resource constraint is:

$$Y_t = \omega C_t^W + (1-\omega)C_t^I + G_t + I_t. \tag{3.15}$$

Bond market participants include unrestricted, restricted households and the central banks from the demand side, the Treasury department from the supply side. Market clearing conditions are:

$$\begin{aligned} B_t &= B_t^S + B_t^L && (3.16)\\ B_t^S &= B_t^{S,ur} + B_t^{S,CB} && (3.17)\\ B_t^L &= B_t^{L,ur} + B_t^{L,r} + B_t^{L,CB} && (3.18) \end{aligned}$$

The three Euler equations in the model demonstrate intuition of bond market segmentation and present channels through which LSAP can impact real economy.

Euler equation: Unrestricted, short

$$\lambda_t^{ur} = \beta_{ur} R_t E_t \left[\frac{1}{\pi_{t+1}} \lambda_{t+1}^{ur} \right] \tag{3.19}$$

Euler equation: Unrestricted, long (given $P_t^L = \frac{1}{r_t^L - \kappa}$)

$$\frac{1+\xi_t}{R_t^L - \kappa} \lambda_t^{ur} = \beta_{ur} E_t \left[\frac{R_{t+1}^L}{R_{t+1}^L - \kappa} \frac{1}{\pi_{t+1}} \lambda_{t+1}^{ur} \right] \tag{3.20}$$

Euler equation: Restricted, long

$$\frac{1}{R_t^L - \kappa} \lambda_t^r = \beta_r E_t \left[\frac{R_{t+1}^L}{R_{t+1}^L - \kappa} \frac{1}{\pi_{t+1}} \lambda_{t+1}^r \right] \tag{3.21}$$

From (3.20) and (3.21), the long-term interest rate R^L_{t+1} can affect the consumption-saving decision for both unrestricted and restricted households, but the transaction cost ξ_t only affects unrestricted households' intertemporal consumption decision. The central bank purchases long-term securities reduces transaction costs, which induces drop in long-term yields for unrestricted households. For restricted households, since they are not subject to transaction costs, fall in long-term yields causes a change in their stochastic discount factor, thus impact the restricted households consumption path, and indirectly influences the investment decisions of capital producers. Eventually, consumption for both groups of households, investment and output respond as well in general equilibrium.

3.3.6 Imperfect Asset Substitution and Limits to Arbitrage

In order for the LSAP to lower the long-term interest rates and further affect the aggregate activities, it is important to incorporate imperfect asset substitution and market segmentation into the DSGE model. In this subsection, I will show how LSAP lowers the long-term interest rate by changing the transaction costs, why it is not sufficient for LSAP to influence the aggregate activities by only including transaction costs capturing imperfect asset substitutes, and why it is crucial for us to further assume market segmentation (i.e. limits to arbitrage) in the model.

3.3.6.1 Transaction Costs

In Andres *et al.*(2004), the imperfect substitutability is modeled to include two frictions. First, households pay an additional time-varying stochastic transaction costs for purchasing long-term bonds. This exogenous transaction costs correspond to Tobin's "exogenous interest differentials". Nevertheless, Tobin(1969) argues that this exogenous spread only captures one part of the wedges between asset prices, the remaining part of the wedge is captured by the relative quantities of assets. Andres *et al.*(2004), then secondly, introduce a portfolio adjustment cost into households' utility function. Because long-term bonds are riskier than short-term bonds, due to a loss of liquidity, households require additional reserve to compensate for this loss of liquidity when purchasing long-term bonds. The reserve requirement is captured by a portfolio adjustment cost in their model.

Following Chen *et.al.*(2012), I combine the exogenous and endogenous component of asset wedge into one transaction cost, which is defined to be a function of the market value of long-term bonds relative to short-term bonds available to households sector, and plus an error

$$\xi_t \equiv f\left(\frac{P_t^L B_t^{L,HH}}{B_t^S}, \epsilon_t^{\xi}\right) \tag{3.22}$$

where $B_t^{L,HH} = B_t^{L,ur} + B_t^{L,r}$, and the function $f(\cdot) > 0$ and its first derivative $f'(\cdot) > 0$ in steady state. Under these two assumptions, the interest rate differential, which is represented by this transaction cost, is always positive at steady state, and a reduction in long-term bonds available to households lowers long-term interest rate from a decrease in transaction cost. In order to show this mechanism, one can linearize equation(23) and (24) around a zero inflation steady state, and combining $R_t = 1+i_t$ and $\frac{R_t^L}{\pi_{t+1}}\frac{P_{t+1}^L}{P_t^L} = 1+r_t^L$ to obtain

$$i_t - \rho + E_t(\hat{\lambda}_{t+1}^{ur} - \hat{\lambda}_t^{ur} - \pi_{t+1}) = E_t r_t^L - \hat{\xi}_t - \rho + E_t(\hat{\lambda}_{t+1}^{ur} - \hat{\lambda}_t^{ur}) \tag{3.23}$$

where ρ is the steady state real return, and $\hat{\xi}_t \equiv \xi_t - \xi$, where ξ is the steady-state of ξ_t. Rearranging terms in equation(27), one can show that

$$E_t r_t^L = i_t - E_t \pi_{t+1} + \hat{\xi}_t$$

Thus, the difference between the expected one-period real returns on long-term and short-term bonds is exactly equal to the transaction cost. Moreover, suppose in the case when the supply of long-term bonds is fixed, because of this one to one change of transaction costs on yield of long-term bonds, the LSAP, which reduces transaction costs by lowering the long-term bonds relative to short-term bonds available to households, will cause r_t^L to decrease by the same amount, but will not affect unrestricted households' inter-temporal consumption decision (i.e. $\hat{\lambda}_{t+1}^{ur} - \hat{\lambda}_t^{ur}$). However, if the supply of long-term bonds also increases at the same time, change in transaction costs would be arbitrary.

3.3.6.2 Limits to Arbitrage

The restricted households are assumed to only invest in long-term bonds, which captures the market segmentation in the model. In reality, there exists a fraction of population who only enters the bond market with their preference on specific bond maturities. In particular, people purchases long-term bonds in the purpose of using predetermined coupon payments to meet their financing need over some future time period, and setting the length of obligation to match with bonds' maturities. Because their investment strategy focuses more on bonds with specific maturities, they are more experienced in analyzing and trading for these financial assets, thus they might purchase these assets at relatively lower prices than those who invest in bonds with different maturities for the purpose of diversification. In other word, the transaction costs for preferred preference investors are likely to be smaller.

The assumption that restricted households only purchase long-term bonds can also be relaxed. Thus restricted households can purchase both short-term and long-term bonds, but due to their specific preference, they will always want to purchase some amount of long-term bonds every period, and use remaining wealth after consumption and purchase of long-term bonds to construct their bond investment portfolio, i.e. purchasing short-term bonds and additional long-term bonds for the purpose of diversification. In this case, the restricted households now obtain limited arbitrage ability. However, this limited arbitrage ability is similar to that of the unrestricted households. Thus, for simplicity and tractability, I assume that restricted households only purchase long-term bonds in this model.

The parameter ω (fraction for unrestricted households) in the model captures this market segmentation, which could be a larger number close to 1, because most of the bond market participants' portfolios are well or at least partially diversified. This market segmentation assumption can also ensure that the LSAP effect the aggregate activities in the model. As shown above in section 3.3.6.1, decrease in transaction costs will decrease the yield on long-term bonds, but will not change the inter-temporal consumption decisions for unrestricted households. However, the reduction in r_t^L has a real effect on restricted households' inter-temporal consumption decisions. Rewrite equation(3.21) yields:

$$\lambda_t^r = \beta_r E_t(1 + r_t^L)\lambda_{t+1}^r.$$

Thus a decrease in r_t^L causes restricted households to increase their consumption today, which further changes decisions of capital goods producers and intermediate goods producers and in consequence, increases current output.

3.4 Parameter Values and IRF

3.4.1 Parameter Values

I solve the model by taking a first-order log-linear approximation around steady states. At this stage, in order to show a preliminary simulation exercise below, I choose most of the parameter values from Chen *et al.*(2012). However, in the future work, I will estimate these parameters on the US data for calibration.

