ESSAYS ON EMPIRICAL FINANCE

A Dissertation
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

by
Wei Gao
May 2021

Examining Committee Members:

Gurdip Bakshi, Advisory Chair, Department of Finance
Xiaohui Gao Bakshi, Department of Finance
Yuanzhi Li, Department of Finance
Steven Balsam, Department of Accounting
Charlotte Ren, External Member, Department of Strategic Management

ABSTRACT

This dissertation has two chapters. Each empirically examines one finance topic. The first chapter focuses on behavioral finance. The second chapter focuses on corporate finance.

The first chapter is motivated by inconclusive theoretical prediction and lack of empirical evidence on the effect of mood on trading volume. This chapter exploits repeated natural experiments from the occurrence of severe smog in Beijing to test the inertia hypothesis (Bad moods cause inactivity and inertia and thus decrease in trading volume) against the mood regulation hypothesis (Bad moods increase trading volume because investors use trading as a way to combat bad moods). Intra-day analysis in this chapter shows that smog in Beijing causes trading volume of stocks headquartered in Beijing to increase, which contradicts the inertia hypothesis. The effect is more pronounced among large stocks, which rules out the possibility that investors seek gambling thrill during smog. Additionally, the effect is more pronounced among low risk stocks, which reflects the risk aversion associated with depression among investors and supports the mood regulation theory.

The second chapter is motivated by the fact that initial public offerings (IPOs) transform private firms into publicly traded ones, thereby improving liquidity of their shares. Better liquidity increases firm value, which I call "liquidity value". I develop a model and hypothesize that issuers and IPO investors bargain over the liquidity value, resulting in a discounted offer price, i.e., IPO underpricing. Consistent with the model, I find that underpricing is positively related to the expected post-IPO liquidity of the issuer.

The relation is stronger when firms are financed by venture capital investors, when the underwriter has more bargaining power, or when a smaller fraction of the firm is sold. I also explore two regulation changes as exogenous shocks to issuers' liquidity before and after IPO, respectively. With a difference-in-difference approach, I find that underpricing is more pronounced with better expected post-IPO liquidity or lower pre-IPO liquidity.

To

my beloved parents, Zhonghua Gao and Zhuqing Deng

for your endless love.

And to

my amazing sister, Ping Gao,

for your faith in me.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor and dissertation committee chair, Dr. Gurdip Bakshi and Dr. Xiaohui Gao Bakshi, for guiding me professionally and personally over the past three years. I have been extremely lucky to have thorough and unrelenting advisors who care so much about my work, and gives me advice, and support all the time.

I would like to thank the rest of my dissertation committee members for their great support and invaluable advice. I am grateful to Dr. Yuanzhi Li for teaching me a great deal about both research and life in general. I am grateful to Dr. Steven Balsam, for his remarks that help improve my dissertation. I am grateful to Dr. Charlotte Ren, for kindly agreeing to serve on my committee, and for bringing in expertise and insights as my external examiner. I am thankful to Dr. Ronald Anderson, Dr. Jonathan Scott, and Dr. Connie Mao for providing financial support.

I am grateful to all other faculty, staff, and doctoral students in the finance departments. They deserve many thanks for their support, kindness, and encouragement along the way.

TABLE OF CONTENTS

Page

ABSTRACT ii

DEDICATION iv

ACKNOWLEDGMENTS v

LIST OF TABLES ix

LIST OF FIGURES xi

CHAPTER

1. DO INVESTORS ENGAGE IN MOOD REGULATION? USING SMOG OCCURRENCES AS REPEATED NATURAL EXPERIMENTS 1

    Introduction 1

    Literature Review 5

    Hypothesis Development 10

    Data and Sample 13

        PM2.5 and Health Risk 13

        Hourly Measure of Smog: PM2.5 Concentration Value 14

        High Frequency Five-Minute Trading Volume 14

        Other Variables 15

        Constituents of Portfolio 15

        Sample 16

        Initial Graphical Evidence 22

    Empirical Results 26

An Exogenous Discrete Smog Variable 26

Alternative Cutoff Values of Smog Dummy 28

Continuous Smog Measure 30

Lagged Smog Variable 33

Smog Positively Impact Trading Volume at the Firm Level 35

Alternative Explanations 37

Evidence against Gambling Hypothesis 38

Evidence Supporting Mood Regulation Hypothesis......................40

Sort Stocks by Trading Activity 42

Conclusions 44

2. LIQUIDITY VALUE AND IPO UNDERPRICING 46

Introduction 46

Relation to Previous Work 55

Hypothesis Development 60

Sample and Variable Construction 61

Empirical Tests and Results 71

The Baseline...72

Cross-sectional Analysis of the Baseline 81

Changes in the Order Handling Rules at Nasdaq...............................85

The National Securities Market Improvement Act (NSMIA) 101

Discussion: Changes to OHR on Nasdaq and the Passage of NSMIA 120

Conclusion 123

REFERENCES CITED..125

APPENDICES

A. CHAPTER 1 VARIABLE DEFINITION 136
B. CHAPTER 1 AIR QUALITY INDEX (AQI) AND PM2.5 VALUE 137
C. CHAPTER 2 VARIABLE DEFINITION 138

LIST OF TABLES

Table Page

1.1. Beijing Smog Distribution by Air Quality Category 18

1.2. Summary Statistics 21

1.3. Impact of Smog Dummy on Trading Volume 27

1.4. Impact of Smog Level on Trading Volume 32

1.5. Impact of Smog Level in Previous Hours on Trading Volume 34

1.6. Impact of Smog on Trading Volume, Firm-Level Evidence 36

1.7. Sort Firms by Firm Size 39

1.8. Sort Firms by Idiosyncratic Volatility 41

1.9. Sort Firms by Trading Activity 43

2.1. Distribution of IPOs by Year, Industry, and Exchange 62

2.2. Summary Statistics 69

2.3. Baseline: IPO Underpricing and Expected Liquidity 73

2.4. Robustness Analysis of the Baseline Regression 79

2.5. Cross-Sectional Analysis of the Baseline 82

2.6. IPO Underpricing and Changes to Order Handling Rules at Nasdaq in 1997 90

2.7. Placebo Tests in the Pre-OHR Period 94

2.8. Individual Peer Liquidity and Order Handling Rules at Nasdaq In 1997 96

2.9. Distribution of VC and PE Firms across States, 1993-2000 105

2.10. IPO Underpricing and National Security Market Improvement Act: Diff-in-Diff 108

2.11. Placebo Tests in the Pre-NSMIA Period 115

2.12. IPO Underpricing and the National Security Market Improvement Act 117

2.13. Changes to OHR On Nasdaq and the Passage of NSMIA 122

LIST OF FIGURES

Figure Page

1.1. Beijing Hourly Smog 17

1.2. Portfolio of Beijing Firms Hourly Trading Volume 19

1.3. Portfolio of Beijing Firms Average Hourly Trading Volume 25

1.4. Smog Dummy with Alternative Cutoff and Trading Volume 29

2.1. Average Underpricing across Quintile Subsamples 50

2.2. IPO Underpricing at Nasdaq and non-Nasdaq Exchanges 89

2.3. IPO Underpricing in States with and without Large Numbers of VC and PE 106

CHAPTER 1

DO INVESTORS ENGAGE IN MOOD REGULATION? USING SMOG OCCURRENCES AS REPEATED NATURAL EXPERIMENTS

Introduction

Under a behavioral theory, there are no clear predictions on the effect of mood on trading volume (Edmans, Garcia, & Norli, 2007). On one hand, one might expect that a bad mood causes inactivity and inertia in traders. Therefore, bad moods cause trading volume to decline. On the other hand, psychology studies suggest human beings are capable of taking actions to fix their mood and engage in mood regulation. Performing challenging tasks is believed to be one way to successfully overcome sad moods (Erber & Tesser, 1992). Trading appears to meet the criterion of a challenging task: Not only is it a mentally intense activity, but it also creates opportunities of monetary reward to offset the negative mood (Edmans et al., 2007). Therefore, trading volume could increase when investors are in bad moods and use trading to regulate their mood.

Due to inconclusive theories on how mood affects trading volume, the role of empirical evidence becomes important. Unfortunately, empirical evidence on this topic is limited. In an attempt to close the gap between theory and empirical evidence, this paper exploits repeated experiments, namely, the frequent occurrence of severe smog in Beijing, and uses intra-day trading data to show that severe smog in Beijing increases the trading volume of stocks headquartered in Beijing, which provides evidence that investors use trading to combat bad moods.

Smog in Beijing is a good mood variable for three reasons. First, neuroscience and medical studies show that smog does affect moods. Fonken et al. (2011) simulate the air pollution exposure a typical worker experiences and treated mice with high levels of fine particulate (main component of smog) eight hours a day, five times a week. Ten months later, they find that mice appeared less interested in sugar water than before, which suggests depression in the treated mice. Secondly, smog is plausibly exogenous. The direction of causality apparently goes from smog to trading volume. It is unlikely that trading volume causes smog to change. Thirdly, the high frequency and severity of smog in Beijing creates strong and statistically detectable mood effects. To measure smog level, I use hourly smog data (PM2.5 concentration value) from US embassy in Beijing, which is high quality and rarely used in the financial studies.

To match the hourly frequency of smog data, I exploit five-minute high frequency trading data, from which I construct hourly trading volume. I then form an equally weighted portfolio of all A share firms listed in Shanghai Stock Exchange and headquartered in Beijing. Shanghai Stock Exchange is the largest stock exchange in China. Unlike B shares, A share stocks are mostly domestically traded. Limiting stocks to those headquarter in Beijing is because investors in Beijing are more exposed to local media coverage of Beijing firms, thus are more familiar with Beijing stocks and hold disproportionately large amount of Beijing stocks (Coval & Moskowitz, 1999; Huberman, 2001; Feng & Seasholes, 2004; Pirinsky & Wang, 2006; Seasholes & Zhu, 2010). To the extent that smog in Beijing affects mood of Beijing investor, stocks headquartered in Beijing are more likely be affected by local smog than stocks headquartered in other cities.

I first use an exogenously decided severe smog threshold to create a severe smog dummy variable. The threshold comes from US Environment Protection Agency (EPA) health guide therefore it is exogenously decided. I find the dummy variable is positively related to trading volumes. To address the concern that the threshold might be arbitrary, I increase or decrease the threshold by small steps and find similar results. I then test whether smog level in an hour has a contemporaneous impact on trading volume. I find that as smog level in Beijing in an hour increases, trading volume of stocks headquartered in Beijing during that hour increases as well.

In my analysis, I also examine the effect firm by firm. Firm level results are consistent with portfolio level results. There could be some lagged effect because as smog increases, it might take one or two hours to change people's mood. Therefore I test whether smog level in previous hours have an impact on trading. I find that smog two or three hours ago do have a statistically significant positive impact on trading volume.

The finding that smog increases trading volume does not support the inertia theory which posits that bad mood induces inactivity and inertia, thus decline in trading volume. It seems to support the mood regulation theory which argues that trading fulfils investors' need to combat bad mood. However, the positive relation between smog and trading can also be explained by the entertainment (D. Dorn & Sengmueller, 2009) and gambling theory (A. Dorn, Dorn, & Sengmueller, 2014; Gao & Lin, 2014) which hypothesizes that investors have limited outdoor entertainments in heavy smog thus seek entertainment and thrill from trading indoors. To rule out this possibility, I sort the stocks by firm size and find that the smog effect is more pronounced among large stocks. However, according gambling theory, the smog effect should be stronger among small stocks because they are lottery-like stocks. Thus, the empirical finding seems to contradict the gambling theory.

If investors use trading to combat sad moods or depression, they should have symptoms of depression. They should also appear to be risk averse because studies show that depression also leads to risk aversion (Kamstra, Kramer, & Levi, 2003). Therefore, if the mood regulation theory is correct, the smog effect should be more pronounced among low risk stocks. I sort the stocks by their idiosyncratic risk and find that smog effect indeed is more pronounced among low idiosyncratic risks, which further supports that mood regulation theory, rather than gambling theory, explains the finding.

I finally sort stocks by how actively they are traded. I find that the effect is more pronounced among actively traded stocks. To the extent that active investors and passive investors prefer and hold different types of stocks, this indicates passive investors remain passive even if they experience mood shift whereas active investors become more active after they experience bad moods.

The contribution of this paper is three folds. First, the theoretical prediction on the relation between mood and trading volume is ambiguous and the supportive empirical findings are scarce. The empirical findings in this paper closes the gap between theory and empirical works. Second, this is the first study that uses intra-day data to examine the impact of smog on stock market in China. Intra-day data provides a more accurate estimation of the effect of mood on stock behaviors (Barclay & Litzenberger, 1988; Busse

& Green, 2002). Sources of variation attributed to other unrelated factors are significantly mitigated by much shorter measurement period, which in return mitigates the possibility that results are driven by outliers and subsamples (Pinegar, 2002; Chang, Chen, Chou, & Lin, 2008). Thirdly, this study uses hourly mood variable to document investor's mood regulation behavior.

The remainder of the paper is organized as follows. Next section conducts a literature review. Then the following section presents the hypothesis development. Following section describes the data and sample in the study. Then the next section describes empirical tests and presents main findings. The last section concludes.

Literature Review

Medical studies have long shown that smog is responsible for respiratory and cardiovascular diseases. Recent neuroscience studies even suggest that smog has a negative impact on brain, which is responsible for intelligence and mood. Particularly, long-term exposure to smog impairs cognition and provokes depressive-like behaviors. For instance, Suglia, Gryparis, Schwartz, and Wright (2007) followed over 200 children in East Boston over the course of 10 years. They found that kids who breathed high levels of black carbon performed worse on tests of memory and IQ.

Fonken et al. (2011) tried to simulate the air pollution exposure a typical worker experiences and treated mice with high levels of fine particulate eight hours a day, five times a week. Ten months later, they discovered physical changes to the nerve cells in mouse brain. In addition, they found that it took treated mice much longer to learn a maze task than it did their untreated peers, which is the evidence that smog damaged the cognition of the mouse. They also find that the treated mice appeared less interested in sugar water than before, which is the evidence that smog provoked depressive-like behaviors in mice.

However, the impact of smog on investor behavior and the financial markets has only been examined by two studies. Hu, Li, and Lin (2014) find no evidence that smog in a stock exchange city in China, Shanghai or Guangzhou, affects stock exchange return. However, they find that air quality in a stock exchange city relative to that in Beijing negatively affects stock prices and trading volume. Li and Peng (2016) find a negative relation between air quality in an exchange city in China and return of the stock exchange index in the subsample period 2010-2014 and a positive relation between the two in the subsample period 2005-2009; that is, the sign reverses. However, in the whole sample period 2005-2014, they find no statistically significant relation.

This paper is related to the studies suggesting stock trading is an entertaining and recreational activity for people who invest in the stock market. Grinblatt and Keloharju (2009) show Finnish investors who seek sensations also trade more often. D. Dorn and Sengmueller (2009) document German investors who enjoy investing also rebalance their portfolios more often. Some researchers go even further and believe that stock trading is as exciting and thrilling as gambling. Barber, Lee, Liu, and Odean (2008) argue that investors trade because they have the desire to gamble. Kumar (2009) reveal that state lottery buyers and lottery-type stock holders share many commonalities. A. Dorn et al. (2014) show that trading activities decline when larger lottery prizes surface. Gao and Lin

(2014) exploit the repeated random occurrences of large jackpot in Taiwan to show that stock trading and gambling are substitute good for each other.

This paper is also related to the earlier literature on weather, mood and financial markets. Saunders (1993) finds that local weather in New York City affects stock prices. Combining evidence that weather affects mood from psychology studies, he argues that security markets are not rational and weather induced mood variation is an important factor of asset prices. Hirshleifer and Shumway (2003) extends Saunders (1993) by examining the question in an international context. Adopting a panel regression approach, they show that the findings in Saunders (1993) also hold in most of the 26 countries in their study. In addition to these studies on mood of trading professionals and asset prices, other studies examine whether mood of investor in general has an impact on the security markets. For instance, Loughran and Schultz (2004) argues that local weather affects the mood of local market participants, since they tend to hold and trade local stocks (Coval & Moskowitz, 1999; Huberman, 2001; Feng & Seasholes, 2004; Pirinsky & Wang, 2006; Seasholes & Zhu, 2010), local weather may affect stocks headquartered in the local area. Loughran and Schultz (2004) find evidence that trading in Nasdaq stocks is localized, but find little evidence that weather in a city affects stock return of firms headquartered in that city.