I choose productivity to be 1 at steady state ($z = 1$), and the steady-state inflation rate is normalized to be at 2%. The consumption fraction of unrestricted to restricted ($\frac{c^{ur}}{c^r}$) is set to be 0.918, which I follow the prior mean of this fraction in Chen et al.*(2012)* results. The share of capital in production α is 0.33, and the depreciation rate δ is set to be 2.5%. I set ω (fraction of unrestricted households), which identifies the degree of bond market segmentation, to be 0.7, which implies that about 70% households are unrestricted households. At steady state, the market value of long-term bonds is about 25% of annual GDP. The ratio of central bank long-term bond holding to total market value of long-term bonds is calibrated to be 18%, which is equivalent to 4.5% of annual GDP at steady state. For the long-term bonds supply policy in Equation (3.13), I follow Chen et al.(2012)[23] to calibrate $\rho_b = 0.8$, and choose two values for parameter $\phi_{B,y} = 0, -1$. $\phi_{B,y} = 0$ implies that the bond issuance policy does not response to output, i.e. without automatic stabilizers. $\phi_{B,y} = -1$ indicates that when GDP growth rate decreases by 1 percentage point, supply of long-term bonds increases by 0.2 percent. Table A.7 presents the other parameter values used in model simulation below.

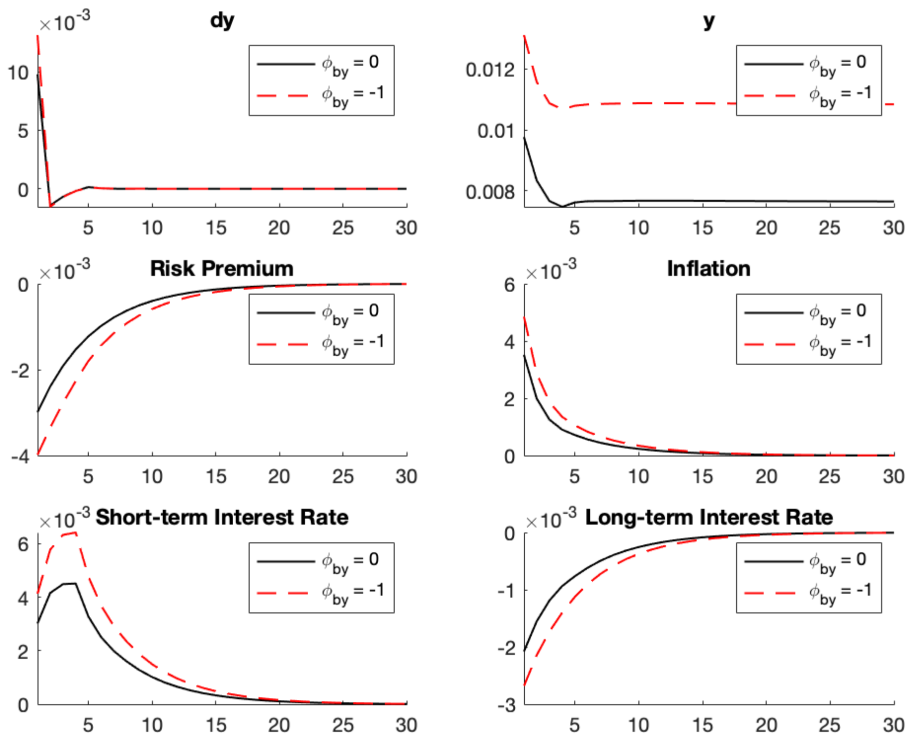
3.4.2 Model Simulations

3.4.2.1 Pure QE effect

Figure 3.4 compares the impulse responses for different values of $\phi_{B,y}$ subject to a positive one standard deviation shock to QE policy ($\epsilon_{qe,t}$). In my baseline model ($\phi_{B,y} = 0$) (shown in black line), after central bank conducts QE policy through purchasing long-term bonds , transaction cost falls, and the expected returns on long-term bonds drops. The drop in long-term interest rate leads to a drop in risk premium [4]. Short-term interest rate and inflation increases due to decrease in long-term interest rate and risk premium, and also increases output. For the case when $\phi_{B,y} = -1$ (shown in

[4]The risk premium is calculated from the difference between $\hat{R_t^L}$ and $\hat{R_t^{L,EH}}$ up to a first order approximation, where $R_t^{L,EH}$ denotes the yield on long-term bonds to the unrestricted households without transaction costs.

Figure 3.4: IRF for a positive QE shock



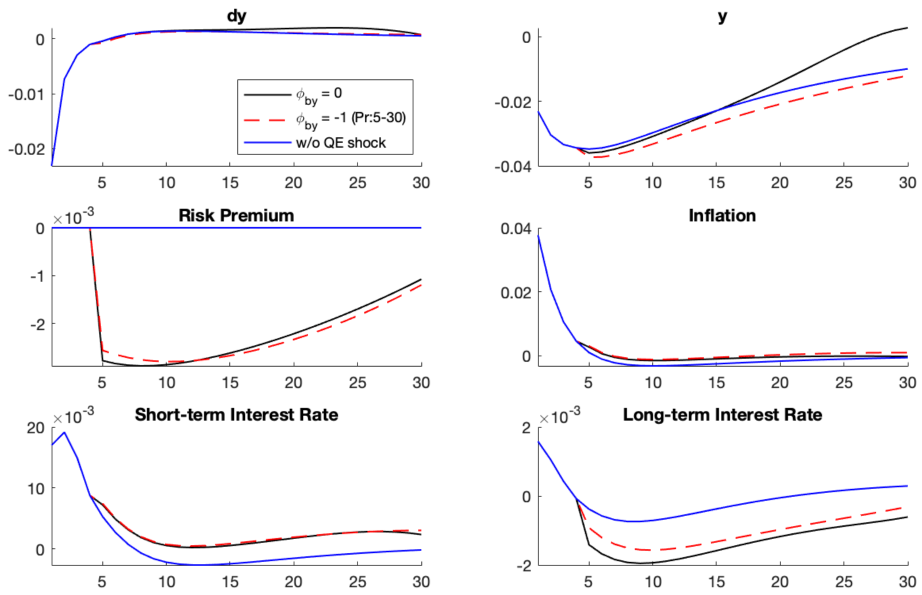
red dashed line), the long-term bond issuance policy now responses to output growth. A positive QE policy shock again lowers the long-term interest rate. However, in the case when the economy starts at steady state, output increases due to a positive QE policy shock, which will further decrease the supply of long-term bonds and lower long-term interest rate as shown in the figure.

3.4.2.2 QE effect Following a Negative Productivity Shock

In order to better investigate the offsetting impact of long-term bonds supply on the effect of LASP, I further simulate the model to be hit by a negative productivity shock at the the initial period, and then impose a positive QE shock from period 5. Figure 3.5 shows the results. The blue line, as the benchmark, shows the impulse response to a negative productivity shock, without any QE policy shock and $\phi_{B,y}$ always equals to 0. The black line shows the impulse response to incorporate a positive QE shock at period 5 with $\phi_{B,y} = 0$, and the red dashed line shows the impulse response to incorporate a positive QE shock at period 5 with $\phi_{B,y} = 0$ for the first 4 periods, and $\phi_{B,y} = -1$ for the remaining 25 periods.

This figure shows clearly, that when the Treasury department actively manages their supply of long-term bonds to be negatively responsive to output growth, the reduction in long-term interest rate, and risk premium are smaller than in the case when long-term bonds supply doesn't respond to output growth. The output in red dashed line reaches back to zero more slowly than the output in black line, which implies that the QE policy in reducing long-term interest rates and stimulating aggregate output is deteriorated by the expansionary long-term bonds supply policy.

Figure 3.5: IRF of One Standard Deviation of Negative Shock to z at period 1, and Positive QE Shock from period 5



3.5 Extended Long-term Bond Supply Policy

I also consider the case when the supply of long-term bonds responds directly to additional government financing need, which is described to be the sum of government expenditure and transfers [5]. The long-term bonds supply policy becomes:

$$\frac{P_t^L B_t^L}{P_t} = \left(\frac{P_{t-1}^L B_{t-1}^L}{P_{t-1}}\right)^{\rho_B} \left[\left(\frac{G_t}{G}\right)^{\phi_{B,g}} \left(\frac{TR_t}{TR}\right)^{\phi_{B,TR}}\right]^{1-\rho_B}. \tag{3.24}$$

where $\phi_{B,g} > 0$ and $\phi_{B,TR} > 0$ capture the response of long-term bonds issuance to the government expenditure and transfers respectively. The government budget becomes:

$$B_t^S + P_t^L B_t^L = R_{t-1} B_{t-1}^S + (1 + \kappa P_t^L) B_{t-1}^L + G_t + TR_t \tag{3.25}$$

Following Leeper et al.(2010)[64], the government conducts fiscal policy following rules which include "automatic stabilizer" component to movements in fiscal instruments, which are also permitted to respond to the state of government debt. The two log-linearized fiscal rules are described as:

$$\hat{g_t} = \Psi_g \hat{y_t} + \nu_g \hat{b_{t-1}} \tag{3.26}$$

$$\hat{tr_t} = \Psi_{TR} \hat{y_t} + \nu_{TR} \hat{b_{t-1}} \tag{3.27}$$

where $\Psi_g < 0$, $\Psi_{TR} < 0$ are response coefficients of automatic stabilizer.

Finally, the amount of short-term bonds issued is set to ensure that the government budget constraint always holds.

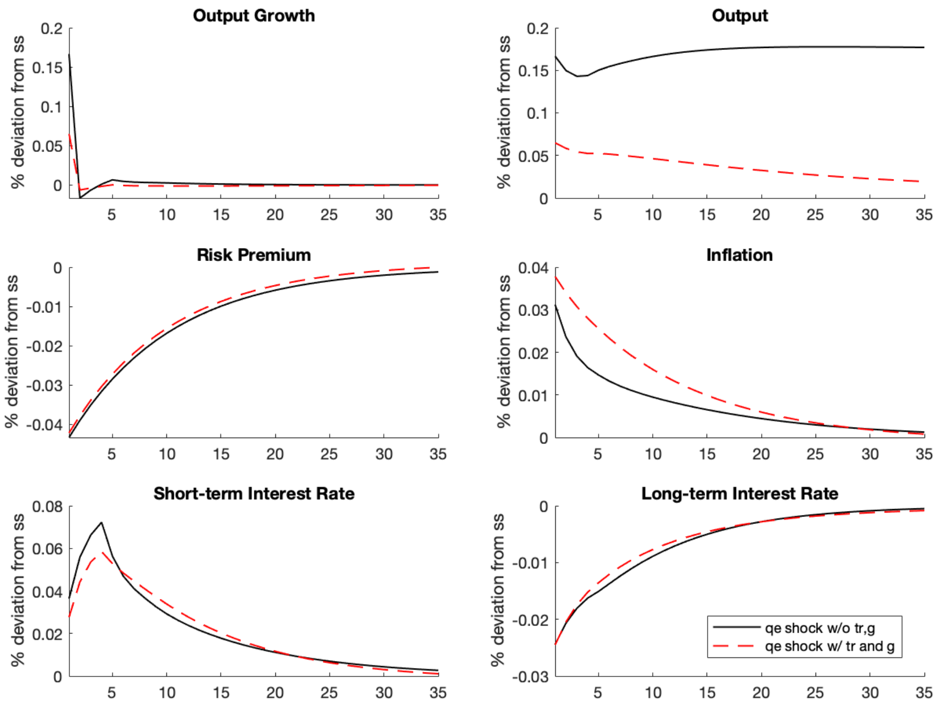
3.5.1 Calibration and Model Simulation

I estimate parameters in the two fiscal policies by running regressions for Equation 3.26 and 3.27 during the QE period. At this stage, for simplicity, I first set ν_g and

[5] Transfers can also interpreted as the net of government transfers minus lump-sum taxes. It is positive when government distributes net transfers to households, and negative when the lump-sum taxes received by the government exceeds the distribution of transfers.

H

Figure 3.6: IRF for a positive QE shock-Extended



ν_{TR} to be zero, so that growth of government spending and transfers depend only on output growth. The estimated Ψ_g and Ψ_{TR} are -0.03 and -0.13 respectively, which implies that government spending and transfers are conducted as automatic stabilizers (the responses are negative). The coefficients $\phi_{B,g}$ and $\phi_{B,TR}$ are estimated to be 0.1 and 1, where in the regression of Equation 3.24 the coefficients of government spending and transfers are 0.02 and 0.2 respectively.

Figure 3.6 show the impulse responses to a positive stock to QE policy. The black lines show the case without government spending and transfers in long-term bond supply policy. The red dashed lines show the responses in a model where government use its spending and transfers to respond as automatic stabilizers and set long-term supply policy respond actively to government financing need. Given a positive QE shock, which causes an decrease in long-term interest rate and rise in output growth, government would actively adjust its spending and transfers to be smaller, and thus deteriorate the growth of output (as shown in the red dashed line in the upper left sub figure). Because the estimation of coefficients on transfers in Equation 3.27 is larger than the one associated with government spending, and change in transfer could directly impact households consumption decisions, the offsetting impact from the fiscal part is thus reflected in the smaller output growth.

In general, the central bank would choose to use the QE policy when short-term interest rate is at zero. Thus model with zero lower bound setup would shed more light on the responses of long-term interest rate and output growth due to the interplay between central bank and fiscal department in determining the supply of long-term treasury bonds. Furthermore, the long term bond supply policy in this project is a Taylor rule type of policy, it is also worthwhile to develop an optimal policy by taking into account of the target and constraint facing by the treasury department. These work should be done in my future study.

3.6 Conclusion

In this chapter, I examine the interaction between central bank's QE policy and bond issuance policy by the Treasury department. Both levels and growth rates of Fed holding of long-term treasury securities and long-term debts outstanding are found to be positively correlated during the QE period. In a bad state of economy, the Treasury department tends to issue more long-term bonds. This increase in supply of long-term bonds would deteriorate the intent of LSAP to decrease long-term interest rates through lowering the private sector long-term bonds. Thus isolating bond issuance policy from its responding to changes in government financing need would cause the estimated impact of LSAP to be biased.

To study the offsetting effect of bond issuance policy on LSAP, this chapter incorporate central bank balance sheet policy and debt issuance policy by fiscal authority in a DSGE model with bond market segmentation and long-term and short-term bonds imperfect substitution. Long-term bond issuance policy is designed to be a Tylor rule type of policy, which responses positively to previous long-term bonds issuance and negatively to output growth or government financing need. My results show that the responsiveness to output in long-term bonds supply policy does have offsetting impact on the effectiveness of LSAP in lowering long-term interest rate and risk premium.

Bibliography

[1] Isha Agarwal, Grace Weishi Gu, and Eswar S Prasad. China's impact on global financial markets. Technical report, National Bureau of Economic Research, 2019.

[2] Javier Andrés, J David López-Salido, and Edward Nelson. Tobin's imperfect asset substitution in optimizing general equilibrium. *Journal of Money, Credit and Banking*, pages 665–690, 2004.

[3] Ye Bai and Darien Yan Pang Chow. Shanghai-hong kong stock connect: An analysis of chinese partial stock market liberalization impact on the local and foreign markets. *Journal of International Financial Markets, Institutions and Money*, 50:182–203, 2017.

[4] Brad M Barber and Terrance Odean. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The review of financial studies*, 21(2):785–818, 2008.

[5] Robert J Barro. On the determination of the public debt. *Journal of political Economy*, 87(5, Part 1):940–971, 1979.

[6] Michael D Bauer and Glenn D Rudebusch. The signaling channel for federal reserve bond purchases. *International Journal of Central Banking, forthcoming*, 2013.

[7] Christiane Baumeister and Lutz Kilian. A general approach to recovering market expectations from futures prices with an application to crude oil. 2014.

[8] Terry Belton, Kristopher Dawsey, David Greenlaw, Huachen Li, Srini Ramaswamy, and Brian Sack. Optimizing the maturity structure of us treasury debt: A model-based framework. Technical report, Hutchins Center Working Paper 46, 2018.

[9] Saroj Bhattarai, Gauti B Eggertsson, and Bulat Gafarov. Time consistency and the duration of government debt: A signalling theory of quantitative easing. Technical report, National Bureau of Economic Research, 2015.

[10] Huixin Bi and Nora Traum. Sovereign risk and fiscal information: A look at the us state default of the 1840s. *Federal Reserve Bank of Kansas City Working Paper*, (19-04), 2019.

[11] Sushil Bikhchandani and Sunil Sharma. Herd behavior in financial markets. *IMF Staff papers*, 47(3):279–310, 2000.

[12] Henning Bohn. The economic consequences of rising us government debt: privileges at risk. *FinanzArchiv: Public Finance Analysis*, 67(3):282–302, 2011.

[13] Michael W Brandt and Lin Gao. Macro fundamentals or geopolitical events? a textual analysis of news events for crude oil. *Journal of Empirical Finance*, 51:64–94, 2019.

[14] Joseph P Byrne, Giorgio Fazio, and Norbert Fiess. Primary commodity prices: Co-movements, common factors and fundamentals. *Journal of Development Economics*, 101:16–26, 2013.

[15] Cristiano Cantore, Paul Levine, Giovanni Melina, and Joseph Pearlman. Optimal fiscal and monetary policy, debt crisis, and management. *Macroeconomic Dynamics*, pages 1–39, 2017.