Building on the findings from psychology literature, which suggests a link between depression an individual suffers and his or her lowered risk-taking behavior, a small, but growing, body of behavioral finance literature believe depression among investors may affect investment decisions and the financial markets. Kuhnen and Knutson (2011) find that people experiencing positive emotions such as excitement are more risk-taking and more confident in their ability to evaluate investment opportunities, while people experiencing negative emotions such as anxiety appear to be the opposite. Kamstra et al. (2003) study how the seasonal depression induced by the seasonal change in the number of daylight hours affects the financial markets. They find that holding-period stock return are higher in the six months of fall and winter when many people suffer from seasonal depression.

Further along this line, Kramer and Weber (2012) find that people who suffer from seasonal affective disorder (SAD) tend to be more financially risk averse. Kamstra, Kramer, and Levi (2015) document an annual cycle in the US Treasury market and show that the seasonal pattern in Treasury returns is driven by seasonal variation in investor risk aversion. Taking a different perspective, Kamstra, Kramer, Levi, and Wermers (2017) examine aggregate investor flow data in Canada and Australia, where seasons are offset by 6 months. They find that investors prefer safe mutual funds in autumn and risky funds in spring, which is consistent with the variation of investor risk-aversion induced by the change of season.

This paper also speaks to the literature on investor's attention and trading outcomes. Payzan-LeNestour and Woodford (2018), build on the neural science finding that the representational capacities of the brain are indeed limited and neural activity is only allocated to the most probable events. They find support for their conjecture that humans pay very little attention to outliers and cannot distinguish one extreme value from another. Some financial economists argue that the assumption investors can costlessly obtain and process information does not hold for most investors. Because information production and processing requires substantial amount of time and efforts, investors can only allocate limited attention to their investments and thus inattention of investors is a rational choice (Bacchetta & Van Wincoop, 2005; Peng & Xiong, 2006; Hong, Torous, & Valkanov, 2007; Huang & Liu, 2007).

Along the same line, two papers use individual level data of attention to test the theories of investor attention. The studies of Karlsson, Loewenstein, and Seppi (2009) and Sicherman, Loewenstein, Seppi, and Utkus (2015) use 401K accounts logins as a proxy for attention and find that attention paid to investment portfolio goes down after stock market declines. Gargano and Rossi (2018) use a brokerage account data set to examine how investors allocate their attention, and the relation between investor attention and performance. They find that investment outcome is better when investors pay more attention.

How smog affects investor behavior in this paper is plausibly a natural experiment. Thus, this paper also contributes to the behavioral finance literature that exploit natural experiment or conduct lab experiment. In 18th century, the public news on English shares was carried by sailboats to investors in Amsterdam. Because of exogenous occurrence of adverse weather conditions, the arrival of public news on these shares was regularly interrupted and there were trading days when such kind of public news were not available. Koudijs (2016) exploits this natural experiment to disentangle the impact of public news from that of private information on day-to-day price movement.

Last but not the least, this paper relates to studies based on lab experiement. Asparouhova, Hertzel, and Lemmon (2009) present their subjects with randomly generated sequences of binary outcomes (UP or DOWN) and ask the participants to assess the probability that the next outcome will be UP. They find that the probability of a streak continuing decreases for short streaks and increases for long streaks, which is consistent with the model in Rabin (2002) and inconsistent with the model in Barberis, Shleifer, and Vishny (1998). Asparouhova, Bossaerts, Eguia, and Zame (2015) create a laboratory environment in which Bayes' rule is difficult to apply and some participants do not apply it correctly. The findings from their experiments answer the question whether and to what extent asset prices reflect correct or incorrect reasoning of different investors. Asparouhova, Bossaerts, Roy, and Zame (2016) emulate the stationary, infinite-horizon setting of the Lucas asset pricing model and motivate participants to smooth consumption over time. They ask participants from Caltech, UCLA and University of Utah to trade two long-lived securities in a continuous open-book system. They find evidence that supports the model.

Hypothesis Development

Psychology studies suggest human beings are capable of taking actions to fix their mood and engage in mood regulation. Performing challenging tasks is believed to be one way to successfully overcome sad moods (Erber & Tesser, 1992). Trading appears to meet the criterion of a challenging task: Not only is it a mentally intense activity, but it also creates opportunities of monetary reward to offset the negative mood (Edmans et al., 2007). Therefore, trading volume could increase when investors are in bad moods and use trading to regulate their mood.

Mood Regulation Hypothesis: As smog levels increases, investors experience bad moods and use trading, a challenging task, to combat bad moods. The relation between smog and trading volume is positive.

An obvious impact of smog on human behavior is reduced outdoor activities, which means increased indoor activities. Outdoor activities are reduced on a smog day for three reasons. First, outdoor activities are not fun anymore. As we all know, sunshine and blue sky are key to outdoor activities. On a smog day, everywhere looks either dark or grey because sunshine is blocked by the dust in the air. As a result, people engage much less in outdoor activities because it is not fun or satisfying any more. Secondly, going outside becomes a hassle. Like fog, severe local smog reduces visibility, which slows down the traffic and makes it more difficult for people to get to outdoor activity destinations, such as shopping districts, parks, hiking trails, camping sites, etc. Thirdly, unlike other weather conditions that have no impact on people's health, smog can actually do harm to people if they do not stay inside. Medical studies show that long time exposure to smog can cause respiratory and cardiovascular disease. Thus, when smog is severe, people, healthy or not, have no other choices but stay inside where air is relative cleaner.

When outdoor activities are not possible and people are forced to stay inside, they usually try to find a way to get entertained, such as watching TV, listening to music, indulging in social media (e.g., WeChat, Twitter, and Instagram), online dating, and so on. For some people, cooking, exercising, reading, or even housing cleaning can be entertaining too. For investors, trading is a great way to spend their time when outdoor activities are not possible. The fun and enjoyment comes from researching trades, executing the order, sharing thoughts with friends, expecting and experiencing the outcome of a trade (Black, 1986; D. Dorn & Sengmueller, 2009; Gao & Lin, 2014).

The following testable prediction is based on the above arguments: a) smog significantly reduces the possibility of outdoor activities and forces people to stay inside; b) trading is a great way to for investors get entertained.

Gambling Hypothesis: As smog levels increases, investors are forced to reduce outdoor entertaining activities and they turn to trading for entertainment and gambling thrill. The relation between smog and trading volume is positive.

The gambling hypothesis predicts a positive relation between smog and trading. However, the relation could be negative as well. When smog darkens the sky and causes serious health problems, it could affect people's mood in a negative way. Medical studies actually find evidence that smog leads to depression of subjects who have been exposed to smog for a significant amount of time (Fonken et al., 2011). In addition, finance literature discover a negative impact of investor depression on stock market (Kamstra et al., 2003; Kramer & Weber, 2012; Kamstra et al., 2014, 2017). Therefore, smog may affect stock market in a negative way through the channel of investor mood. The following alternative hypothesis predicts a negative sign between smog and trading volume.

Inertia Hypothesis: As smog level increases, investors are depressed and show less interest in trading, which was exciting and entertaining to them when they were not depressed. On the other hand, as smog diminishes, investors are in a good mood and enjoy doing things that are fun to them, such as trading. The relation between smog and trading volume is negative.

To summarize, inertia hypothesis predicts a negative relation between smog and trading volume, whereas mood regulation hypothesis and gambling hypothesis both predict a positive relation. The difference between mood regulation hypothesis and gambling hypothesis is the channel through which smog affects trading volume. Gambling hypothesis believes that smog reduces outdoor entertainment opportunities and investors seek excitement and thrill from trading indoor. However, mood regulation hypothesis posits that smog causes depressionamong investors and investors engage in trading to improve their mood.

Data and Sample

PM2.5 and Health Risk

PM2.5 is a national standard recognized by the U.S. Environmental Protection Agency (EPA).[1] It refers to particulates less than 2.5 micrometers in diameter (PM2.5). They are of particular concern because they are very small and can directly enter the lungs and even the blood stream. Therefore, they pose the largest health risks.

In general, higher PM2.5 concentration value indicates worse air quality and higher health risk. Specifically, it is considered very unhealthy and everyone should limit outdoor exertion if PM2.5 concentration value is greater than 150.5; It is considered hazardous and everyone should avoid all outdoor exertion if PM2.5 concentration value is greater than 250.5 (See Appendix).

---

[1] https://www.epa.gov/pm-pollution

Hourly Measure of Smog: PM2.5 Concentration Value

I obtain smog data from U.S. Department of State Air Quality Monitoring Program-Mission China.[2] The location of the air quality monitor is at the U.S. Embassy in Beijing.[3] A special feature of the data is that it is an hourly data. An hourly measurement is taken for the 60 minutes of each hour. For example, the data reported at 5/27/2017 03:00 is taken from 5/27/2017 03:00 through 03:59. Each day, there are 24 reported measurements (24 PM2.5 concentration values) corresponding to 24 hours (0:00 to 23:00).

High Frequency Five-Minute Trading Volume

High frequency five-minute trading volume is from Wind Information Inc. (WIND), the largest and most prominent financial data provider in China. WIND serves 90% of China's financial institutions and 70% of the Qualified Foreign Institutional Investors (QFII) operating in China (Liu, Stambaugh and Yuan 2019). Trading of stock each day starts at 9:30 and ends at 15:00. Trading volume is recorded every five minutes with the first entry at 9:35 and the last entry at 15:00. Each entry represents the trading volume during the past five minutes. Trading lunch break takes place at 11:30 and trading resumes at 13:00.

---

[2] https://airnow.gov/index.cfm?action=airnow.global_summary#China$Beijing

[3] Prior to February 16, 2009, the Beijing air quality monitor was located at No. 3 Xiu Shui Bei Jie, Chaoyang District, Beijing (geographic coordinates 39.31, 116.44). Beginning on February 17, 2009, at 23:00 BJT, the air monitor began operating at No. 55 An Jia Lou Rd., Chaoyang District, Beijing. See US Department of State air quality data file fact sheet.

Trading data in this paper spans from 07/08/2016 through 03/29/2019. To match the hourly smog data, I average the entries of trading volume during an hour to get the hourly trading volume. Thus I have hourly trading volume for Hour 9, Hour 10, Hour 11, Hour 13 and Hour 14 for each firm. I do not sum the entries of trading volume within an hour to get the hourly trading volume because some hours have 30 minutes of trading while others have 60 minutes of trading. For instance, Trading starts at 9:30. Hour 9 only has 30 minutes of trading. The sum of five minute trading volumes in Hour 9 will be automatically much less than the sum of five minute trading volumes in Hour 10 simply because Hour 10 has 60 minutes of trading. Taking the average of all trading volume within an hour to get the hourly trading volume can avoid this problem and make sure Hour 9 and Hour 10 trading volume are comparable to each other.

Other Variables

Daily stock related variables (stock trading volume, stock closing price, common shares outstanding, and stock headquarter city) come from Datastream.

Constituents of Portfolio

I focus on A-share stocks because they are China's domestically traded stocks. I search Datastream for A-share stocks that are listed on the Shanghai Stock Exchange and also headquartered in Beijing. The total number of firms is 139. I then use their security code to extract five-minute trading data from WIND one firm at a time.

In my analysis, I average the five minute trading volume for each firm to get hourly trading volume for each firm. I then form an equally weighted portfolio of these firms.

Hourly trading volume of the equally-weighted portfolio is the simple average of hourly trading volume of 139 firms in the portfolio.

Sample

Sample period is from 07/08/2016 through 03/29/2019. The number of firms listed on Shanghai Stock Exchange and headquartered in Beijing is 139. Beijing smog data is at hourly frequency. It is recorded at the beginning of each hour. Figure 1.1 shows the smog level hour by hour during the trading hours from 07/08/2016 -03/29/2019. A clear pattern of smog level is that it is low in June and high in December.

Table 1.1 shows the smog distribution by air quality category. In the sample, “Good” air quality (PM2.5 concentration value between 0.0 and 12.0) occur in 817 hours, 25.15% of the sample. “Very unhealthy” air quality (PM2.5 concentration value between 150.5 and 250.4) occur in 135 hours, 4.16% of the sample. 5.73% (4.16%+1.14%+0.43%=5.73%) of the sample, air quality is considered “Very unhealthy” or even worse.

As I mention above, the original trading volume data is at five-minute frequency. To match hourly smog data with five-minute trading volume data, I average the five-minute trading volumes within an hour to get the hourly trading volume of that hour. Figure 1.2 shows the trading volume of an equally weighted portfolio of Beijing firms hour by hour during the same period of Figure 1.1.

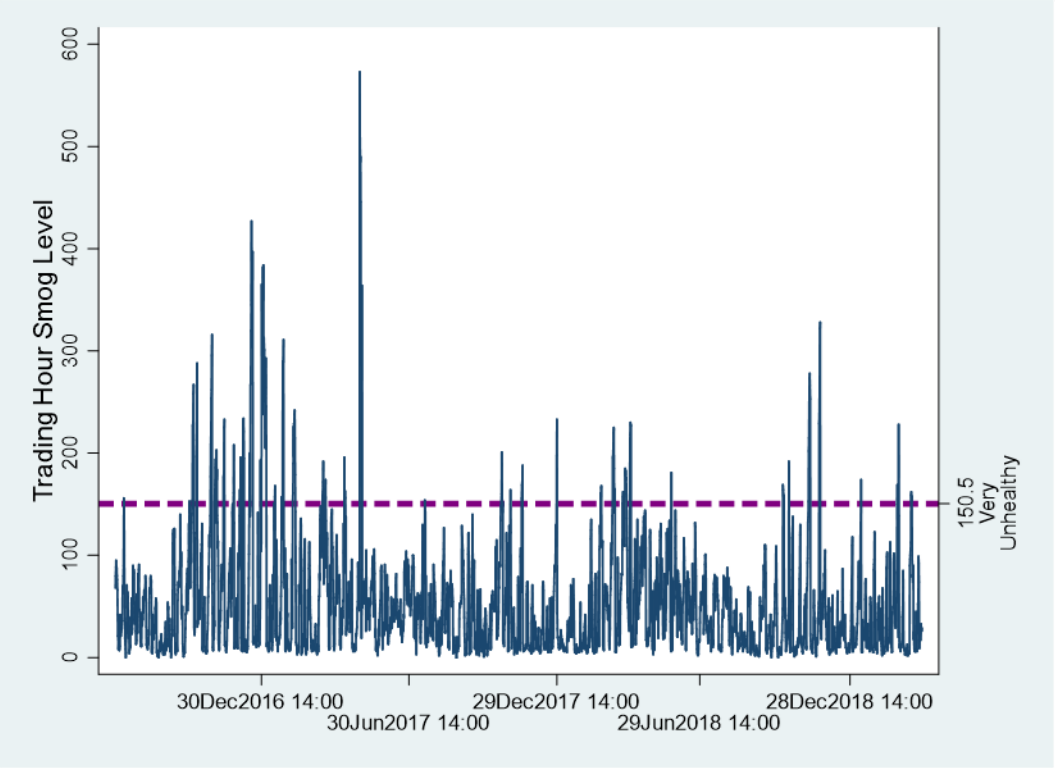


Figure 1.1. Beijing Hourly Smog

This figure shows hourly smog level during trading hours (9:30-11:30, 13:00-15:00) from 07/08/2016 10:00-03/29/2019 14:00. Smog level is measured by PM2.5 concentration value. Horizontal purple dash line marks the “Very Unhealthy” threshold (PM2.5 concentration value = 150.5).

Table 1.1. Beijing Smog Distribution by Air Quality Category

This table shows the distribution of hourly smog level (PM2.5 concentration value) during trading hours (9:30-11:30, 13:00-15:00) by air quality category for the period of 07/08/2016 10:00-03/29/2019 14:00 in Beijing. The “Frequency” Column gives the total number of hours when smog level is in a certain category during the sample period. The range of smog level for each air quality category is in parenthesis (also see Appendix B).