[16] Charles T Carlstrom, Timothy S Fuerst, and Matthias Paustian. Targeting long rates in a model with segmented markets. *American Economic Journal: Macroeconomics*, 9(1):205–42, 2017.

[17] Paul Cashin, Hong Liang, and C John McDermott. How persistent are shocks to world commodity prices? *IMF Staff Papers*, 47(2):177–217, 2000.

[18] Paul Cashin, C John McDermott, and Alasdair Scott. Booms and slumps in world commodity prices. *Journal of development Economics*, 69(1):277–296, 2002.

[19] Wesley S Chan. Stock price reaction to news and no-news: drift and reversal after headlines. *Journal of Financial Economics*, 70(2):223–260, 2003.

[20] Eric C Chang, Joseph W Cheng, and Ajay Khorana. An examination of herd behavior in equity markets: An international perspective. *Journal of Banking & Finance*, 24(10):1651–1679, 2000.

[21] An-Sing Chen and Bi-Shia Hong. Institutional ownership changes and returns

around analysts' earnings forecast release events: Evidence from taiwan. *Journal of Banking & Finance*, 30(9):2471–2488, 2006.

[22] Gong-Meng Chen, Bong-Soo Lee, and Oliver Rui. Foreign ownership restrictions and market segmentation in china's stock markets. *Journal of Financial Research*, 24(1):133–155, 2001.

[23] Han Chen, Vasco Cúrdia, and Andrea Ferrero. The macroeconomic effects of large-scale asset purchase programmes. *The economic journal*, 122(564):F289–F315, 2012.

[24] Yu-Fen Chen, Chih-Yung Wang, and Fu-Lai Lin. Do qualified foreign institutional investors herd in taiwan's securities market? *Emerging Markets Finance and Trade*, 44(4):62–74, 2008.

[25] Nicole Choi and Richard W Sias. Institutional industry herding. *Journal of Financial Economics*, 94(3):469–491, 2009.

[26] Nicole Choi and Hilla Skiba. Institutional herding in international markets. *Journal of Banking & Finance*, 55:246–259, 2015.

[27] William G Christie and Roger D Huang. Following the pied piper: Do individual returns herd around the market? *Financial Analysts Journal*, 51(4):31–37, 1995.

[28] Vicentiu Covrig, Sie Ting Lau, and Lilian Ng. Do domestic and foreign fund managers have similar preferences for stock characteristics? a cross-country analysis. *Journal of International Business Studies*, 37(3):407–429, 2006.

[29] Vasco Curdia and Michael Woodford. The central-bank balance sheet as an instrument of monetarypolicy. *Journal of Monetary Economics*, 58(1):54–79, 2011.

[30] Angus Deaton and Guy Laroque. On the behaviour of commodity prices. *The review of economic studies*, 59(1):1–23, 1992.

[31] Marco Del Negro, Gauti Eggertsson, Andrea Ferrero, and Nobuhiro Kiyotaki. The great escape? a quantitative evaluation of the fed's liquidity facilities. *American Economic Review*, 107(3):824–57, 2017.

[32] Elizabeth Demers, Clara Vega, et al. Soft information in earnings announcements: News or noise? 2008.

[33] Rıza Demirer and Ali M Kutan. Does herding behavior exist in chinese stock markets? *Journal of international Financial markets, institutions and money*, 16(2):123–142, 2006.

[34] Janice Eberly, James H Stock, and Jonathan H Wright. The federal reserve's current framework for monetary policy: A review and assessment. *NBER Working Paper*, (w26002), 2019.

[35] Gauti B Eggertsson et al. Zero bound on interest rates and optimal monetary policy. *Brookings papers on economic activity*, 2003(1):139–233, 2003.

[36] John Elder, Hong Miao, and Sanjay Ramchander. Jumps in oil prices: the role of economic news. *The Energy Journal*, 34(3), 2013.

[37] Hao Fang, Jehn-Yih Wong, Yu-Chun Wang, and Ralph Lu. Information contents of qfiis' cascades in the taiwan stock market. *Available at SSRN 1343433*, 2009.

[38] Lei Feng and Mark S Seasholes. Correlated trading and location. *The Journal of Finance*, 59(5):2117–2144, 2004.

[39] John Fernald and John H Rogers. Puzzles in the chinese stock market. *Review of Economics and Statistics*, 84(3):416–432, 2002.

[40] Joseph Gagnon, Matthew Raskin, Julie Remache, Brian Sack, et al. The financial market effects of the federal reserve's large-scale asset purchases. *international Journal of central Banking*, 7(1):3–43, 2011.

[41] Diego Garcia. Sentiment during recessions. *The Journal of Finance*, 68(3):1267–1300, 2013.

[42] Mark Gertler and Peter Karadi. A model of unconventional monetary policy. *Journal of monetary Economics*, 58(1):17–34, 2011.

[43] Mark Gertler, Peter Karadi, et al. Qe 1 vs. 2 vs. 3...: A framework for analyzing large-scale asset purchases as a monetary policy tool. *international Journal of central Banking*, 9(1):5–53, 2013.

[44] Mark Gertler and Nobuhiro Kiyotaki. Financial intermediation and credit policy in business cycle analysis. In *Handbook of monetary economics*, volume 3, pages 547–599. Elsevier, 2010.

[45] Mark Gertler, Nobuhiro Kiyotaki, and Albert Queralto. Financial crises, bank risk exposure and government financial policy. *Journal of Monetary Economics*, 59:S17–S34, 2012.

[46] Reuven Glick and Michael Hutchison. China's financial linkages with asia and the global financial crisis. *Journal of International Money and Finance*, 39:186–206, 2013.

[47] David Greenlaw, James D Hamilton, Ethan Harris, and Kenneth D West. A skeptical view of the impact of the fed's balance sheet. Technical report, National Bureau of Economic Research, 2018.

[48] Robin Greenwood, Samuel Hanson, Joshua S Rudolph, and Lawrence H Summers. *Government debt management at the zero lower bound*. Hutchins Center on Fiscal & Monetary Policy at Brookings Institution, 2014.

[49] Robin Greenwood, Samuel G Hanson, and Jeremy C Stein. A comparative-advantage approach to government debt maturity. *The Journal of Finance*, 70(4):1683–1722, 2015.

[50] Stephane Guibaud, Yves Nosbusch, and Dimitri Vayanos. Preferred habitat and the optimal maturity structure of government debt. *Available at SSRN 1098195*, 2008.

[51] James D Hamilton. Understanding crude oil prices. *The Energy Journal*, 30(2), 2009.

[52] James D Hamilton and Jing Cynthia Wu. The effectiveness of alternative monetary policy tools in a zero lower bound environment. *Journal of Money, Credit and Banking*, 44:3–46, 2012.

[53] Henning Hesse, Boris Hofmann, and James Weber. The macroeconomic effects of asset purchases revisited. *Journal of Macroeconomics*, 58:115–138, 2018.

[54] Rui Huo and Abdullahi D Ahmed. Return and volatility spillovers effects: Evaluating the impact of shanghai-hong kong stock connect. *Economic Modelling*, 61:260–272, 2017.

[55] Lutz Kilian. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review*, 99(3):1053–69, 2009.

[56] Lutz Kilian and Clara Vega. Do energy prices respond to us macroeconomic news? a test of the hypothesis of predetermined energy prices. *Review of Economics and Statistics*, 93(2):660–671, 2011.

[57] Martin Kliem and Alexander Kriwoluzky. Toward a taylor rule for fiscal policy. *Review of Economic Dynamics*, 17(2):294–302, 2014.

[58] Gerhard Kling and Lei Gao. Chinese institutional investors' sentiment. *Journal of International Financial Markets, Institutions and Money*, 18:374–387, 2008.

[59] Arvind Krishnamurthy and Annette Vissing-Jorgensen. The aggregate demand for treasury debt. *Journal of Political Economy*, 120(2):233–267, 2012.

[60] Kenneth N Kuttner. Outside the box: Unconventional monetary policy in the great recession and beyond. *Journal of Economic Perspectives*, 32(4):121–46, 2018.

[61] Josef Lakonishok, Andrei Shleifer, and Robert W Vishny. Contrarian investment, extrapolation, and risk. *The journal of finance*, 49(5):1541–1578, 1994.

[62] Kinger Lau, Timothy Moe, Ben Bei, and Chenjie Liu. Sh-hk connect: New regime, unprecedented opportunity. *Goldman Sachs*, 2014.

[63] Eric M Leeper. Why central banks should care about fiscal rules. *Sveriges Riksbank Economic Review*, 3:109–125, 2016.