Air quality group	PM2.5 value range	Frequency (in hour)	Percentage
Good	(0.0-12.0)	817	25.15%
Moderate	(12.1-35.4)	808	24.87%
Unhealthy for sensetitive group	(35.5-55.4)	503	15.48%
Unhealthy	(55.5-150.4)	935	28.78%
Very unhealthy	(150.5-250.4)	135	4.16%
Hazardous	(250.5-350.4)	37	1.14%
Beyond index	(>=350.5)	14	0.43%
Total		3,249	100.00%

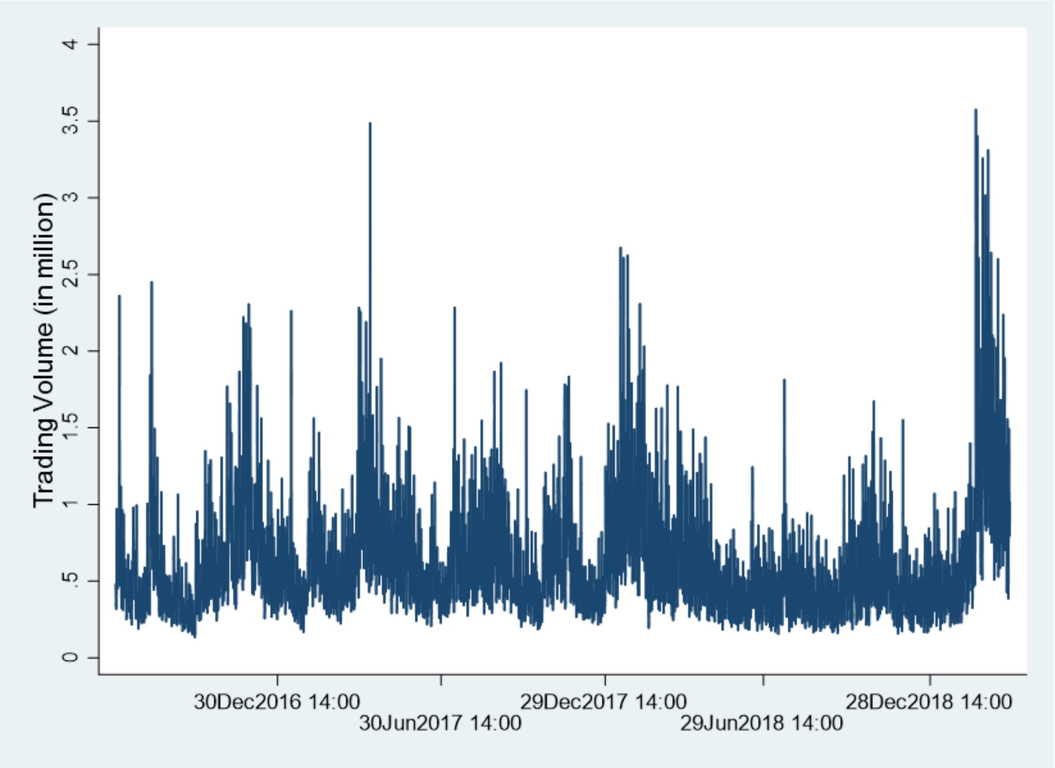


Figure 1.2. Portfolio of Beijing Firms Hourly Trading Volume

This figure shows hourly trading volume of a portfolio of Beijing firms from 07/08/2016 10:00-03/29/2019 14:00. I obtain five minute trading volume for all Beijing firms. I compute hourly trading volume as the average of all five-minute trading volumes within that hour. For instance, there are 12 five-minute trading volumes within the hour 10:00. The average of those 12 trading volume is the trading volume for the hour 10:00. The portfolio is an equally weighted portfolio.

In Table 1.2 Summary Statistics Panel A, the mean (median) of natural log of the portfolio trading volume at hour t is 13.12 (13.07). To capture past trading volume, I compute Log(Volume_(t-1, t-55)), the average trading volume between hour t-1 and hour t-55. Since there are five trading hours each day, there are 11 trading days from hour t-1 and hour t-55. Thus Log(Volume_(t-1, t-55)) measures the trading volume of the past 11 days.

Trading volume is persistent, which can be seen in Table 1.2 Summary Statistics Panel A. The distribution of trading volume at hour t and the distribution of average trading volume between hour t-1 and t-55 are very close. Table 1.2 Summary Statistics Panel B also shows that the correlation between the two is high (Pearson correlation = 0.52). In Table 1.2 Panel A, the mean (median) of smog at hour t is 52.45 (35.00). Panel B shows that smog is also persistent. The Pearson correlation between Smog_t and Smog_t-1 is 0.86.

Table 1.2. Summary Statistics

This table provides summary statistics of main variables. Sample period is 07/08/2016 10:00-03/29/2019 14:00. The number of firms that appear in the portfolio is 139. Panel A shows the distribution statistics, and Panel B presents the correlation matrix. In Panel B, Pearson correlations are below diagonal and Spearman correlations are above diagonal. The portfolio is equally weighted. Variables construction is described in Appendix.

Panel A: Distribution statistics

	N	Mean	Median	Std. Dev	Min	P25	P75	Max
Log(Volume_t)	3,249	13.12	13.07	0.52	11.81	12.75	13.43	15.09
Log(Volume_(t-1, t-55))	3,249	13.21	13.20	0.31	12.55	12.96	13.40	14.17
Smog_t	3,249	52.45	35.00	57.59	0.00	12.00	71.00	573.00
Smog_t-1	3,249	54.07	36.00	65.44	0.00	13.00	71.00	985.00
Smog_t-2	3,249	54.64	36.00	66.31	0.00	13.00	72.00	985.00
Smog_t-3	3,249	55.92	37.00	70.41	0.00	14.00	74.00	985.00

Panel B: Correlation matrix

	Log(Volume_t)	Log(Volume_(t-1, t-55))	Smog_t	Smog_t-1	Smog_t-2	Smog_t-3
Log(Volume_t)	1	0.50	0.04	0.03	0.03	0.03
Log(Volume_(t-1, t-55))	0.52	1	0.00	0.01	0.01	0.01
Smog_t	0.06	0.03	1	0.96	0.92	0.87
Smog_t-1	0.06	0.04	0.86	1	0.95	0.91
Smog_t-2	0.07	0.04	0.83	0.77	1	0.96
Smog_t-3	0.06	0.04	0.75	0.77	0.73	1

Initial Graphical Evidence

In Figure 1.3, portfolio of Beijing firms trading volumes are sorted into seven air quality groups (See the cutoff of these air quality groups in Appendix 2). Mean trading volume for each air quality group are plotted on the y-axis. The red vertical line segment represents the 95% confidence interval of the mean. In Fig 3.1, trading volume is sorted by contemporaneous smog level. In Fig 3.2 through 3.4, trading volume is sorted by lagged smog level.

Two patterns shown in these four graphs are: 1) Very Unhealthy group has the highest average trading volume; 2) The lower bound of its 95% confidence interval is also above the upper bound of four better air quality groups (Good, Moderate, Unhealthy to Sensitive Group, Unhealthy). These patterns trading volume is distinguishably higher in Very Unhealthy group than in better air quality groups. The graphs thus are consistent with the mood regulation hypothesis.

Fig 3.1 shows a contemporaneous increase in trading volume when air quality is Very Unhealthy. However Fig 3.2, Fig 3.3 and Fig 3.4 show a lagged increase in trading volume. Both the contemporaneous and lagged increase in trading volume are plausible since it takes time for smog to cause bad moods and for investors to take actions to regulate their moods.

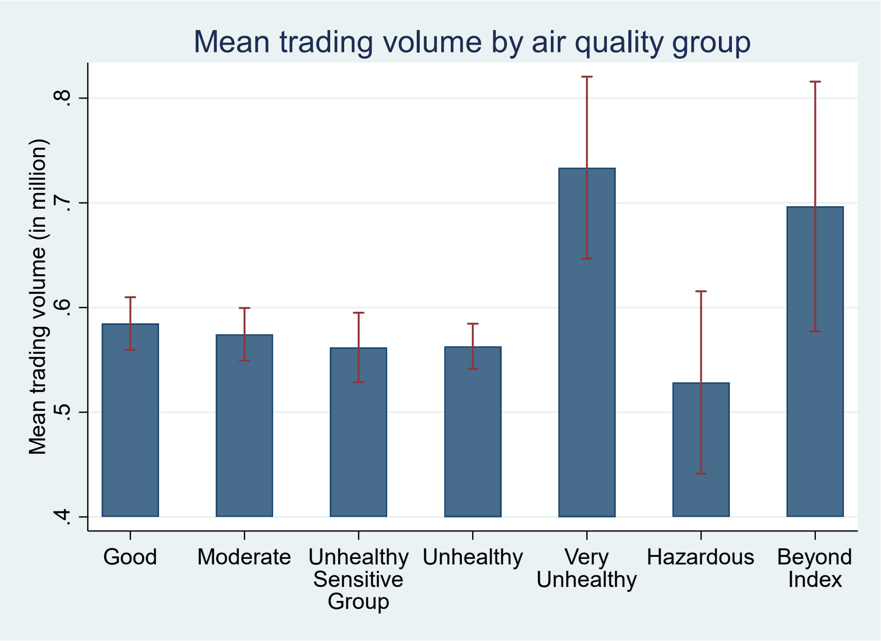


Figure 1.3.1

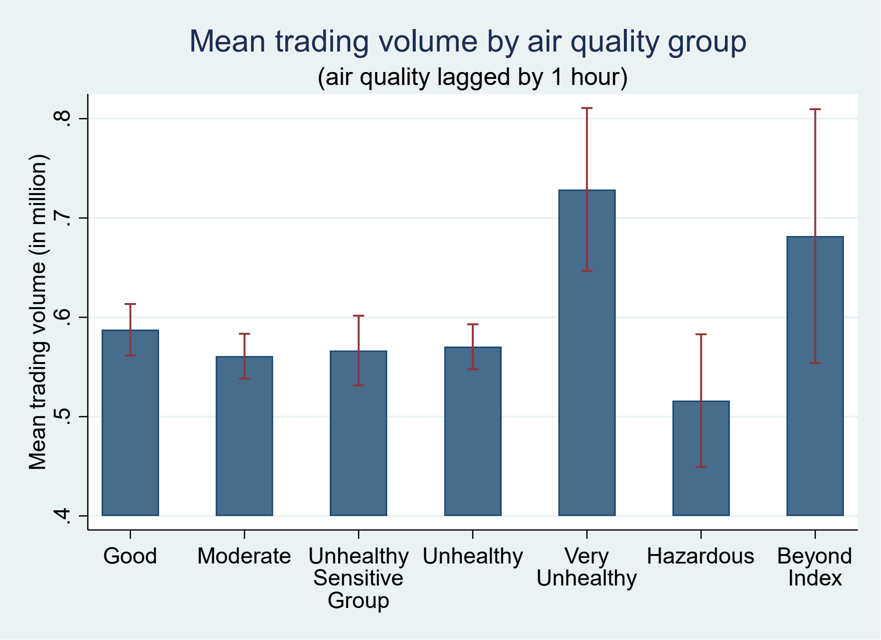


Figure 1.3.2

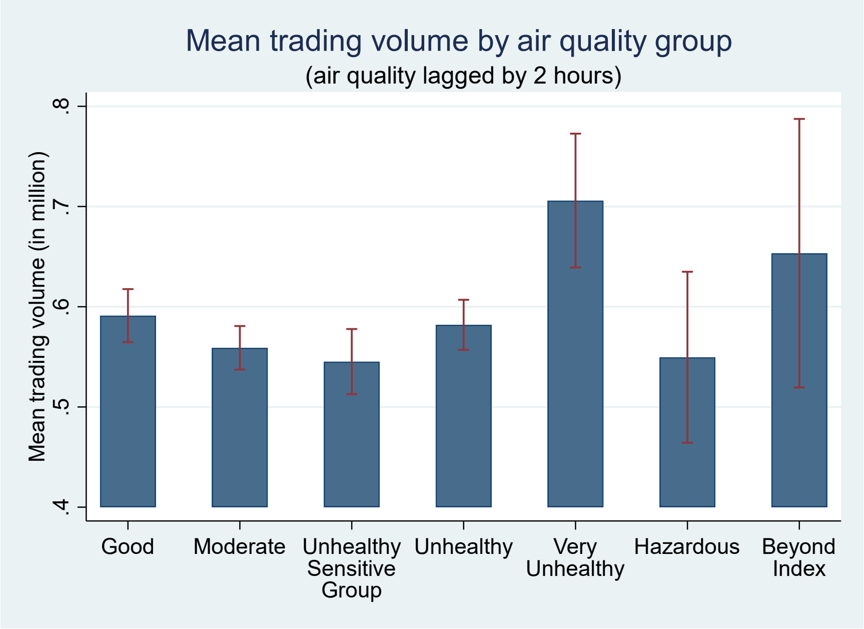


Figure 1.3.3

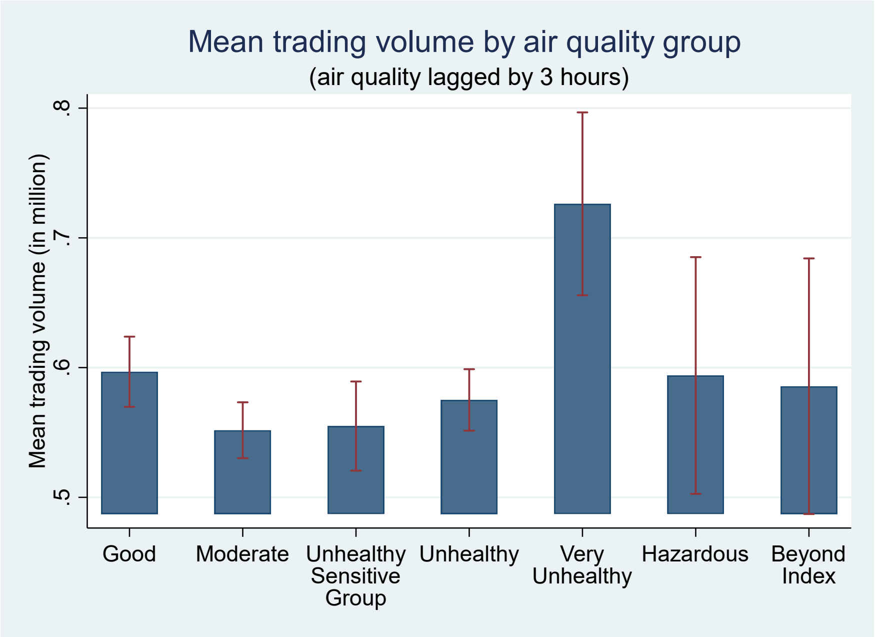


Figure 1.3.4

Figure 1.3. Portfolio of Beijing Firms Average Hourly Trading Volume

This figure shows portfolio of Beijing firm average hourly trading volume by air quality group. There are seven air quality groups: good, moderate, unhealthy for sensitive group, unhealthy, very unhealthy, hazardous, beyond index. Hourly trading volumes are sorted into these seven groups by corresponding smog level, measured by PM2.5 concentration value. The vertical bars in each graph show the average trading volume of each air quality group. The red vertical line segment represents 95% confidence interval of the mean trading volume. The first group sorts the trading volume by the smog level during the same hour. The second, third and fourth group sort the trading volume by the smog level with one hour lag, two hour lag and three hour lag respectively.

One feature in these graph is trading volume doesn’t increase monotonically as air quality worsens. Trading volume peaks in Very Unhealthy group. As air quality continues to worsen, trading volume doesn’t continue to increase. Also notice that, due to the very small number of observations for the Hazardous and Beyond Index air quality group (1.14% and 0.43% of the sample respectively, see Table 1.1), the 95% confidence interval of average trading volume for these two groups are very wide, which suggests that the average trading volume in these two groups may not reflect the true population mean.

Empirical Results

An Exogenous Discrete Smog Variable

According to EPA, it is very unhealthy when PM2.5 concentration value is above 150.5 (see Appendix 2), thus I create a smog dummy variable which take the value 1 if PM2.5 concentration value is above 150.5, otherwise 0. This variable is truly exogenous because the threshold is randomly determined.

I then run OLS regression with dummy smog measure as follows:

$$\begin{aligned} Log(Volume)_t = {} & Very\ Unhealty_t + Log(Volume_{(t-1,\ t-55)}) + \\ & Hour\ of\ the\ Day_t + \ Day\ of\ the\ Week_t + \epsilon_t \qquad (1) \end{aligned}$$

where all variables are the same as defined in (1) except that Very Unhealthy is a dummy variable which takes value 1 if PM2.5 concentration value is greater than 150.5, otherwise 0.

Table 1.3. Impact of Smog Dummy on Trading Volume

This table reports the results of OLS regression of trading volume of equally weighted Beijing portfolio on Very Unhealthy dummy. Very Unhealthy dummy takes value 1 if smog level is above the threshold 150.5, otherwise 0. Controls are Log(Volume_(t-1, t-55)), Hour9, Hour10, Hour11, Hour13, MON, TUE, WED, THR (See appendix for variable definition). t-statistics (reported in parentheses) are calculated using the Newey and West (1987) heteroscedasticity and autocorrelation consistent covariance matrix estimator with automatically selected lag, as in Newey and West (1994). Asterisks denote statistical significance at 1% (***), 5% (**), or 10% (*) level.

	(1)	(2)	(3)
Very Unhealthy_t	0.1172***	0.0910**	0.0913**
	(2.84)	(2.43)	(2.37)
Log(Volume_(t-1,t-55))	0.8644***	0.8702***	0.8694***
	(23.40)	(23.84)	(23.79)
Hour9		0.6062***	0.6063***
		(43.95)	(44.04)
Hour10		0.0511***	0.0511***
		(4.38)	(4.37)
Hour11		-0.2833***	-0.2832***
		(-22.03)	(-22.01)
Hour13		-0.2627***	-0.2628***
		(-23.72)	(-23.69)
MON			0.0555**
			(2.21)
TUE			0.0062
			(0.22)
WED			0.0037
			(0.14)
THR			0.0155
			(0.75)
Constant	1.6931***	1.5952***	1.5907***
	(3.48)	(3.32)	(3.30)
Observations	3,249	3,249	3,249
Adjusted R-squared	0.268	0.651	0.652

The results are reported in Table 1.3. As in previous table, t-statistics (reported in parentheses) is calculated using the Newey and West (1987) heteroscedasticity and autocorrelation consistent covariance matrix estimator with automatically selected lag, as in Newey and West (1994). In Column (3), the coefficient on Very Unhealthy_t is positive and statistically significant (coefficient = 0.0913, t-statistics = 2.37). The results in the table support the mood regulations hypothesis and reject the inertia hypothesis.

Alternative Cutoff Values of Smog Dummy

The mood regulation hypothesis relies on the severity of smog. Smog has to be severe enough to make the effect observable and testable. The cutoff of 150.5 for smog dummy is plausible because EPA healthy guides suggest that when PM2.5 concentration value is greater than 150.5, everyone should limit outdoor activity. However, some might still have the concern that the cutoff is somewhat arbitrary. To accommodate the concern, I rerun the regression in equation (3) with alternative threshold levels 110.5, 130.5, 170.5, and 190.5, representing 89th, 92th, 96th, and 97th percentiles of the smog distribution in the sample, respectively.

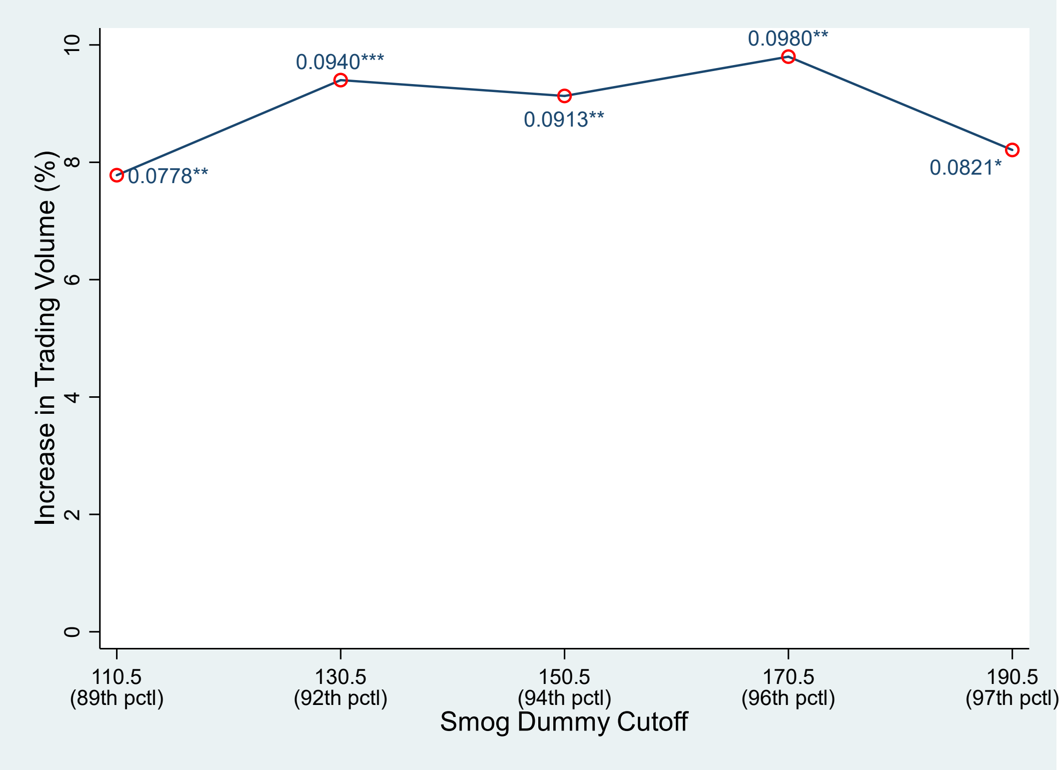


Figure 1.4. Smog Dummy with Alternative Cutoff and Trading Volume

This figure summarizes the regression results of trading volume on smog dummy with alternative cutoff. The alternative cutoffs and their corresponding percentile are on x-axis. The numbers next to the red circle in the graph represents the coefficient on the smog dummy. Asterisks denote statistical significance at 1% (***), 5% (**), or 10% (*) level. Dependent variable is the log of trading volume. A positive coefficient translates into 100*coefficient percent increase in trading volume when smog dummy increases from zero to one.

Rather than reporting all the coefficient estimates, I use Figure 1.4 to convey the most important message: when the estimated coefficients on the smog dummy are plotted against alternative thresholds, the sign, magnitude of coefficient, and statistical significance level are qualitatively the same as the 150.5 cutoff. For example, when 110.5 (89th percentile) is used as the cutoff for smog dummy, the coefficient on the dummy variable is 0.0778 and statistically significant at 5% level, indicating that trading volume increases by 7.78% when PM2.5 concentration value is above 110.5.

Overall, Figure 1.4 suggests that the positive impact of smog dummy on trading volume is robust to alternative cutoff values of smog dummy.

Continuous Smog Measure

To test the hypothesis, I run OLS regression with continuous smog measure.

$$Log(Volume)_t = Smog_t + Log(Volume_{(t-1,\ t-55)}) + Hour\ of\ the\ Day_t + Day\ of\ the\ Week_t + \epsilon_t \qquad (2)$$

where Log(Volume_t) is the natural log of equal-weighted Beijing portfolio trading volume at hour t, smog_t is smog level for trading hour, Log(Volume_(t-1, t-55)) is the natural log of average portfolio trading volume between hour t-1 and t-55 which captures average trading volume of the portfolio during the past 11 days, Hour of the Day are dummies for Hour 9, Hour 10, Hour 11, and Hour 13 with Hour 14 as the base to capture intraday trading patterns, Day of the Week are day of the week dummies with Friday as the base to capture intraweek trading patterns, ϵ_t is zero-mean disturbance term which can be correlated with ϵ_{t-1}.

Table 1.4 reports the coefficient on Smog_t with t-statistics (reported in parentheses) calculated using the Newey and West (1987) heteroscedasticity and autocorrelation consistent covariance matrix estimator with automatically selected lag, as in Newey and West (1994).

In Column (1), I only control for past trading volume. In Column (2), I control for both past trading volume and hour of the day. In Column (3), I control for past trading volume, hour of the day and day of the week as in equation (1). All three columns show that smog has a positive contemporaneous impact on trading volume and past trading volume is a good predictor of current trading volume.

As noted in Column (3), the coefficient on is positive (=0.003) and statistically significant on the 10% level (t stat=1.88). This results supports the mood regulation hypothesis which predicts a positive sign between smog and trading volume.

In Column (2) and (3), the coefficient on Monday Dummy is positive and highly statistically significant (at 5% level) which is consistent with the fact that investor trade on the information released over the weekend.

Also in In Column (2) and (3), the coefficient on Hour 9 Dummy and Hour 10 is positive and highly statistically significant (at 1% level) which is consistent with the fact that investor trade on the information released over night.

Table 1.4. Impact of Smog Level on Trading Volume

This table reports the results of OLS regression of trading volume of equally weighted Beijing portfolio on smog level. Controls are Log(Volume_(t-1, t-55)), Hour9, Hour10, Hour11, Hour13, MON, TUE, WED, THR (See appendix for variable definition). t-statistics (reported in parentheses) are calculated using the Newey and West (1987) heteroscedasticity and autocorrelation consistent covariance matrix estimator with automatically selected lag, as in Newey and West (1994). Asterisks denote statistical significance at 1% (***), 5% (**), or 10% (*) level.

	(1)	(2)	(3)
Smog_t	0.0004**	0.0003*	0.0003*
	(2.17)	(1.83)	(1.88)
Log(Volume_(t-1,t-55))	0.8674***	0.8725***	0.8716***
	(23.49)	(23.95)	(23.89)
Hour9		0.6066***	0.6067***
		(44.04)	(44.14)
Hour10		0.0502***	0.0502***
		(4.30)	(4.29)
Hour11		-0.2837***	-0.2836***
		(-22.06)	(-22.03)
Hour13		-0.2627***	-0.2627***
		(-23.72)	(-23.68)
MON			0.0579**
			(2.30)
TUE			0.0094
			(0.33)
WED			0.0048
			(0.18)
THR			0.0151
			(0.73)
Constant	1.6410***	1.5550***	1.5499***
	(3.37)	(3.24)	(3.22)
Observations	3,249	3,249	3,249
Adjusted R-squared	0.267	0.651	0.652

Lagged Smog Variable

When smog level increases, it might take some time for the mood to change and for people to act on their mood. So smog level in previous hours could matter for trading volume. I replace the contemporaneous smog variable with lagged smog variable.

Consequently, I perform the following equation.

$$\begin{aligned} Log(Volume)_t = Smog_{t-1} + Log(Volume_{(t-1,\ t-55)}) \\ +Hour\ of\ the\ Day_t \\ +Day\ of\ the\ Week_t + \epsilon_t \end{aligned} \tag{3}$$

where all variables are the same as defined in (1) except that Smog_t is the smog level one hour before. I also replace Smog_t-1 with Smog_t-2 (smog level two hours ago) or Smog_t-3 (smog level three hours ago).

Table 1.5 reports the results. The table shows that smog level hours ago has an impact on trading volume with the stronger impact taking place two hour or three hours ago.

Table 1.5. Impact of Smog Level in Previous Hours on Trading Volume

This table reports the results of OLS regression of trading volume of equally weighted Beijing portfolio on smog level in previous hours. Smog_t is contemporaneous smog level. Smog_t-1 is smog level one hour ago. Similarly for Smog_t-2, Smog_t-3. Controls are Log(Volume_(t-1, t-55)), Hour9, Hour10, Hour11, Hour13, MON, TUE, WED, THR (See appendix for variable definition). t-statistics (reported in parentheses) are calculated using the Newey and West (1987) heteroscedasticity and autocorrelation consistent covariance matrix estimator with automatically selected lag, as in Newey and West (1994). Asterisks denote statistical significance at 1% (***), 5% (**), or 10% (*) level.

	(1)	(2)	(3)	(4)
Smog_t	0.0003*			
	(1.88)			
Smog_t-1		0.0002*		
		(1.67)		
Smog_t-2			0.0003**	
			(1.97)	
Smog_t-3				0.0003**
				(2.11)
Log(Volume_(t-1,t-55))	0.8716***	0.8715***	0.8715***	0.8712***
	(23.89)	(23.88)	(23.93)	(23.93)
Hour9	0.6067***	0.6063***	0.6059***	0.6063***
	(44.14)	(44.02)	(44.04)	(44.00)
Hour10	0.0502***	0.0498***	0.0493***	0.0494***
	(4.29)	(4.24)	(4.20)	(4.21)
Hour11	-0.2836***	-0.2840***	-0.2841***	-0.2848***
	(-22.03)	(-22.09)	(-22.09)	(-22.11)
Hour13	-0.2627***	-0.2625***	-0.2629***	-0.2627***
	(-23.68)	(-23.67)	(-23.68)	(-23.78)
MON	0.0579**	0.0573**	0.0577**	0.0575**
	(2.30)	(2.28)	(2.30)	(2.28)
TUE	0.0094	0.0090	0.0096	0.0097
	(0.33)	(0.32)	(0.34)	(0.34)
WED	0.0048	0.0037	0.0049	0.0041
	(0.18)	(0.13)	(0.18)	(0.15)
THR	0.0151	0.0147	0.0154	0.0156
	(0.73)	(0.71)	(0.74)	(0.75)
Constant	1.5499***	1.5559***	1.5534***	1.5568***
	(3.22)	(3.23)	(3.23)	(3.24)
Observations	3,249	3,249	3,249	3,249
Adjusted R-squared	0.652	0.652	0.652	0.652

Smog Positively Impact Trading Volume at the Firm Level

I perform regression in equation (2) (regression with continuous smog measure) firm by firm. The dependent variable is the log of hourly trading volume of a firm. The key independent variable is continuous smog measure (PM2.5 concentration value). Key control variable is the log of average past 55 hour trading volume of a firm. The motivation is to control for persistence in firm trading volume. Other controls are hour of the day dummies and day of the week dummies. Similarly, I perform regression in equation (1) (smog dummy) firm by firm.

The results are summarized in three dimensions. First, the mean coefficients on key independent variables are computed to assess the mean effect. Second, t-statistics is computed for the mean coefficient under the assumption that coefficients across firms are independent and identically distributed (i.i.d) and H0: the mean coefficient =0. Third, the number of firms that share the same sign as the mean coefficient and have p-values less than 0.05 or 0.1, according to the HAC standard errors.

Table 1.6. Impact of Smog on Trading Volume, Firm-Level Evidence

This table summarizes the firm by firm results of OLS regression of trading volume on smog. The dependent variable is the log of the firm's hourly trading volume. Controls include log of firm's past 55 hour average trading volume, hour of the day dummy and day of the week dummy. In Panel A, a continuous variable, PM2.5 concentration value, is used to measure smog level. In Panel B, a dummy variable, Very Unhealthy, is usded to measure smog level. Reported are the (1) average of the estimated coefficients, (2) t-statististics on the average coefficient, under the assumption that the coefficients are distributed i.i.d and H0: the average coefficient = 0, and (3) #, p-val < 0.05 (and #, p-value < 0.1), which corresponds to the number of coefficients with a p-value less than 0.05 (0.1) and with a sign that is the same as the average coefficient; these are reported in curly brackets, and are based on the Newey and West (1987) heterosecedasticity and autocorrelation covariance matrix estimator.

Panel A: continuous smog variable

PM2.5 Concentration Value	Average	0.0002
	t-statistics	(8.5045)
	#, p-value <0.05	{16}
	#, p-value <0.1	{31}
Log of average of past 55 hours of trading volume	Average	0.8526
	t-statistics	(167.7416)
	#, p-value <0.05	{139}
	#, p-value <0.1	{139}
Average Adjusted R2		0.4986
Number of firms		139

Panel B: dummy smog variable

Very Unhealthy Dummy	Average	0.0759
	t-statistics	(11.0455)
	#, p-value <0.05	{32}
	#, p-value <0.1	{48}
Log of average of past 55 hours of trading volume	Average	0.8529
	t-statistics	(169.1898)
	#, p-value <0.05	{139}
	#, p-value <0.1	{139}
Average Adjusted R2		0.4987
Number of firms		139

Table 1.6 Panel A shows that smog has a positive impact on the trading volume at firm level. Trading volume increases, on average, by a highly statistically significant 1.15% (0.0002*57.59*100=1.15%, t-stat=8.5045) when the smog level increases by one standard deviation in the same hour (see Panel A, the first row). Equally important, there are 16 out of 139 firms (see the row marked #, p-value <0.05) that support a positive coefficient on the continuous smog variable in equation (1) and yet have p-values less than 0.05 associated with the coefficient, according to the HAC standard errors. The number of firms rises to 31 when the criterion is relaxed to p-value less than 0.1.