[64] Eric M Leeper, Michael Plante, and Nora Traum. Dynamics of fiscal financing in the united states. *Journal of Econometrics*, 156(2):304–321, 2010.

[65] Canlin Li and Min Wei. Term structure modelling with supply factors and the federal reserve's large scale asset purchase programs. 2012.

[66] Hong Li. The impact of china's stock market reforms on its international stock market linkages. *The Quarterly Review of Economics and Finance*, 52(4):358–368, 2012.

[67] Wei Li, Ghon Rhee, and Steven Shuye Wang. Differences in herding: Individual vs. institutional investors. *Pacific-Basin Finance Journal*, 45:174–185, 2017.

[68] Wei Li and Steven Shuye Wang. Do institutional trades stabilize the retail investor dominated market. Technical report, Citeseer, 2008.

[69] Anchor Y Lin and Peggy E Swanson. The behavior and performance of foreign investors in emerging equity markets: Evidence from taiwan. *International Review of Finance*, 4(3-4):189–210, 2003.

[70] Andy Lin and Chih-Yuan Chen. The impact of qualified foreign institutional investors on taiwan's stock market. *Web Journal of Chinese Management Review*, 9(2):1–27, 2006.

[71] Wensheng Lin. Modeling volatility linkages between shanghai and hong kong stock markets before and after the connect program. *Economic Modelling*, 67:346–354, 2017.

[72] Ningyue Liu, Don Bredin, Liming Wang, and Zhihong Yi. Domestic and foreign institutional investors' behavior in china. *The European Journal of Finance*, 20(7-9):728–751, 2014.

[73] Tim Loughran and Bill McDonald. Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4):1187–1230, 2016.

[74] Robert E Lucas Jr and Nancy L Stokey. Optimal fiscal and monetary policy in an economy without capital. *Journal of monetary Economics*, 12(1):55–93, 1983.

[75] Yu Luo, Chengsi Zhang, and Yueteng Zhu. Openness and financial development in china: The political economy of financial resources distribution. *Emerging Markets Finance and Trade*, 52(9):2115–2127, 2016.

[76] Maxime Merli and Tristan Roger. What drives the herding behavior of individual investors? *Finance*, 34(3):67–104, 2013.

[77] John R Nofsinger and Richard W Sias. Herding and feedback trading by institutional and individual investors. *The Journal of finance*, 54(6):2263–2295, 1999.

[78] Mitchell A Petersen. Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies*, 22(1):435–480, 2009.

[79] Avihai Rapaport. Supply and demand shocks in the oil market and their predictive power. *Available at SSRN 2472379*, 2010.

[80] Walker Ray. Monetary policy and the limits to arbitrage: Insights from a new keynesian preferred habitat model. Technical report, Technical report, 2019.

[81] Stephanie Schmitt-Grohé and Martin Uribe. Optimal fiscal and monetary policy in a medium-scale macroeconomic model. *NBER Macroeconomics Annual*, 20:383–425, 2005.

[82] Michael Schuppli and Martin T Bohl. Do foreign institutional investors destabilize china's a-share markets? *Journal of International Financial Markets, Institutions and Money*, 20(1):36–50, 2010.

[83] Adam Hale Shapiro, Moritz Sudhof, and Daniel Wilson. Measuring news sentiment. Federal Reserve Bank of San Francisco, 2020.

[84] Susan Sunila Sharma, Paresh Narayan, and Kannan Thuraisamy. Time-varying

herding behavior, global financial crisis, and the chinese stock market. *Review of Pacific Basin Financial Markets and Policies*, 18(02):1550009, 2015.

[85] Robert J Shiller. *Irrational exuberance: Revised and expanded third edition.* Princeton university press, 2015.

[86] Richard W Sias. Institutional herding. *The Review of Financial Studies*, 17(1):165–206, 2004.

[87] Eric R Sims and Jing Cynthia Wu. Evaluating central banks' tool kit: Past, present, and future. *Present, and Future (February 28, 2019)*, 2019.

[88] Cindy Soo. Quantifying animal spirits: news media and sentiment in the housing market. *Ross School of Business Paper*, (1200), 2015.

[89] Eric T Swanson. Let's twist again: A high-frequency event-study analysis of operation twist and its implications for qe2. *Brookings Papers on Economic Activity*, 2011(1):151–188, 2011.

[90] Eric T Swanson. Measuring the effects of federal reserve forward guidance and asset purchases on financial markets. Technical report, National Bureau of Economic Research, 2017.

[91] Lin Tan, Thomas C Chiang, Joseph R Mason, and Edward Nelling. Herding behavior in chinese stock markets: An examination of a and b shares. *Pacific-Basin Finance Journal*, 16(1-2):61–77, 2008.

[92] John B Taylor. A historical analysis of monetary policy rules. In *Monetary policy rules*, pages 319–348. University of Chicago Press, 1999.

[93] Paul C Tetlock. Giving content to investor sentiment: The role of media in the stock market. *The Journal of finance*, 62(3):1139–1168, 2007.

[94] James Tobin. A general equilibrium approach to monetary theory. *Journal of money, credit and banking*, 1(1):15–29, 1969.

[95] Dimitri Vayanos and Jean-Luc Vila. A preferred-habitat model of the term structure of interest rates. Technical report, National bureau of economic research, 2009.

[96] Neil Wallace. A modigliani-miller theorem for open-market operations. *The American Economic Review*, 71(3):267–274, 1981.

[97] Russ Wermers. Mutual fund herding and the impact on stock prices. *the Journal of Finance*, 54(2):581–622, 1999.

[98] Richard A Werner. Enhanced debt management: Solving the eurozone crisis by linking debt management with fiscal and monetary policy. *Journal of International Money and Finance*, 49:443–469, 2014.

[99] Stephen D Williamson. Liquidity, monetary policy, and the financial crisis: A new monetarist approach. *American Economic Review*, 102(6):2570–2605, 2012.

Appendix A

Appendix

A.0.1 Log-linearized Equations:

$$
\begin{aligned}
\hat{\lambda}_t^{ur} &= \hat{R}_t + E_t(\hat{\lambda}_{t+1}^{ur} - \hat{\pi}_{t+1}) \\
\hat{\lambda}_t^{ur} + \hat{\xi}_t &= \frac{R^L}{R^L - \kappa}\hat{R}_t^L + E_t\left[\hat{\lambda}_{t+1}^{ur} - \hat{\pi}_{t+1} - \frac{\kappa}{R^L - \kappa}\hat{R}_{t+1}^L\right] \\
\hat{\lambda}_t^{r} &= \frac{R^L}{R^L - \kappa}\hat{R}_t^L + E_t\left[\hat{\lambda}_{t+1}^{r} - \hat{\pi}_{t+1} - \frac{\kappa}{R^L - \kappa}\hat{R}_{t+1}^L\right] \\
\hat{\lambda}_t^{ur} &= -\sigma \hat{c}_t^{ur} \\
\hat{\lambda}_t^{r} &= -\sigma \hat{c}_t^{r} \\
\hat{H}_t^j &= [1 - \beta_j \omega_w \pi^{-\theta_n(1+\eta)}][\theta_n(1+\eta)\hat{w}_t + (1+\eta)\hat{\ell}_t] + \beta_j \omega_w \pi^{-\theta_n(1+\eta)} E_t \hat{H}_{t+1}^j, j = ur, r \\
\hat{F}_t^j &= [1 - \beta_j \omega_w \pi^{1-\theta_n}][1 + \hat{\lambda}_t^w + \theta_n \hat{w}_t + \hat{\ell}_t] + \beta_j \omega_w \pi^{1-\theta_n} E_t \hat{F}_{t+1}^j, j = ur, r \\
(\theta_n \eta + 1)\hat{\tilde{w}}_t^j &= \hat{H}_t^j - \hat{F}_t^j, j = ur, r \\
\hat{w}_t &= (1 - \omega_w)(\omega \hat{\tilde{w}}_t^{ur} + (1-\omega)\hat{\tilde{w}}_t^r) + \omega_w(\hat{w}_{t-1} - \hat{\pi}_t)
\end{aligned}
$$