Table 1.6 Panel B provides additional evidence that smog has a positive impact on the trading volume at firm level. Trading volume increases, on average, by a highly statistically significant 7.59% (0.0759*100=7.59%, t-stat=11.0455) when air quality is Very Unhealthy or worse in that hour (see Panel B, the first row).

Equally important, there are 32 out of 139 firms 9 (see the row marked #, p-value<0.05) that support a positive coefficient on the smog dummy variable in equation (3) and yet have p-value less than 0.05 associated with the coefficient, according to the HAC standard errors. The number of firms rises to 48 when the criterion is relaxed to p-value less than 0.1.

To summarize, when smog level is high outside, stock trading increases.

Alternative Explanations

In previous sections, the results support the mood regulation hypothesis and reject the inertia hypothesis. However, there is an alternative hypothesis, gambling hypothesis, that could explain the positive relation between smog and stock trading.

Gambling hypothesis argues that when smog is high outside, investors seek gambling excitement indoors. Thus they trade more when smog is high. In the following sections, I try to disentangle one from each other.

Evidence against Gambling Hypothesis

To test the gambling hypothesis, I sort stocks by firm size. To measure firm size, I use firm’s daily market capitalization, defined as closing price times shares outstanding. Similar to sorting stocks by trading activity, I sort all stocks in the sample by their market capitalization each day. The 30th percentile and 70th percentile are used as cutoffs. I then form three equally weighted portfolios: large, medium and small. I rerun the regressions in equation (1) and (2) for these portfolios respectively.

The results are shown in Table 1.7. In Panel A, the coefficient on Very Unhealthy smog dummy variable is highly statistically significant (t stat=3.07) in Column (1), the large stocks. However, it is not significant in Column (2), the medium stocks, and Column (3), the small stocks. This suggests that smog has more impact on trading volume among larger firms. Panel B shows the results for the continuous smog variable. Similar to Panel A, the effect is significant among large firms than in small firms.

This evidence is against the gambling hypothesis because if investors seek gambling thrill, the smog effect should be more pronounced among small stocks rather than big stocks.

Table 1.7. Sort Firms by Firm Size

This table reports results of sorting firms by their size. By their daily market capitalization, firms are sorted into three groups: large (Column 1), medium (Column 2) and small (Column 3). The cutoff points are 30th and 70th percentile of market capitalization. Dependent variable is hourly trading volume. Controls are Log(Volume_(t-1, t-55)), Hour9, Hour10, Hour11, Hour13, MON, TUE, WED, THR (See appendix for variable definition). In Panel A, a dummy variable, Very Unhealthy, is used to measure smog level. In Panel B, a continuous variable, PM2.5 concentration value, is used to measure smog level. t-statistics (reported in parentheses) are calculated using the Newey and West (1987) heteroscedasticity and autocorrelation consistent covariance matrix estimator with automatically selected lag, as in Newey and West (1994). Asterisks denote statistical significance at 1% (***), 5% (**), or 10% (*) level.

Panle A: Dummy Smog Variable

	(1) Large	(2) Medium	(3) Small
Very Unhealthy	0.1517***	0.0571	0.0606
	(3.07)	(1.36)	(1.50)
Log(Volume_(t-1,t-55))	0.7625***	0.8698***	0.9052***
	(11.06)	(18.02)	(27.83)
Hour of the Day	YES	YES	YES
Days of the Week	YES	YES	YES
Observations	1,779	1,779	1,779
Adjusted R-squared	0.4874	0.6563	0.7130
F-test	177.8	269	308.4

Panle B: Continuous Smog Variable

	(1) Large	(2) Medium	(3) Small
Smog_t	0.0006**	0.0003	0.0002
	(2.46)	(1.40)	(1.40)
Log(Volume_(t-1,t-55))	0.7684***	0.8677***	0.9021***
	(11.08)	(17.95)	(27.50)
Hour of the Day	YES	YES	YES
Days of the Week	YES	YES	YES
Observations	1,779	1,779	1,779
Adjusted R-squared	0.4866	0.6565	0.7130
F-test	178.6	265.5	313.1

Evidence Supporting Mood Regulation Hypothesis

I then sort stocks by their idiosyncratic volatility. Idiosyncratic volatility is not directly observable. I define idiosyncratic volatility as the sum of squared errors form the following regression model:

$$r_{it} = \alpha_i + \beta_i Market\ Return_t + \varepsilon_{it} \tag{4}$$

where r_{it} is daily stock return of stock i, market return is daily SSE Index return, ε_{it} is the error term. I obtain the predicted value of the error term, square it and sum over a month to get the monthly idiosyncratic volatility of a stock. Similar to sorting stocks by trading activity, I sort all stocks in the sample by their idiosyncratic volatility each month. The 30th percentile and 70th percentile are used as cutoffs. I then form three equally weighted portfolios: high volatility, medium volatility and low volatility. I rerun the regressions in equation (1) and (2) for these portfolios respectively.

The results are shown in Table 1.8. In Panel A, the coefficient on the Very Unhealthy smog dummy is statistically significant across all three portfolios but has a larger magnitude and greater t statistics in low volatility portfolio. Panel B, the coefficient on the continuous smog variable exhibits similar pattern. In general, Table 1.8 shows that the smog effect is more pronounced among low idiosyncratic firms. This is evidence to support the mood regulation hypothesis. Investors go after low risk stocks suggests that they are risk averse, which is a sign a bad mood or depression because psychology studies show that depression and risk averse are related. Hit by smog, investors become depressed and risk averse. They trade low risk stocks to combat bad mood. This is consistent with the mood regulation hypothesis.

Table 1.8. Sort Firms by Idiosyncratic Volatility

This table reports results of sorting firms by monthly their idiosyncratic volatility. To obtain monthly idiosyncratic volatility, I first regress daily stock return on market return for each firm. Square the daily residual and sum over a month for each firm. By their monthly idiosyncratic volatility, firms are sorted into three groups: high volatility (Column 1), medium volatility (Column 2) and low volatility (Column 3). The cutoff points are 30th and 70th percentile of monthly idiosyncratic volatility. Dependent variable is hourly trading volume. Controls are Log(Volume_(t-1, t-55)), Hour9, Hour10, Hour11, Hour13, MON, TUE, WED, THR (See appendix for variable definition). In Panel A, a dummy variable, Very Unhealthy, is used to measure smog level. In Panel B, a continuous variable, PM2.5 concentration value, is used to measure smog level. t-statistics (reported in parentheses) are calculated using the Newey and West (1987) heteroscedasticity and autocorrelation consistent covariance matrix estimator with automatically selected lag, as in Newey and West (1994). Asterisks denote statistical significance at 1% (***), 5% (**), or 10% (*) level.

Panle A: Dummy Smog Variable

	(1) High volatility	(2) Medium volatiltiy	(3) Low volatility
Very Unhealthy	0.0994*	0.0923**	0.1095**
	(1.72)	(2.11)	(2.32)
Log(Volume_(t-1,t-55))	0.8695***	0.8870***	0.8418***
	(29.55)	(25.85)	(19.73)
Hour of the Day	YES	YES	YES
Days of the Week	YES	YES	YES
Observations	3,249	3,249	3,249
Adjusted R-squared	0.6886	0.6428	0.5090
F-test	589.9	427.3	278.2

Panle B: Continuous Smog Variable

	(1) High volatility	(2) Medium volatiltiy	(3) Low volatility
Smog_t	0.0003	0.0003*	0.0004*
	(1.26)	(1.66)	(1.86)
Log(Volume_(t-1,t-55))	0.8702***	0.8881***	0.8446***
	(29.53)	(25.91)	(19.82)
Hour of the Day	YES	YES	YES
Days of the Week	YES	YES	YES
Observations	3,249	3,249	3,249
Adjusted R-squared	0.6882	0.6425	0.5083
F-test	590.5	427.3	277.5

Sort Stocks by Trading Activity

I use the trading volume to measure how actively the stocks are traded. Each hour, I sort all stocks in the sample by trading volume. The 30th percentile and 70th percentile are used as cutoffs. I then form three equally weighted portfolios: high trading activity, medium trading activity and low trading activity. I rerun the regressions in equation (1) and (2) for these portfolios respectively.

The results are shown in Table 1.9. In Panel A, the coefficient on Very Unhealthy smog dummy is statistically significant across three portfolios. However, the t-statistics decreases from Column (1) to Column (3), so does the magnitude of the coefficient, which indicates that the effect is more pronounced among more actively traded firms. In Panel B, the coefficient on the continuous smog variable is significant in Column (1), the high trading activity portfolio, but insignificant in Column (2), the medium trading activity portfolio, and Column (3), the low trading activity portfolio, which also indicates that the effect is more pronounced among more actively traded stocks.

To the extent that active investors and passive investors prefer and hold different types of stocks, this indicates passive investors remain passive even if they experience mood shift whereas active investors become more active after they experience bad moods.

Table 1.9. Sort Firms by Trading Activity

This table reports results of sorting firms by their trading activity. By their hourly trading volume, firms are sorted into three groups: high activity (Column 1), medium activity (Column 2) and low activity (Column 3). The cutoff points are 30th and 70th percentile of trading volume. Dependent variable is hourly trading volume. Controls are Log(Volume_(t-1, t-55)), Hour9, Hour10, Hour11, Hour13, MON, TUE, WED, THR (See appendix for variable definition). In Panel A, a dummy variable, Very Unhealthy, is used to measure smog level. In Panel B, a continuous variable, PM2.5 concentration value, is used to measure smog level. t-statistics (reported in parentheses) are calculated using the Newey and West (1987) heteroscedasticity and autocorrelation consistent covariance matrix estimator with automatically selected lag, as in Newey and West (1994). Asterisks denote statistical significance at 1% (***), 5% (**), or 10% (*) level.

Panel A: Continuous Smog Variable

	(1) High activity	(2) Medium activity	(3) Low activity
Very Unhealthy	0.0963**	0.0681**	0.0688*
	(2.36)	(2.16)	(1.90)
Log(Volume_(t-1,t-55))	0.8646***	0.9170***	0.9581***
	(23.86)	(25.10)	(33.85)
Hour of the Day	YES	YES	YES
Days of the Week	YES	YES	YES
Observations	3,249	3,249	3,249
Adjusted R-squared	0.6416	0.6969	0.7282
F-test	558.7	596.9	545.5

Panel B: Continuous Smog Variable

	(1) High activity	(2) Medium activity	(3) Low activity
Smog_t	0.0003*	0.0002	0.0002
	(1.94)	(1.41)	(1.07)
Log(Volume_(t-1,t-55))	0.8667***	0.9180***	0.9602***
	(23.96)	(25.14)	(34.05)
Hour of the Day	YES	YES	YES
Days of the Week	YES	YES	YES
Observations	3,249	3,249	3,249
Adjusted R-squared	0.6412	0.6963	0.7276
F-test	560.5	595.4	542.3

Conclusions

Behavioral theoretical prediction on the effect of mood on trading volume is inclusive. This paper exploits repeated natural experiments from the frequent occurrence of severe smog in Beijing to test two competing hypothesis: the inertia hypothesis, which posits that bad moods cause inactivity and inertia and thus decreasing in trading volume, and the mood regulation hypothesis, which suggests that bad moods increase trading volume because investors use trading as a way to combat bad moods.

This paper takes advantage of hourly smog and trading volume data to conduct intra-day level analysis. The smog data is from US EPA, an accurate and reliable source. The high frequency trading data is from WIND, a Chinese data set used in multiple top-tier finance journal publications.

There are a number of core empirical findings. For one, this paper finds that smog in Beijing causes trading volume of stocks headquartered in Beijing to increase, which contradicts the inertia hypothesis. The effect is more pronounced among large stocks, which rules out the possibility that investors seek gambling thrill during smog. In particular, the effect is more pronounced among low risk stocks, which reflects the risk aversion associated with depression among investors and supports the mood regulation theory. The paper also finds that the effect is more pronounced among actively traded stocks. To the extent that active investors and passive investors prefer and hold different types of stocks, this indicates passive investors remain passive even if they experience mood shift whereas active investors become more active after they experience bad moods.

Overall, my evidence provides the takeaway that investor do engage in mood regulation activities.

CHAPTER 2

LIQUIDITY VALUE AND IPO UNDERPRICING

Introduction

Investors demand lower returns and give higher valuations to more liquid assets (Amihud and Mendelson 1986).[4] IPOs transform private companies into publicly traded ones and therefore significantly improve the liquidity of their shares. Amihud and Mendelson (1988) point out that "Going public is the most fundamental form of increasing liquidity". Better liquidity can raise investors' willingness to invest, and directly increase firms' market value. In addition, founders and early investors benefit from better liquidity because they can exit their investments with smaller price impact in a liquid public market.[5] I refer to such increased payoffs to founders and investors due to improved liquidity as "liquidity value" in this chapter. Despite IPOs' significant impact on the liquidity of stocks, little attention has been paid to the relation between liquidity value and the determination of IPO offer prices. In this study, I posit that, issuers share the benefit of liquidity value with IPO investors by setting a discounted offer price, i.e., IPO underpricing – a stylized fact that IPO stocks typically yield large first-day returns after going public.[6]

---

[4] Relatedly, Balakrishnan, Billings, Kelly, and Ljungqvist (2014) show that managers actively shape firms' information environments in order to improve the liquidity of their shares and achieve higher valuation.

[5] Aggarwal, Krigman, and Womack (2002) point out that managers and early investors of issuers do not sell their own stakes at IPOs, but rather wait until the end of the lockup period, which is typically six months after IPO.

This chapter is closely related to Loughran and Ritter (2002) in spirit, (i) I take the stand that the underwriting business is not completely competitive, because there are significant barriers to entry, due to the limited supply of influential analysts; (ii) I explore the conflict of interest between underwriters and issuers. Since the IPO offer price is determined during the book building process, and underwriters have the discretion in share allocation, underwriters could intentionally leave more money on the table than necessary and allocate these shares to favored regular buy-side clients. The key is why issuers appear content and allow this to happen. Loughran and Ritter (2002) explain this using the prospect theory, which shows people care about the change of wealth, rather than the level of wealth. They argue that with great recent increase in wealth, founders and early investors spend little bargaining effort in their negotiations over the offer prices with underwriters. Instead, I posit the rational entrepreneur is willing to leave money on the table for IPO investors because he relies on the underwriter and IPO investors to have a successful IPO, in order to realize the higher firm value caused by public listing.[7] Hence the higher the total value gain, the larger the payoffs received by IPO investors.

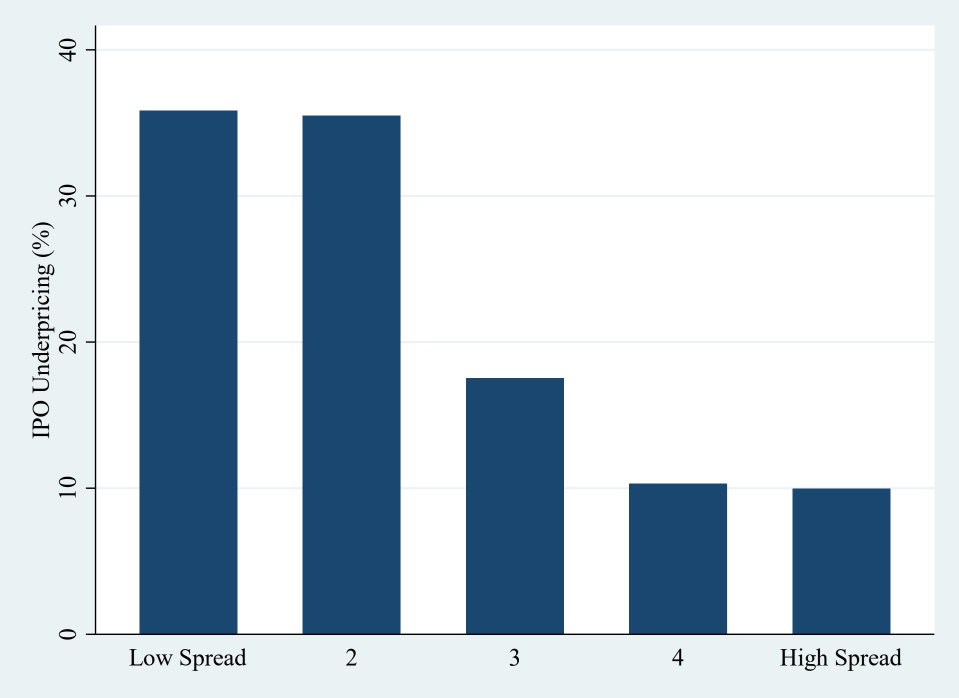
---

[6] In my sample period of 1981-2015, the average underpricing is 21.5%.

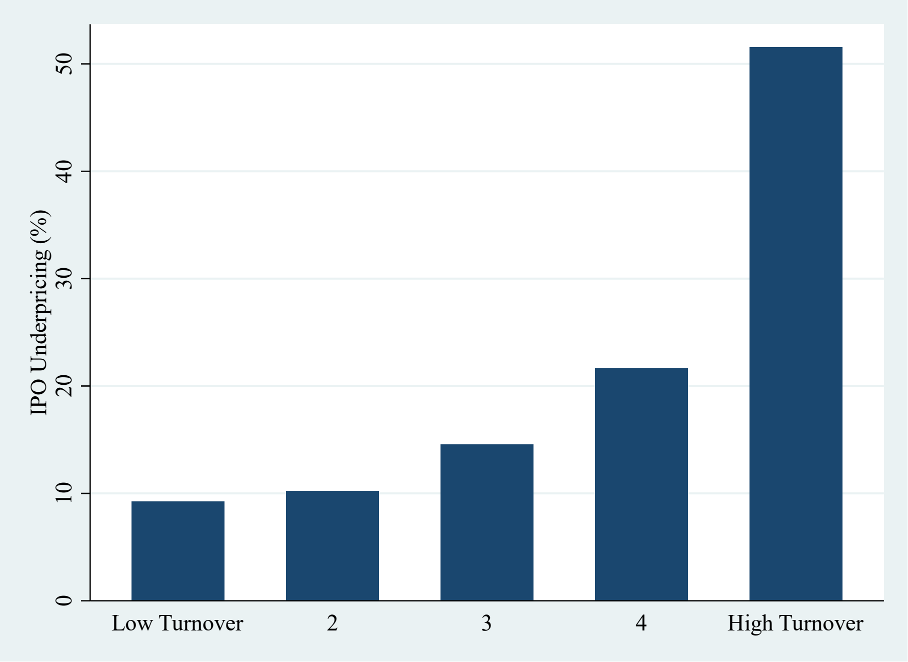
[7] Enough subscription interests from IPO investors are crucial for IPO success. Around 20% of IPO are withdrawn in the book-building process, and withdrawn issuers receive lower valuations when they return to the market later. See Dunbar and Foerster (2008), among others.

Admittedly, conditional on the fact that issuers choose to go public, the source of increase in firm value includes, but is not limited to, improved liquidity. For example, firms can benefit from less constraints in financing which enables them to invest in more positive-NPV projects, lower cost of capital with reduced information asymmetry, and/or attracting more faith from banks, consumers, and suppliers with its public status. I choose to focus on the aspect of liquidity only, because it is the most tractable and testable one of all the channels. If the liquidity value does not make up a significant part of the value gain, it will simply bias against me finding any results.

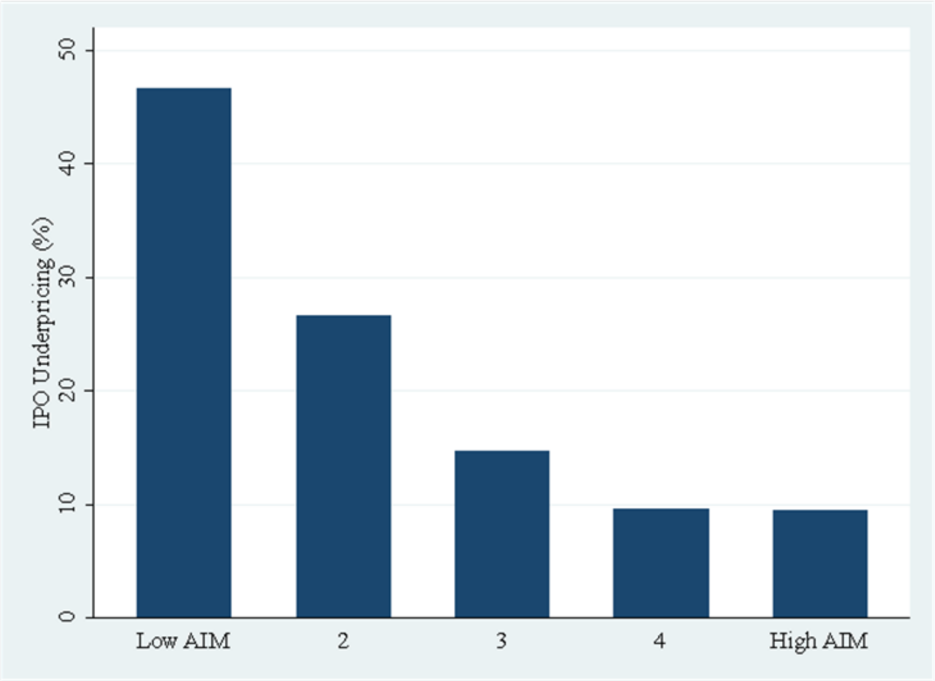
The degree of liquidity improvement is equal to the difference between the issuer's expected post-IPO liquidity and its pre-IPO liquidity. Empirically I test the relation between underpricing and the expected post-IPO liquidity (henceforth, "expected liquidity", for brevity) in the baseline regressions, since the pre-IPO liquidity cannot be directly measured. I measure expected liquidity by spread, turnover, and log AIM (Amihud's Illiquidity Measure) of the issuer's peer public firms in the 12 months before the issuer's IPO time (henceforth, referred to as Peer Spread, Peer Turnover, and Peer AIM). I use peer liquidity before the IPO time to measure expected liquidity because it wards off reverse causality concerns. An issuer's IPO price and its subsequent secondary market first-day price of are unlikely to affect its public peers' trading environment on the secondary market a year before the IPO time.



Panel A: Underpricing and Peer Spread



Panel B: Underpricing and Peer Turnover



Panel C: Underpricing and Peer AIM

Figure 2.1. Average Underpricing across Quintile Subsamples

The figure shows the average underpricing of deals in five quintiles in 1981-2015, sorted by Peer Spread, Peer Turnover, or Peer AIM. Peer firms are publicly traded firms that are in the same industry (SIC two-digit code), listed on the same exchange, and within the same size quintile by market capitalization in the COMPUSTAT-CRSP universe as the issuer. Peer Spread (Turnover) is the peer public firms' average daily spread (turnover) in the 12-month period prior to the issuer's IPO time. Peer AIM is the natural log of one plus peer public firms' average daily AIM in the 12-month period prior to the issuer's IPO time.

Figure 2.1 summarizes the key takeaway of the baseline regression. It plots the average underpricing across subsamples of issuers in five quintiles sorted by expected liquidity. Underpricing is monotonically increasing in expected liquidity across the five groups and the pattern is economically large. The difference in underpricing exceeds 25% between the lowest and highest quintile. I find similar results in multi-variate regressions, controlling for various deal and issuer characteristics, and time-varying industry effects. For instance, one standard deviation increase (decrease) in peer turnover (AIM) increases underpricing by 13.3% (9.6%), in absolute terms. These results hold in various robustness checks.

I then provide cross-sectional tests of the baseline specification, conjecturing that some issuers may benefit more from the improved liquidity compared to others. First, venture capital (hereafter, VC) funds typically have limited duration, and therefore have stronger needs to unload shares in their portfolio companies in time. Hence, I hypothesize that the relation between underpricing and expected liquidity is stronger for VC-backed issuers, compared to non-VC-backed issuers.[8] In the data, I find that the interaction term between the dummy variable of being VC-backed and expected liquidity has a positive and significant effect on underpricing. Second, my theoretical model shows that the relation between underpricing and expected liquidity should be stronger when the underwriter has more bargaining power or when the issuance size is smaller. I find consistent results supporting these predictions.

---

[8] Black and Gilson (1998) show that IPOs are not so much exits for founders as they are for VC investors, as founders often regain control from VC investors in VC-backed companies at IPO.

One alternative interpretation of the baseline results could be that some underlying unobserved variables drive both peer liquidity and underpricing simultaneously, even when reverse causality is ruled out. In addition, since I leave out the pre-IPO liquidity in the regression, it could cause estimation bias if expected liquidity is correlated with pre-IPO liquidity. Using peer liquidity as the measure for expected liquidity is unlikely to be the case. Still, to thoroughly address these concerns, I explore two regulation changes as exogenous shocks to firms' expected liquidity value, one to the expected post-IPO liquidity, and the other to the pre-IPO liquidity.

I use the enactment of changes in Order Handling Rules (hereafter, OHR) at Nasdaq in 1997 as an exogenous shock to expected liquidity of issuers listed on Nasdaq. Triggered by a scandal that Nasdaq dealers collude to maintain large bid-ask spread, SEC mandated several changes of OHR for all Nasdaq securities in 1997. Most changes aim to decrease quoted spread and promote trading. For example, one of the changes requires a limit order to be posted on the trading system if it is better than a dealer's quotes. These changes in OHR promoted the competition between limit orders and dealer quotes on Nasdaq and improved the liquidity of stocks listed on Nasdaq (see Bessembinder 1999, Barclay et al. 1999, among others). Meanwhile, other exchanges did not undergo a similar regulation change and IPO stocks listed on those exchanges were not subject to the same liquidity shock. More importantly, this change in liquidity of public stocks does not affect the pre-IPO liquidity should the firm stay private. Using a difference-in-difference approach, I show that Nasdaq IPOs exhibit more underpricing after changes to OHR compared to non-Nasdaq IPOs, due to their improved liquidity, and the magnitude is economically large. I find that, on average, underpricing of Nasdaq deals is less than that of non-Nasdaq deals by 5.6% before 1997, but higher by 15% after 1997, and both differences are quoted in absolute terms.[9]

Second, I use the passage of the National Securities Market Improvement Act (hereafter, NSMIA) in 1996 as an exogenous shock to the pre-IPO liquidity of issuers. Liquidity of public shares is measured by how easily transactions take place among existing shareholders and other investors, which is bid-ask spread or price impact of trades. However, since a secondary market for private shares does not widely exist, I measure pre-IPO liquidity by how easily founders can sell part of their ownership and raise capital, specifically by how large the pool of potential investors is.[10] When it is easier for founders to sell ownership in exchange for external equity capital, the private firm receives higher valuations, which is equivalent to the concept of smaller price impact for public shares. In the era before NSMIA, if a private firm raises capital in multiple states, it must comply with varying state-level disclosure and registration rules in each state. NSMIA lowers the regulatory barriers for private firms to raise capital from multiple states by exempting firms from those rules (Ewens and Farre-Mensa, 2017). Furthermore, NSMIA also expanded exemptions that VC and PE funds typically use to avoid registering with SEC and subsequent disclosures. It made it possible for these funds to manage more capital. Hence the adoption of NSMIA facilitates financing by private firms, broadens the pool of potential buyers for private shares, and enhances their liquidity.

---

[9] One potential concern is that this law change coincides with the tech bubble period of 1998-2000 and many tech stocks are listed on Nasdaq. I repeat the analysis excluding all tech firms and results remain. I follow the definition in Loughran and Ritter (2004) for tech firms.

[10] Most private firms have restrictions on shareholders selling their shares to a third party, for concerns of control rights. Normally the agreement between the company and shareholders (the stock issuance agreement) stipulates that the company has a right of first refusal over their shares. A right of first refusal requires shareholders to offer their shares to the company to purchase before they can sell those shares to the third party. See Belt (2018).

I observe that, even though NSMIA is a federal regulation impacting all private firms in US, firms located in states with scarce in-state capital should be affected more than firms located in states with abundant in-state capital. I regard issuers located in the former states as the treatment group, and issuers in the latter states as the control group, and conduct a difference-in-difference test. I find that the treatment group exhibit less underpricing than the control group post NSMIA, while controlling for expected liquidity. For example, when I use deals in the top four states with the largest number of VC and PE firms as the control group (CA, NY, MA, and TX) and deals in all other states as the treatment group, I find that underpricing of the treatment group is 14% lower than the control group post-NSMIA, in absolute terms. Overall, the evidence from the empirical analysis supports the model prediction that underpricing is positively related to the issuer's expected liquidity and negatively related to pre-IPO liquidity.

The traditional IPOs fundamentally serve two roles for the issuer: raising capital and improving the liquidity of its shares. Traditional theories in the IPO literature tends to focus on the first role but had little to say about the latter. However, given the recent trend of direct-listing promotion in the community of successful private firms and the ever-growing amount of capital under the management of VC and PE funds, I believe analyzing IPO phenomena only through the lens of raising capital is somewhat restricted. This chapter recognizes the liquidity improvement role via IPO and shows that it is relevant in determining a key aspect of IPO: underpricing.

The remainder of the chapter is organized as follows. Next section summarizes the IPO underpricing literature and discusses the relation of this chapter to previous studies. The following section develops hypotheses. The next section describes the sample and variable construction in the study. The next section describes empirical tests, presents main findings and conducts robustness checks. The final section concludes.

Relation to Previous Work

IPO underpricing is a longstanding puzzle in finance, and there is a large literature aiming to explain the phenomenon. Ritter and Welch (2002) provide an excellent and extensive review. The early studies focus on information asymmetry issues. For example, Rock (1986) models information asymmetry among investors, which causes the winner's curse for uninformed investors, and argue that average underpricing is necessary to attract uninformed investors to invest in IPOs. Other studies model information asymmetry between firm insiders and IPO investors, and view underpricing as a credible signal for firm quality, since high-quality firms can benefit from returning to the equity market for subsequent financing (Welch, 1989), receiving more favorable market reactions to future dividend news (Allen and Faulhaber, 1989), or attracting more attention to stimulate information production (Chemmanur, 1993). Benveniste and Spindt (1989) model information asymmetry between informed investors and the underwriter during the book-building process, and argue that underpricing is a compensation to informed investors for them to reveal private information. Some other explanations do not rely on information asymmetry, such as Hughes and Thakor (1992), which argue that issuers use underpricing to reduce legal liabilities.

This chapter is different from these studies in that I take the determination of IPO offer prices as a result of negotiation between the underwriter acting on behalf of their buy-side clients and the issuer. This approach is closely related to Loughran and Ritter (2002), which also focuses on the conflicts of interests between the underwriter and the issuer. This view is based on two bodies of empirical evidence. First, the IPO market is not completely competitive, otherwise the underwriter does not have any bargaining power. Krigman, Shaw, and Womack (2001) base their studies on questionnaires to firms and document that the two most important determinants of choosing underwriters are underwriter prestige and analyst coverage, instead of the magnitude of underpricing. Similarly, Loughran and Ritter (2002) argue that the perceived importance of analyst coverage allows high-prestige investment banks to attract issuers despite underpricing. Liu and Ritter (2011) show that the IPO market is characterized by local underwriter oligopolies, because issuers care about non-price dimensions such as all-star analyst coverage and industry expertise. Bradley et al. (2004) study whether the offer price is an integer and conclude that the office price is likely to be the result of negotiation and bargaining between issuers and IPO investors.