$$\begin{aligned}
\hat{mc}_t &= \alpha\hat{r}_t^k + (1-\alpha)\hat{w}_t - \hat{z}_t \\
\hat{k}_{t-1} &= \hat{w}_t - \hat{r}_t^k + \hat{\ell}_t \\
\hat{y}_t &= \hat{z}_t + \alpha\hat{k}_{t-1} + (1-\alpha)\hat{\ell}_t \\
\hat{H}_t^{p,j} &= (1-\omega_f\beta_j\pi^{-\theta_f})(\hat{\lambda}_t^j + \hat{mc}_t + \hat{y}_t) + \omega_f\beta_j\pi^{-\theta_f}\hat{H}_{t+1}^{p,j}, j = ur, r \\
\hat{F}_t^{p,j} &= (1-\omega_f\beta_j\pi^{1-\theta_f})(\hat{\lambda}_t^j + \hat{y}_t) + \omega_f\beta_j\pi^{1-\theta_f}\hat{F}_{t+1}^{p,j}, j = ur, r \\
\hat{\tilde{p}}_t &= \chi_p\hat{H}_t^{p,ur} + (1-\chi_p)\hat{H}_t^{p,r} - \chi_p\hat{F}_t^{p,ur} - (1-\chi_p)\hat{F}_t^{p,r} \\
\text{where} \quad \chi_p &\equiv \frac{\omega}{\omega + (1-\omega)\frac{\lambda^r}{\lambda^{ur}}} \\
\hat{\pi} &= \frac{1-\omega_f}{\omega_f}(\chi_p\hat{H}_t^{p,ur} + (1-\chi_p)\hat{H}_t^{p,r} - \chi_p\hat{F}_t^{p,ur} - (1-\chi_p)\hat{F}_t^{p,r}) \\
\hat{k}_t &= (1-\delta)\hat{k}_{t-1} + \delta\hat{i}_t \\
\hat{q}_t &= E_t\left[\bar{\beta}\hat{q}_{t+1} + \bar{\beta}r^k\hat{r}_{t+1}^k + \bar{\omega}(\frac{\beta_{ur}}{\bar{\beta}}\hat{\lambda}_{t+1}^{ur} - \hat{\lambda}_t^{ur}) + (1-\bar{\omega})(\frac{\beta_r}{\bar{\beta}}\hat{\lambda}_{t+1}^r - \hat{\lambda}_t^r)\right] \\
\text{where} \quad \bar{\omega} &\equiv \frac{\omega\lambda^{ur}}{\omega\lambda^{ur} + (1-\omega)\lambda^r} \\
\hat{q}_t &= S''(\hat{i}_t - \hat{i}_{t-1}) \\
\hat{z}_t &= \rho_z\hat{z}_{t-1} + \varepsilon_{B,t} \\
\hat{y}_t &= \frac{\omega c^{ur}}{y}\hat{c}_t^{ur} + \frac{(1-\omega)c^r}{y}\hat{c}_t^r + \frac{i}{y}\hat{i}_t + \frac{g}{y}\hat{g}_t \\
-\frac{R^L}{R^L-\kappa}\hat{R}_t^L + \hat{b}_t^L &= -\frac{R^L}{R^L-\kappa}\rho_B\hat{R}_{t-1}^L + \rho_B\hat{b}_{t-1}^L + \phi_{B,y}(1-\rho_B)\hat{y}_t + \varepsilon_{B,t} \\
\hat{b}_t^s + \frac{b^L/b^s}{R^L-\kappa}\hat{b}_t^L &= \frac{1}{\beta_{ur}}(\hat{R}_{t-1} - \hat{b}_{t-1}^s) + \frac{b^L/b^s}{R^L-\kappa}\frac{1}{\beta_r}\hat{b}_{t-1}^L - \left(\frac{1}{\beta_{ur}} + \frac{b^L/b^s}{R^L-\kappa}\frac{1}{\beta_r}\right)\hat{\pi}_t \\
&\quad + \frac{(1-\kappa\pi^{-1}R^L)}{R^L-\kappa}\frac{b^L/b^s}{R^L-\kappa}\hat{R}_t^L + \frac{g}{b^s}\hat{g}_t \\
\hat{R}_t &= \rho_m\hat{R}_{t-1} + (1-\rho_m)[\phi_\pi\hat{\pi} + \phi_y(\hat{y}_t - \hat{y}_{t-4})] + \varepsilon_{m,t}
\end{aligned}$$

$$\hat{\xi}_t = \rho_\xi(\hat{b}^L - \frac{R^L}{R^L - \kappa}\hat{R}_t^L)$$

$$\hat{g}_t = \Psi_g \hat{y}_t + \nu_g \hat{b_{t-1}}$$

$$\hat{tr}_t = \Psi_{TR} \hat{y}_t + \nu_{TR} \hat{b_{t-1}}$$

A.0.2 Tables:

Table A.1: Tests for Herding (Buy Side) on Other Types – Include SH-HK

	Institutional Buy					
Dependent Variable	(1) FUND	(2) QFII	(3) GOV.	(4) INSUR.	(5) OTHER	(6) SH-HK
lag FUND	-0.4619 ***	0.0883	0.0740	0.1700	-0.2668	-0.1140
	(0.1152)	(0.2271)	(0.1523)	(0.1616)	(0.1615)	(0.1335)
Lag QFII	-0.0530	0.0450	-0.0083	-0.0522	0.0934	-0.0648
	(0.0331)	(0.1288)	(0.1036)	(0.0916)	(0.0751)	(0.0754)
Lag GOV.	-0.0369	-0.0871	0.1052	0.3874 ***	-0.0692	0.1387 *
	(0.0424)	(0.0844)	(0.1216)	(0.0504)	(0.0797)	(0.0738)
Lag INSUR.	-0.0308	0.1708 **	0.0707	0.0742	-0.0798	0.0142
	(0.0388)	(0.0661)	(0.0656)	(0.0940)	(0.0687)	(0.0504)
Lag OTHER	0.0557 *	-0.0099	0.0151	0.1350	-0.0214	0.1096 ***
	(0.0285)	(0.0711)	(0.1333)	(0.1589)	(0.0849)	(0.0362)
Lag SH-HK	-0.0206	0.0154	-0.0219	0.1222	0.0363	0.3503 **
	(0.0264)	(0.0758)	(0.1247)	(0.0863)	(0.1383)	(0.1514)
EPS	-0.0014	-0.0515 ***	-0.0147	-0.0292	-0.0368 ***	0.0195
	(0.0111)	(0.0184)	(0.0185)	(0.1300)	(0.0086)	(0.0130)
P/E ratio	1.2261 *	0.0831	-0.1218	0.6186	-0.0751	-2.5924 ***
	(0.6908)	(1.0975)	(1.0972)	(1.2982)	(1.2111)	(0.8074)
ROE	0.3422 ***	0.9400 ***	0.0020	-0.0465	0.2230 *	0.0770
	(0.0729)	(0.3046)	(0.1119)	(0.1404)	(0.1316)	(0.1288)
A/L Ratio	0.0076	-0.1835	0.0394	0.2582 ***	-0.1264 **	0.1301 **
	(0.0338)	(0.1579)	(0.1561)	(0.0955)	(0.0331)	(0.0531)
Observations	125	101	120	94	112	126
Time Effect	YES	YES	YES	YES	YES	YES
Std. Err. Clustered	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH
R-Squared	0.3852	0.1566	0.0359	0.2142	0.1004	0.2993

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

Table A.2: Tests for Herding (Sell Side) on Other Types – Include SH-HK

	Institutional Sell					
Dependent Variable	(1) FUND	(2) QFII	(3) GOV.	(4) INSUR.	(5) OTHER	(6) SH-HK
lag FUND	-0.4719 ***	-0.0033	0.0224	-0.0768	-0.0417	-0.2204
	(0.1494)	(0.1502)	(0.2306)	(0.2576)	(0.1990)	(0.2025)
Lag QFII	0.0342	0.2742	0.0026	-0.1026	0.1201 *	0.0138
	(0.0213)	(0.1830)	(0.0095)	(0.1193)	(0.0722)	(0.0474)
Lag GOV.	-0.0076	0.0505	0.1112	0.1759	-0.0846 ***	-0.0243
	(0.0464)	(0.1025)	(0.0795)	(0.1429)	(0.0247)	(0.0722)
Lag INSUR.	-0.0067	-0.0217	0.0615	(0.0825)	-0.0374	-0.0208
	(0.0416)	(0.1632)	(0.0915)	(0.1060)	(0.0609)	(0.0463)
Lag OTHER	0.1198 ***	0.0591	-0.0588	0.3829 **	0.0049	0.1098
	(0.0393)	(0.1278)	(0.1357)	(0.1557)	(0.0685)	(0.0645)
Lag SH-HK	0.0332	0.0914	0.0351	0.2603 **	0.0307	0.2290 ***
	(0.0572)	(0.1280)	(0.1928)	(0.1266)	(0.0514)	(0.0831)
EPS	-0.0128	0.0542	0.0124	0.1210	0.0200	0.0235
	(0.0119)	(0.0428)	(0.0227)	(0.1491)		(0.0177)
P/E ratio	-1.0773 *	-2.1950 *	-2.3610 **	0.5321	-4.5709 ***	-0.5588
	(0.6323)	(1.2856)	(1.1541)	(1.2407)	(0.8555)	(1.5302)
ROE	-0.3136 ***	-1.1108 *	(0.0104)	-0.5282 ***	-0.0680	-0.0816
	(0.0451)	(0.5692)	(0.2875)	(0.1630)	(0.0742)	(0.1005)
A/L Ratio	0.0709	0.0959	0.0742	-0.4471 ***	0.1629 **	0.0127
	(0.0447)	(0.2893)	(0.3009)	(0.1559)	(0.0798)	(0.0868)
Observations	125	101	120	94	112	126
Time Effect	YES	YES	YES	YES	YES	YES
Std. Err. Clustered	BOTH	BOTH	BOTH	BOTH	BOTH	BOTH
R-Squared	0.3933	0.1453	0.0517	0.2252	0.1961	0.0908