Second, the underwriter has an incentive to leave money on the table to benefit its buy-side clients. There is abundant evidence showing that underwriters and institutional IPO investors are no strangers to each other. Aggarwal, Prabhala, and Puri (2002) show that institutional investors get more allocations in IPOs with strong premarket demand. Cornelli and Goldreich (2001) document that regular investors receive favorable allocations, especially when the issue is heavily oversubscribed. Underwriters in general rely on the same group of investors for future deals, so the interaction between investment banks and IPO investors can be characterized as repeated games. Thus, it is beneficial for investment banks to "leave a good taste in investors' mouth". In addition, Loughran and Ritter (2002) point out that underwriters benefit from underpricing because investors engage in rent-seeking behavior to receive favorable allocations in hot IPOs. The observation that underwriters can act in the interests of IPO investors is also supported by the "partial adjustment phenomenon" during book building. Hanley (1993) first documents that underwriters do not fully adjust the offer price based on information collected when the demand is strong. Bradley and Jordan (2002), Loughran and Ritter (2002), and Lowry and Schwert (2004) all find that the offer price is not fully adjusted to reflect publicly available information either, proxied by recent market rally. So underwriters seem to leave money on the table on purpose.

This chapter contributes to the literature in several ways. First, I explore a novel factor determining the magnitude of underpricing: the size of the firm value gain via IPO. I test this channel by focusing on the liquidity value when the firm is transformed from a private to a public one. My empirical analysis shows that this is an economically significant factor for underpricing magnitude.

Second, this chapter deviates from several previous papers that study liquidity and underpricing, as they mostly focus on the issuer's realized liquidity in the secondary market and the theory is vastly different. For example, Booth and Chua (1996) model how the issuer's needs for ownership dispersion and secondary market liquidity jointly determine the equilibrium level of underpricing with asymmetric information. Issuers achieve broad ownership dispersion through over-subscription, which increases both the secondary-market liquidity and information costs borne by investors, who are compensated through larger underpricing. Their model predicts that more underpricing causes better secondary-market liquidity through the channel of diverse ownership. In their empirical analysis, they do not test the relation between underpricing and liquidity directly, but focus on the relation between underpricing and over-subscription. However, given that some IPO investors quickly turn around to sell their shares (dubbed as "flipping"), it is unclear how the ownership structure at IPO benefits control-seeking issuer executives and facilitates secondary-market trading.[11] Furthermore, the direction of causality studied in their chapter is the opposite to ours, and I carefully select peer liquidity as expected liquidity measures before the IPO time to rule out reverse causality concerns.

Ellul and Pagano (2006) hypothesize that deals with higher liquidity risk and less post-IPO liquidity need more underpricing to attract investors to participate, extending the line of research linking information asymmetry with underpricing. Their theory predicts a negative relationship between post-IPO liquidity and underpricing, that is, shares that are expected to be illiquid should command more underpricing. This is opposite to the hypothesis in this chapter. They use IPO data from UK and confirm the negative relation, by documenting a positive relation between underpricing and the PIN variable measuring asymmetric information. My study helps shed light on this relation by focusing on expected liquidity of the issuer and using exogenous shocks to post-IPO expected liquidity and pre-IPO liquidity for identification. My empirical finding is opposite to Ellul and Pagano (2006), possibly due to the fact that they use UK data and focus on the relation between underpricing and liquidity risk rather than liquidity value.[12]

---

[11] Field and Sheehan (2004) find that the link between underpricing and ownership structure is weak.

Lastly, this chapter adds to the literature on why firms go public. Zingales (1995) models the profit maximization problem of an entrepreneur selling his stake in the firm, and argues that the public status enables him to maximize proceeds from selling cash flow rights and the private status enables him to maximize proceeds from selling control rights. Hence the decision to go public depends on the initial ownership structure. Chemmanur and Fulghieri (1999) develop a life cycle theory of the going public decision based on the pros and cons of diverse ownership brought by the public status. Maksimovic and Pichler (2001) point out that public trading can inspire more faith in the firm from investors, customers, creditors, and suppliers. There are also market-timing theories that firms go public when market valuations are high (Lucas and McDonald, 1990, among others). I argue that IPOs provide issuers access to liquidity traders in the secondary market and improve the firm's liquidity, which increases firm value and decreases the price impact of investment exit for founders and early investors.

---

[12] For my own curiosity, I download the PIN measure for information asymmetry from Prof. Stephen Brown's website: http://scholar.rhsmith.umd.edu/sbrown/pin-data?destination=node/998 and replicate Ellul and Pagano (2006) with US IPO data. I find the opposite result: the relationship between underpricing and PIN is negative, instead of positive.

Hypothesis Development

The liquidity value is generated by the difference between the issuer's expect liquidity and its pre-IPO liquidity. Expected liquidity is the market's consensus expectation at the time of IPO for the issuer's post-IPO liquidity. Formally, I test the following two hypotheses.

H1: The magnitude of IPO underpricing should be positively related to the expected liquidity of the issuer, ceteris paribus.

H2: The magnitude of IPO underpricing should be negatively related to the pre-IPO liquidity of the issuer when it is private, ceteris paribus.

The expected liquidity can be measured using the "comparables" approach, while the issuer's pre-IPO liquidity as a private firm is not observable. In the next two sections, I describe in detail the data and empirical methodology employed to test these two hypotheses.

Sample and Variable Construction

I start sample construction by identifying US firm-commitment IPOs from Thomson Financial's SDC Global New Issues database from 1981 to 2015.[13] Following the literature, I exclude deals with offer prices less than $5, unit offerings, ADRs, financial and utility offerings (SIC codes 6000–6999 and 4900–4999), certificates, shares of beneficial interest, companies incorporated outside the U.S., Americus Trust components, closed-end funds, REITs, and limited partnerships. Variables related to the deal and issuer characteristics (pre-issue assets, offer price, underwriters, VC-backed or not, and proceeds) are from SDC. Stock price and liquidity data (first-day closing price, bid price, ask price, number of shares outstanding, stock exchange, and trading volume) are from CRSP. I use underwriter ranking data on Prof. Jay Ritter's website. The monthly market sentiment data are from Prof. Jeffrey Wurgler's website (Baker and Wurgler 2006). I download the Fama-French 10-industry classification from Prof. Kenneth French's website. The final sample consists of 3,775 deals.

---

[13] Some IPO studies examine both firm-commitment and best-effort issues (see Ritter, 1987; Booth and Chua, 1996) Meanwhile, many other IPO studies only examine firm-commitment issues (Carter and Manaster, 1990; Loughran and Ritter, 2004, among others). I exclude best-effort offers because best-effort offers are typically very small offerings.

Table 2.1. Distribution of IPOs by Year, Industry, and Exchange

The table shows sample distribution of IPOs in US across year, industry, and exchange in 1981-2015. Panel A shows the distribution by year and exchange, and Panel B shows the distribution by year and industry. The exchanges include the New York Stock Exchange (NYSE), American Stock Exchange (ASE), and Nasdaq. Industries are defined by the Fama-French 10-industry classification. Nine industries are presented because I exclude utilities firms. CN, CD, M, E, BE, T, W, H, O stands for Consumer Non-durables, Consumer Durables, Manufacturing, Energy, Business Equipment, Telecommuniciatons, Wholesale and Retail, Healthcare, Others.

Panel A: Distribution by year and exchange

Year	**NYSE**	**ASE**	**Nasdaq**	**Total**	**%**
1981	0	0	3	3	0.1%
1983	0	1	4	5	0.1%
1984	0	0	2	2	0.1%
1985	0	1	12	13	0.3%
1986	8	16	166	190	5.0%
1987	6	12	129	147	3.9%
1988	5	5	41	51	1.4%
1989	4	3	44	51	1.4%
1990	3	3	39	45	1.2%
1991	20	3	112	135	3.6%
1992	20	1	185	206	5.5%
1993	19	1	233	253	6.7%
1994	15	3	202	220	5.8%
1995	13	5	212	230	6.1%
1996	26	11	352	389	10.3%
1997	31	5	229	265	7.0%
1998	29	4	147	180	4.8%
1999	15	3	331	349	9.2%
2000	10	2	230	242	6.4%
2001	9	2	27	38	1.0%
2002	13	1	27	41	1.1%
2003	7	1	28	36	1.0%
2004	16	3	79	98	2.6%
2005	22	5	49	76	2.0%
2006	10	3	56	69	1.8%
2007	15	0	58	73	1.9%
2008	3	0	4	7	0.2%
2009	6	0	10	16	0.4%
2010	14	2	29	45	1.2%
2011	16	1	27	44	1.2%
2012	23	0	24	47	1.2%
2013	27	0	45	72	1.9%
2014	19	1	66	86	2.3%
2015	13	0	38	51	1.4%
Total	437	98	3,240	3,775	
%	11.6%	2.6%	85.8%		

Panel B: Distribution by year and industry

Year	CN	C	M	E	BE	T	W	H	O	Total	%
1981	0	0	0	0	2	0	1	0	0	3	0.1%
1983	0	1	1	0	0	0	1	0	2	5	0.1%
1984	0	0	0	0	0	0	2	0	0	2	0.1%
1985	3	0	1	0	4	0	4	1	0	13	0.3%
1986	17	9	32	0	38	8	43	16	27	190	5.0%
1987	13	1	26	2	32	5	26	16	26	147	3.9%
1988	4	2	13	0	12	2	9	6	3	51	1.4%
1989	2	0	6	2	15	1	6	8	11	51	1.4%
1990	2	2	2	4	10	0	7	13	5	45	1.2%
1991	9	5	10	2	38	3	24	38	6	135	3.6%
1992	16	8	10	2	43	9	47	54	17	206	5.5%
1993	18	12	41	9	56	14	42	27	34	253	6.7%
1994	9	9	29	4	61	14	35	26	33	220	5.8%
1995	9	3	20	2	98	11	22	36	29	230	6.1%
1996	16	6	21	9	145	14	55	64	59	389	10.3%
1997	17	4	25	4	100	5	33	40	37	265	7.0%
1998	13	2	12	2	66	12	28	16	29	180	4.8%
1999	8	3	5	1	212	35	23	11	51	349	9.2%
2000	0	1	6	2	137	20	10	42	24	242	6.4%

2001	1	0	4	1	13	0	3	11	5	38	1.0%
2002	2	0	2	1	13	1	7	7	8	41	1.1%
2003	0	1	2	1	12	2	4	7	7	36	1.0%
2004	2	3	5	4	28	3	10	31	12	98	2.6%
2005	5	2	7	5	21	4	8	18	6	76	2.0%
2006	3	2	8	2	17	4	4	21	8	69	1.8%
2007	1	1	7	6	22	2	4	21	9	73	1.9%
2008	0	0	3	0	2	0	0	1	1	7	0.2%
2009	1	0	0	0	3	0	4	4	4	16	0.4%
2010	0	2	4	0	11	2	5	11	10	45	1.2%
2011	0	0	2	3	17	0	7	11	4	44	1.2%
2012	1	1	5	4	15	1	10	8	2	47	1.2%
2013	1	1	5	2	21	0	6	27	9	72	1.9%
2014	1	1	1	2	22	0	6	47	6	86	2.3%
2015	2	0	1	0	8	0	5	33	2	51	1.4%
Total	176	82	316	76	1,294	172	501	672	486	3,775	
%	4.7%	2.2%	8.4%	2.0%	34.3%	4.6%	13.3%	17.8%	12.9%		

Panel A of Table 2.1 presents the distribution of IPO deals by listing venues. There are three exchanges that issuer stocks are listed on: the New York Stock Exchange (NYSE), the American Stock Exchange (ASE), and Nasdaq. Nasdaq IPOs account for 85.8% of the whole sample, while NYSE deals account for 11.6%. The imbalance across exchanges is more severe during the tech bubble period in the 1990s. Over time, there are hot IPO periods such as the 1990s, and cold years such as 2001, 2002, and 2003, in the aftermath of bubble burst. Panel B shows the distribution of IPOs in each industry over time.[14] Consistent with stylized facts, the two industries with largest numbers of IPOs in the sample are Business Equipment (Computers, Software, and Electronic Equipment) and Healthcare. They each account for 34.3% and 17.8% of all deals.

The key variables in the baseline analysis are underpricing and expected liquidity. Underpricing is the percentage change from the offer price to the first-day closing price. Liquidity is measured by spread, turnover, or log AIM. Spread is the difference between ask price and bid price divided by the average of the two prices (mid-point price). Turnover is the daily trading volume divided by the number of shares outstanding. AIM is the Amihud's (2002) illiquidity measure, which is the absolute value of daily returns divided by daily dollar volumes, scaled by 10,000,000. Peer firms are defined by industry, size, and the listing stock exchange.[15] I divide the COMPUSTAT-CRSP universe of stocks into five quintiles by market capitalization each year, and select publicly traded stocks that are in the same size quintile, with the same two-digit SIC code, and listed on the same stock exchange as the issuer. For each issuer, its expected liquidity is defined as the average daily liquidity measures across its peer firms over a 12-month window before the IPO date, denoted by Peer Spread, Peer Turnover, and Peer AIM respectively. These measures are observable to the issuer and investors during book-building when the offer price is determined, so there is no look-ahead bias. In addition, they are unlikely to cause reverse causality concerns, because the issuer's offer price and its subsequent first-day close price on the secondary market should not affect its peer public firms' trading environment before its IPO time.

---

[14] Nine industries are presented because Industry nine (Utilities) are excluded.

[15] A common practice to identify peer firms is to use just industry and size (Albuquerque, 2009, among others). My results are robust to constructing peer firms only by industry and size. I include listing exchanges because some studies suggest that different institutional

The control variables are as follows. I use Top Underwriter as a proxy for bargaining power, which is a dummy variable that is one if the lead underwriter has an updated Carter and Manaster (1990) rank of eight or more, and zero otherwise. The literature also document a significant relation between the prestige of the underwriter and underpricing (see Carter and Manaster, 1990; Beatty and Welch ,1996; Liu and Ritter, 2011, among others). Habib and Ljungqvist (2001) propose that the dilution of shares as a result of new issuance should matter for IPO underpricing. To measure issuance size, I define New Shares Ratio as the percentage of IPO shares in the firm's total number of shares outstanding. Computationally, it is equivalent to the ratio of IPO proceeds divided by the product of offer price multiplied by the number of shares outstanding. Bradley et al.

---

designs in different stock exchanges can affect liquidity of listed stocks (see Huang and Stoll, 1996; Bessembinder and Kaufman, 1997).

(2004) hypothesize that the integer versus fractional dollar offer prices are results of negotiations between issuers and underwriters. They find that deals with integer offer prices are associated with more underpricing. I include the dummy variable of Integer Price. To control for investor sentiment, I include Sentiment in the regression, which is equal to the monthly market sentiment index downloaded from Prof. Jeffrey Wurgler's website at IPO time.[16] The literature has shown that whether the issuer is backed by VC funds is related to underpricing (see Lee and Wahal, 2004, among others). I include the dummy variable VC-backed as a control. Finally, I also control for the issuer's asset size and age. More details of variable construction can be found in Appendix. All variables are winsorized at the 1st and 99th percentile level.

---

[16] Lee, Shleifer, and Thaler (1991) argue that individual investor sentiment is an important factor that determines when companies go public. Ljungqvist, Nanda, and Singh (2006) hypothesize that regular investors sell IPO stocks to sentiment investors.

Table 2.2. Summary statistics

The table provides summary statistics of variables. Panel A shows the distribution statistics, and Panel B presents the correlation matrix. In Panel B, Pearson correlations are below diagonal and Spearman correlations are above diagonal. Variables construction is described in Appendix. I winsorize all variables at the 1st and 99th percentile levels.