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

Table A.3: Tests for Herding (Buy Side) on Other Types – Exclude SH-HK

	Institutional Buy									
Dependent Variable	(1) FUND		(2) QFII		(3) GOV.		(4) INSUR.		(5) OTHER	
lag FUND	-0.4639	***	0.0445		0.0684		0.0599		0.0002	
	(0.0910)		(0.0953)		(0.0575)		(0.0711)		(0.0619)	
Lag QFII	-0.0119		0.1275	*	-0.0275		0.0525		0.0411	
	(0.0140)		(0.0664)		(0.0536)		(0.0438)		(0.0371)	
Lag GOV.	0.0276		-0.0262		0.2510	***	0.1266	**	-0.0019	
	(0.0233)		(0.0512)		(0.0529)		(0.0618)		(0.0350)	
Lag INSUR.	-0.0172		0.0417		0.0723	*	0.2134	***	0.0022	
	(0.0142)		(0.0430)		(0.0427)		(0.0465)		(0.0363)	
Lag OTHER	0.0558	**	0.0561		-0.0855		0.0275		-0.0567	
	(0.0218)		(0.0409)		(0.0581)		(0.0730)		(0.0456)	
EPS	0.0008		-0.0168		-0.0140	*	-0.0302	**	-0.0098	*
	(0.0095)		(0.0109)		(0.0081)		(0.0129)		(0.0051)	
P/E ratio	0.0953		-0.3406		(0.1227)		0.5662		0.0675	
	(0.0858)		(1.1283)		(0.2285)		(0.3802)		(0.2156)	
ROE	0.0720		0.3848	***	0.1140		0.2028		0.1192	
	(0.0659)		(0.1376)		(0.0886)		(0.1480)		(0.0792)	
A/L Ratio	-0.0204		0.0255		0.0516		-0.0608		-0.0491	
	(0.0389)		(0.0970)		(0.0606)		(0.0766)		(0.0413)	
Observations	498		377		440		400		441	
Time Effect	YES		YES		YES		YES		YES	
Std. Err. Clustered	BOTH		BOTH		BOTH		BOTH		BOTH	
R-Squared	0.2910		0.0508		0.0997		0.1074		0.0130	

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

Table A.4: Tests for Herding (Sell Side) on Other Types – Exclude SH-HK

	Institutional Sell									
Dependent Variable	(1) FUND		(2) QFII		(3) GOV.		(4) INSUR.		(5) OTHER	
lag FUND	-0.4765	***	0.1297		0.0463		0.0861		0.0385	
	(0.0904)		(0.1036)		(0.0882)		(0.0960)		(0.0606)	
Lag QFII	0.0211		0.1731	**	-0.0138		0.0519		0.0254	
	(0.0165)		(0.0804)		(0.0332)		(0.0484)		(0.0394)	
Lag GOV.	0.0373	*	-0.0282		-0.1448	***	0.0607		0.0042	
	(0.0214)		(0.0572)		(0.0530)		(0.0636)		(0.0272)	
Lag INSUR.	-0.0074		-0.0183		0.0337		0.0779		-0.0575	
	(0.0142)		(0.0583)		(0.0362)		(0.0495)		(0.0354)	
Lag OTHER	0.0764	***	0.0609		-0.1086	*	0.0852		-0.1057	***
	(0.0168)		(0.0601)		(0.0610)		(0.0779)		(0.0423)	
EPS	-0.0108		0.0241	*	0.0138	**	0.0542	***	0.0113	**
	(0.0095)		(0.0130)		(0.0062)		(0.0070)		(0.0060)	
P/E ratio	-0.0920		-2.0815	**	(0.9072)		-1.1850	***	-0.3286	
	(0.0850)		(0.8862)		(0.6980)		(0.4079)		(0.2546)	
ROE	(0.1204)	**	-0.2198		0.0888		0.0214		0.0103	
	(0.0494)		(0.1571)		(0.1272)		(0.0739)		(0.0859)	
A/L Ratio	0.0525		0.0438		(0.0458)		-0.0068		-0.0105	
	(0.0335)		(0.1444)		(0.0968)		(0.0999)		(0.0433)	
Observations	498		377		440		400		441	
Time Effect	YES		YES		YES		YES		YES	
Std. Err. Clustered	BOTH		BOTH		BOTH		BOTH		BOTH	
R-Squared	0.3183		0.0502		0.0443		0.0408		0.0233	

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

Table A.5: Tests for Herding by Firm Size – Buy Side – Own Type

Dependent Variable	(1) FUND		(2) QFII		(3) GOV.		(4) INSUR.		(5) OTHER		(6) SH-HK	
Panel A: Firm size <= 20% percentile												
Lag Buy by Own Type	-0.1076	***	0.2178	***	0.2592	***	0.1976	***	0.1941	***	0.4071	***
	(0.0247)		(0.0602)		(0.0517)		(0.0446)		(0.0201)		(0.0421)	
Observations	3,834		312		495		579		2,827		351	
R-Squared	0.0360		0.0960		0.1306		0.0635		0.0740		0.3187	
Panel B: Firm size <= 40% percentile												
Lag Buy by Own Type	-0.1303	***	0.1077	***	0.2256	***	0.1670	***	0.1596	***	0.4166	***
	(0.0290)		(0.0381)		(0.0330)		(0.0341)		(0.0166)		(0.0867)	
Observations	5,566		500		1,054		992		3,137		707	
R-Squared	0.0265		0.0423		0.1067		0.0513		0.0443		0.2305	
Panel C: Firm size <= 60% percentile												
Lag Buy by Own Type	-0.1918	***	0.1794	***	0.2534	***	0.1577	***	0.0778	***	0.4593	***
	(0.0231)		(0.0371)		(0.0285)		(0.0170)		(0.0148)		(0.0579)	
Observations	6,671		633		1,906		1,507		3,551		1,230	
R-Squared	0.0450		0.0586		0.0949		0.0390		0.0172		0.2674	
Panel D: Firm size <= 80% percentile												
Lag Buy by Own Type	-0.2566	***	0.1312	***	0.2297	***	0.1586	***	0.0687	***	0.4938	***
	(0.0281)		(0.0399)		(0.0243)		(0.0249)		(0.0160)		(0.0471)	
Observations	7,541		686		3,300		1,913		4,375		1,822	
R-Squared	0.0817		0.0385		0.0741		0.0370		0.0111		0.2679	
Panel E: Firm size > 80% percentile												
Lag Buy by Own Type	-0.5110	***	0.1788	***	0.1628	***	0.2006	***	0.0062		0.3824	***
	(0.0596)		(0.0404)		(0.0309)		(0.0201)		(0.0203)		(0.0393)	
Observations	8,230		1,151		4,808		2,564		5,942		2,489	
R-Squared	0.2765		0.0481		0.0423		0.0595		0.0070		0.1968	

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

Table A.6: Tests for Herding by Firm Size – Sell Side – Own Type

Dependent Variable	(1) FUND		(2) QFII		(3) GOV.		(4) INSUR.		(5) OTHER		(6) SH-HK	
Panel A: Firm size <= 20% percentile												
Lag Sell by Own Type	-0.1931	***	0.1427		0.1240	*	0.1225	**	0.1230	***	0.1607	**
	(0.0240)		(0.0873)		(0.0632)		(0.0610)		(0.0265)		(0.0624)	
Observations	3,834		312		495		579		2,827		351	
R-Squared	0.0512		0.0260		0.0242		0.0245		0.0274		0.0531	
Panel B: Firm size <= 40% percentile												
Lag Sell by Own Type	-0.2027	***	0.1431	**	0.2268	***	0.0680		0.0994	***	0.3373	***
	(0.0249)		(0.0625)		(0.0737)		(0.0413)		(0.0232)		(0.0415)	
Observations	5,566		500		1,054		992		3,137		707	
R-Squared	0.0472		0.0226		0.0526		0.0050		0.0133		0.1223	
Panel C: Firm size <= 60% percentile												
Lag Sell by Own Type	-0.2549	***	0.2453	***	0.1566	***	0.1086	***	0.0306		0.3323	***
	(0.0281)		(0.0581)		(0.0475)		(0.0310)		(0.0224)		(0.0367)	
Observations	6,671		633		1,906		1,507		3,551		1,230	
R-Squared	0.0687		0.0457		0.0258		0.0166		0.0030		0.1238	
Panel D: Firm size <= 80% percentile												
Lag Sell by Own Type	-0.2923	***	0.1137	**	0.1409	***	0.0799	***	-0.0062		0.2570	***
	(0.0321)		(0.0548)		(0.0287)		(0.0256)		(0.0200)		(0.0339)	
Observations	7,541		686		3,300		1,913		4,375		1,822	
R-Squared	0.0984		0.0131		0.0239		0.0056		0.0050		0.0720	
Panel E: Firm size > 80% percentile												
Lag Sell by Own Type	-0.5385	***	0.2434	***	0.1275	***	0.1423	***	-0.0346	*	0.1770	***
	(0.0617)		(0.0462)		(0.0298)		(0.0265)		(0.0202)		(0.0335)	
Observations	8,230		1,151		4,808		2,564		5,942		2,489	
R-Squared	0.3031		0.0642		0.0268		0.0271		0.0046		0.0527	

***,**, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

Table A.7: Parameter Value Used

Parameter	Value	Description
ω	0.7	Fraction of consumers that are workers
η	2	Labor supply elasticity
β_w	0.9975	Discount rate for workers
ρ_m	0.7	Taylor rule coef. on previous interest rate
ϕ_π	1.75	Taylor rule coef. on inflation
ϕ_y	0.4	Taylor rule coef. on output growth
ρ_{qe}	0.8	QE Policy persistence coef. on previous bonds
ρ_b	0.7	Supply Policy persistence coef. on previous bonds
ρ_z	0.6	Productivity persistence coef.
ω_f	0.5	Probability of resetting prices
ω_w	0.5	Probability of resetting wages
θ_f	7.667	Price markup coef. (Price markup = 1.15)
θ_n	7.667	Wage markup coef. (Wage markup = 1.15)

Table A.8: Scatter Plots Slope Coefficients

	$G_{Fedholding}$before	$G_{Fedholding}$within	$G_{Fedholding}$after	$G_{Fedholding}$BS
G_{debt}within	0.337*			
	(0.176)			
G_{debt}within		0.588*		
		(0.303)		
G_{debt}within			0.053	
			(0.061)	
G_{debt}within				−1.114
				(0.876)
Observations	69	74	32	19

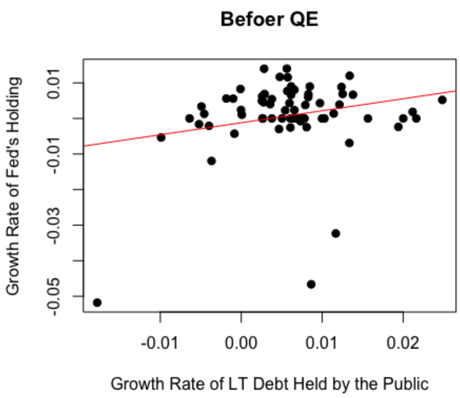
*p<0.1; **p<0.05; ***p<0.01

Note: $G_{Fedholding}$ and G_{debt} denote for the growth rate of Fed's holding and LT debt outstanding, which are obtained by log-differentiating Fed's holding of LT Treasury debt and LT debt outstanding. The four sample periods are denoted as before (200301 – 200810), within (200811–201412), after (201501–201708) and B/S reduction (201709–201903).

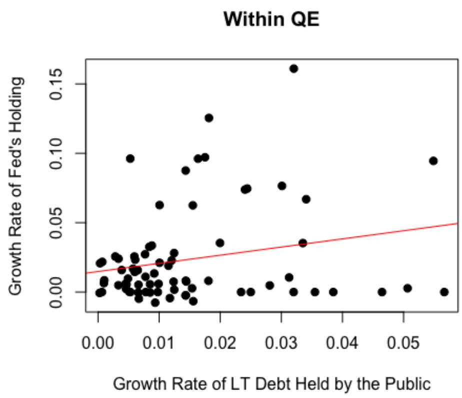
A.0.3 Figures:

Figure A.1: Scatter Plots for Growth Rates between Fed's Holding and Longer-term Treasury Debts Held by the Public

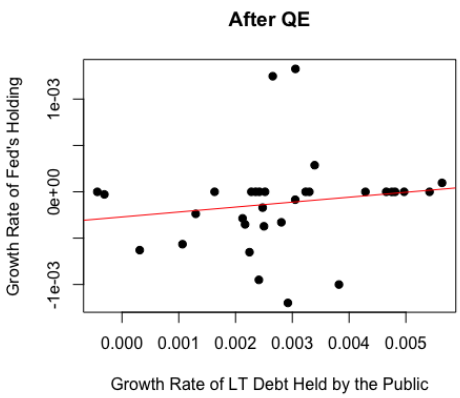
(a) Scatter Plot before QE



(b) Scatter Plot within QE



(c) Scatter Plot after QE



(d) Scatter Plot in Balance Sheet Reduction

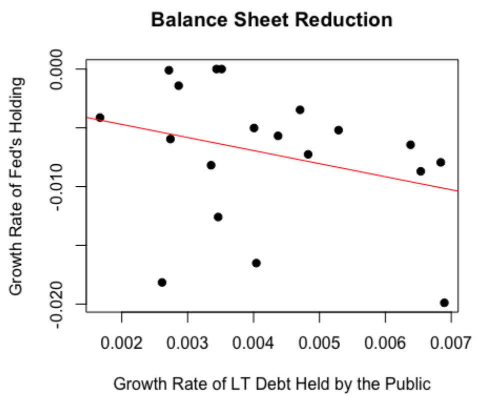
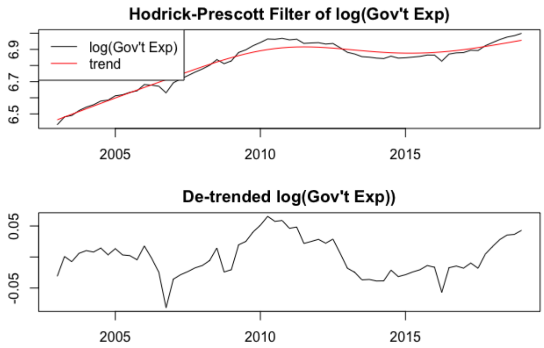
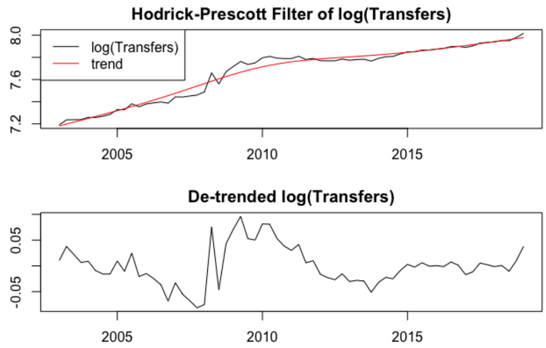


Figure A.2: Government Expenditure, Transfer, Tax Revenue, Real GDP from 2003Q1 to 2019Q1

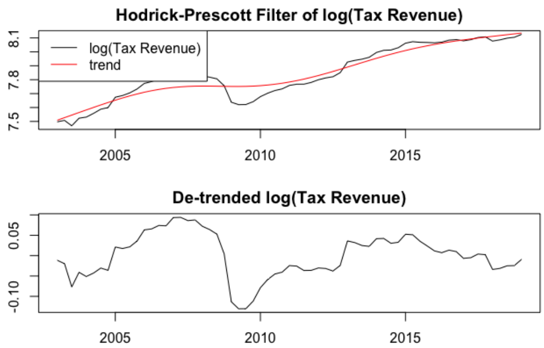
(a) Government Expenditure



(b) Transfer



(c) Tax Revenue



(d) Real GDP