Panel A: Distribution statistic

	N	Mean	Median	Std. Dev.	P25	P75
IPO Underpricing	3,775	21.5%	8.7%	38.8%	0.0%	25.4%
Peer Spread	3,658	3.50%	2.90%	3.00%	1.30%	4.90%
Peer Turnover	3,773	0.80%	0.70%	0.50%	0.40%	1.00%
Peer AIM	3,773	1.83	1.47	1.60	0.36	2.89
Top Underwriter	3,775	60.7%	1	48.8%	0	1
Integer Price	3,775	82.1%	1	38.3%	1	1
New Shares Ratio	3,775	32.2%	28.9%	17.0%	21.6%	38.0%
Sentiment	3,766	0.32	0.31	0.46	-0.06	0.63
Assets	3,775	183.65	29.70	577.74	11.00	90.10
Age	3,775	14.83	8.00	19.75	4.00	16.00
VC-backed	3,775	44.7%	0	49.7%	0	1
Tech	3,775	0.392	0	0.488	0	1
Price Revision	3,775	-0.008	0	0.148	-0.033	0
Inverse Price	3,775	0.094	0.083	0.039	0.067	0.111
30-day Underpricing	3,775	0.235	0.151	0.253	0.092	0.246

15-day Market Return	3,775	0.016	0.018	0.035	-0.003	0.037
Analyst Coverage	3,775	2.543	2	1.782	1	3.2

Panel B: Correlation matrix

	IPO Under-pricing	Peer Spread	Peer Turnover	Peer AIM	Top Under-writer	Integer Price	New Shares Ratio	Sentiment	Log (Assets)	Log (1+Age)	VC
IPO Underpricing	1.00	-0.22	0.31	-0.32	0.12	0.13	-0.21	-0.09	-0.01	-0.14	0.15
Peer Spread	-0.23	1.00	-0.60	0.89	-0.46	-0.18	0.41	0.40	-0.54	-0.08	-0.24
Peer Turnover	0.47	-0.53	1.00	-0.59	0.29	0.20	-0.39	-0.20	0.15	-0.12	0.34
Peer AIM	-0.30	0.90	-0.55	1.00	-0.53	-0.18	0.48	0.25	-0.58	-0.06	-0.20
Top Underwriter	0.16	-0.47	0.27	-0.54	1.00	0.12	-0.28	-0.14	0.43	0.05	0.25
Integer Price	0.15	-0.15	0.20	-0.17	0.12	1.00	-0.10	-0.08	0.05	-0.04	0.12
New Shares Ratio	-0.21	0.29	-0.32	0.33	-0.16	-0.08	1.00	0.06	-0.14	0.16	-0.25
Sentiment	-0.09	0.28	-0.14	0.18	-0.12	-0.07	0.01	1.00	-0.16	-0.03	-0.11
Log(Assets)	-0.01	-0.49	0.16	-0.55	0.44	0.05	0.01	-0.14	1.00	0.34	-0.10
Log(1+Age)	-0.17	-0.08	-0.12	-0.07	0.08	-0.03	0.16	-0.04	0.40	1.00	-0.23
VC-backed	0.19	-0.25	0.32	-0.24	0.25	0.12	-0.24	-0.09	-0.12	-0.23	1.00

Table 2.2 presents summary statistics of variables. Panel A describes the distribution of each variable and Panel B is the correlation matrix. The average underpricing is 21.5%, with a standard deviation of 38.8%. The average daily spread and turnover of public peer firms are 3.50% and 0.80% in the 12-month period before the issuer's IPO time. About 60.7% of IPO underwriters are large investment banks with the Top Underwriter status. 82.1% of IPOs have integer offer prices. On average, the issuer sells 32.2% of the ownership to raise capital in IPOs. The average sentiment is 0.32. The average asset value of issuers is $184 million, with a large variation as its standard deviation is $578 million. The average age of issuers is 14.8 years. Around 44.7% of sample issuers are backed by VC funds.

Panel B shows that underpricing has negative correlations with Peer Spread (-0.23 for Pearson correlation, and -0.22 for Spearman correlation), positive correlations with Peer Turnover (0.47 for Pearson correlation, and 0.31 for Spearman correlation), and negative correlations with Peer AIM (-0.30 for Pearson correlation, and -0.32 for Spearman correlation). Additionally, underpricing is positively correlated with Top Underwriter, Integer Price, and VC-backed, and negatively correlated with New Shares Ratio, Sentiment, and Age.

Empirical Tests and Results

In this section, I test the two hypotheses developed previous section. First, I develop a baseline regression testing the relation between underpricing and expected liquidity. I then test the cross-sectional variation of this relation based on certain firm characteristics.

Next, I use changes to OHR at Nasdaq in 1997 as an exogenous shock to expected liquidity of Nasdaq deals, so I can study the effect of expected liquidity on underpricing in a difference-in-difference setting. Lastly, I use NSMIA as an exogenous shock to the pre-IPO liquidity of issuers to test how pre-IPO liquidity affects underpricing, while controlling for expected liquidity.

The Baseline

The baseline regression tests the hypothesis of H1 directly by investigating whether there is a positive relation between underpricing and the issuer's expected liquidity. I regress underpricing on expected liquidity and control variables. The regression equation is specified as follows,

$$Underpricing_{ijt} = \alpha_{jt} + \beta_1 Expected\ Liquidity_{ijt} + \gamma' X_{ijt} + \varepsilon_{ijt} \qquad (1)$$

where i,j,t index firms, industries, and years, respectively. X_{ijt} is a vector of control variables, explained in the last section. H1 predicts that $\beta_1 > 0$. I use public peers' liquidity measures Peer Spread, Peer Turnover, and Peer AIM to proxy for the expected liquidity of the issuer. These measures can ward off reverse causality concerns, as it is unlikely that the offer price and the later secondary market first-day price of an issuer can affect its public peers' trading environment on the secondary market a year before IPO time.

Table 2.3. Baseline: IPO Underpricing and Expected Liquidity

The table reports coefficients and t-statistics in the parenthesis of OLS regressions of underpricing on the issuer's expected liquidity and control variables, specified by the regression Equation (1). Expected liquidity is measured by Peer Turnover in Columns (1) and (2), by Peer Spread in Columns (3) and (4), and by Peer AIM in Columns (5) and (6). Spread (Turnover) is the peer public firms' average daily spread (turnover) in the 12-month period prior to the issuer's IPO time. Peer AIM is the natural log of one plus peer public firms' average daily AIM in the 12-month period prior to the issuer's IPO time. Variable construction is described in Appendix. The industry-year fixed effects control for time-varying industry effects. Industries are defined by the Fama-French 10-industry classification. The t-statistics are computed based on standard errors adjusted for heteroskedasticity and industry-year clustering. Asterisks denote statistical significance at 1% (***), 5% (**), or 10% (*) level.

	Peer Spread		Peer Turnover		Peer AIM	
	(1)	(2)	(3)	(4)	(5)	(6)
Liquidity	**-0.568****	**-0.967*****	**20.827*****	**21.980*****	**-0.034*****	**-0.037*****
	(-2.27)	**(-2.82)**	**(3.73)**	**(2.74)**	**(-5.53)**	**(-5.60)**
Underwriter	-0.004	-0.006	-0.002	-0.006	-0.007	-0.009*
	(-0.85)	(-1.30)	(-0.51)	(-1.12)	(-1.43)	(-1.81)
New Shares	-0.227***	-0.179***	-0.148***	-0.126***	-0.173***	-0.133***
	(-4.23)	(-4.03)	(-4.74)	(-3.86)	(-3.35)	(-3.11)
Integer Price	0.054***	0.045***	0.038***	0.040***	0.049***	0.043***
	(5.98)	(4.83)	(4.43)	(4.81)	(5.70)	(4.76)
Sentiment	0.015	-0.074	0.011	-0.069	0.013	-0.080
	(0.69)	(-1.21)	(0.75)	(-1.18)	(0.58)	(-1.37)

Log(Assets)	-0.031***	-0.028***	-0.031***	-0.025***	-0.037***	-0.033***
	(-4.33)	(-5.15)	(-4.31)	(-5.02)	(-5.01)	(-6.20)
Log(1+Age)	-0.030***	-0.032***	-0.024***	-0.028***	-0.027***	-0.030***
	(-4.34)	(-4.35)	(-4.83)	(-4.94)	(-4.09)	(-4.07)
VC-backed	0.013	0.043**	-0.008	0.032**	0.010	0.043**
	(0.72)	(2.18)	(-0.49)	(2.00)	(0.58)	(2.24)
Tech	0.060***	0.040*	0.026	0.037	0.054***	0.039*
	(3.29)	(1.86)	(1.58)	(1.51)	(3.10)	(1.83)
Price Revision	0.377***	0.355***	0.352***	0.340***	0.340***	0.318***
	(11.04)	(10.30)	(11.80)	(11.90)	(9.37)	(9.11)
Inverse Price	-1.428***	-1.110**	-1.252***	-1.056***	-1.130**	-0.863*
	(-2.98)	(-2.13)	(-3.17)	(-2.63)	(-2.31)	(-1.66)
30-day prior	0.518***	0.220**	0.386***	0.204**	0.502***	0.219**
	(7.44)	(2.50)	(9.44)	(2.42)	(7.20)	(2.47)
Market Return	1.198***	1.547***	1.302***	1.543***	1.212***	1.541***
	(5.39)	(5.84)	(4.91)	(5.97)	(5.47)	(5.92)
Analyst	0.022***	0.030***	0.012***	0.025***	0.018***	0.027***
	(3.96)	(3.45)	(3.40)	(3.64)	(3.36)	(3.13)
Constant	0.377***	0.337***	0.220***	0.155*	0.428***	0.394***
	(5.69)	(4.21)	(3.95)	(1.71)	(6.33)	(5.13)
Industry-Year Fixed Effect	NO	YES	NO	YES	NO	YES

Observations	3,649	3,649	3,764	3,764	3,764	3,764
Adjusted R2	0.307	0.331	0.342	0.357	0.315	0.339

To further address endogeneity concerns, I control for any unobserved time-varying industry effects that can affect both underpricing and expected liquidity by including the industry-year fixed effects α_{jt}.[17] To address for potential correlations of error terms among deals in the cohort of the same industry and year, I cluster standard errors in every industry-year group.

The results of regression Equation (1) are presented in Table 2.3. Expected liquidity is measured by Peer Spread in Columns (1) and (2), by Peer Turnover in Columns (3) and (4), and by Peer AIM in Columns (5) and (6). For each measure, I run the regression with two specifications: with and without the interacted industry-year fixed effects α_{jt}. In all specifications, underpricing is found to be positively related to expected liquidity, as the coefficients on Peer Spread and Peer AIM are negative, and the coefficient on Peer Turnover is positive. The coefficients are statistically and economically significant across all specifications. In Column (2), one standard deviation increase in peer spread (3%) reduces underpricing by 6.5% (=3%×-2.152) in absolute terms. Considering that the sample average underpricing is 21.5%, this is a 30% reduction from the mean (=6.5/21.5). In Column (4), one standard deviation increase in peer turnover (0.5%) increases underpricing by 13.3% (=0.5%×26.584), which is a 61.9% increase from its mean (=13.3/21.5). In Column (6), one standard deviation increase in peer AIM (1.6) reduces underpricing by 9.6% (=1.6×-0.06), which is a 44.7% reduction from the mean (=9.6/21.5).

---

[17] I use Fama-French 10-industry classification, since with the industry-year interaction fixed effects, using Fama-French 48-industry classification generates too many explanatory variables and lowers the power of the test.

The results on expected liquidity are robust to the inclusion of industry-year fixed effects, suggesting that underpricing is higher when an issuer has better expected liquidity than peer issuers in the same industry in the same IPO year.

I also find supporting evidence for predictions regarding relations between underpricing and the underwriter's bargaining power and the issuance size. Top Underwriter is positively related to underpricing, and its coefficient is statistically significant in all specifications. New Shares Ratio is negatively and significantly related to underpricing, that is, the larger the fraction of the ownership an issuer sells at IPO, the smaller the underpricing. Results on other control variables are also consistent with existing literature. When the offer price is an integer, there is more underpricing. Investor sentiment is found to be negatively related to underpricing. In most specifications, larger and older issuers are associated with lower underpricing, which is likely due to less asymmetric information, or more bargaining power of the issuer. VC-backed IPOs have higher underpricing compared to non-VC-backed ones.

I conduct the following robustness checks to the baseline and the results remain. (i) I use the issuer's own secondary-market liquidity measures (Issuer Spread, Issuer Turnover, and Issuer AIM) in the 12-month period following IPO as proxies for expected liquidity.[18] (ii) I use alternative time horizons of six months prior to the IPO date, when I construct peer liquidity measures. (iii) I follow the definition of tech firms in Loughran and Ritter (2004), and find that about 42.4% of the deals are tech stocks.[19] I exclude these deals from the sample, to address the possibility that one particular industry drives the results, or that the results are driven by the tech-bubble period. The results are robust to all the above alternative specifications, and shown in row (1), (2), and (3) of Table 2.4.

---

[18] Using the issuer's realized liquidity post-IPO assumes that the issuer and the underwriter can predict its post-IPO liquidity on the secondary market accurately. Peer liquidity and issuer liquidity measures are highly correlated. For example, the correlation between *Peer Spread* and *Issuer Spread* is 0.73 and they have similar magnitudes. All results with the baseline remain and some are stronger with these alternative measures. Admittedly, using the issuer's own realized liquidity involves look-ahead bias and is likely to generate

Next, I investigate whether the relation between underpricing and expected liquidity holds in sub-periods of the sample. Since there are many more deals in the 1990s than 1980s and 2000s, I divide the sample period of 1981-2015 to four sub-periods: 1981-1990, 1991-1995, 1996-2000, and 2000-2015, for a roughly even distribution of deal numbers across the subsamples. There are 505, 1044, 1425, 790 deals in the four sub-periods. Row (4) to (7) of Table 2.4 show the results and confirm the same finding in the baseline with the full sample. It suggests that the relation does not just exist in a particular sub-period.[20]

---

endogeneity concerns. That's why I do not include this analysis in the main text of the chapter.

[19] In Loughran and Ritter (2004), tech stocks are defined as those in SIC codes 3571, 3572, 3575, 3577, 3578 (computer hardware), 3661, 3663, 3669 (communication equipment), 3671, 3672, 3674, 3675, 3677, 3678, 3679 (electronics), 3812 (navigation equipment), 3823, 3825, 3826, 3827, 3829 (measuring and controlling devices), 3841, 3845 (medical instruments), 4812, 4813 (telephone equipment), 4899 (communications services), and 7371, 7372, 7373, 7374, 7375, 7378, and 7379 (software).

[20] Across the 12 specifications (with four subsamples and three liquidity measures), only in row (7) and column (2), the coefficient for Peer Turnover is statistically insignificant with a t-statistics value of 1.37.

Table 2.4. Robustness Analysis of the Baseline Regression

The table reports coefficients and t-statistics in the parenthesis of OLS regressions, which are robustness analysis of the baseline, specified by the regression Equation (1). In row (1), I use the issuers' own secondary-market liquidity in the 12-month period post IPO to proxy for Expected Liquidity. In row (2), I use the peer liquidity measures in the 6-month period (instead of the 12-month period in the baseline) before the issuer's IPO to proxy for Expected Liquidity. In row (3), I exclude all technology deals from the sample, which account for 42% of the total sample. Technology firms are defined in Loughran and Ritter (2004), and the details are described in Footnote 19. From row (4) to row (7), I show results of subperiod analysis by dividing the sample period of 1981-2015 to four sub-periods: 1981-1990, 1991-1995, 1996-2000, and 2000-2015, for a roughly even distribution of deal numbers across the subsamples. For the sake of brevity, estimation results for control variables including Top Underwriter, New Shares Ratio, Integer Price, Sentiment, Log(Assets), Log(1+Age), VC-backed and the intercept are not presented. Variable construction is described in Appendix. Industries are defined by the Fama-French 10-industry classification. The industry-year fixed effects control for time-varying industry effects. The t-statistics are computed based on standard errors adjusted for heteroskedasticity and industry-year clustering. Asterisks denote statistical significance at 1% (***), 5% (**), or 10% (*) level.

Row	Robustness to the baseline regression	Spread	Turnover	AIM
		(1)	(2)	(3)
(1)	Expected liquidity measured by the issuer's own secondary-market	-2.662***	18.236***	-0.064***
	liquidity in the 12-month period post IPO	(-5.29)	(7.90)	(-6.97)
(2)	Expected liquidity measured by peer liquidity in the 6-month period	-0.868***	18.189***	-0.029***
	(instead of the 12-month period) before the issuer's IPO	(-2.63)	(3.02)	(-3.84)
(3)	Excluding 42% of the deals in the sample, which are technology firms	-0.976***	10.059*	-0.034***
	defined in Loughran and Ritter (2004), # of observations = 2214	(-3.20)	(1.67)	(-5.85)
	Sub-period Anaysis:			
(4)	The baseline regression is run in the subsample during 1981-1990	-2.366***	3.909	-0.037***

	# of observations = 505	(-3.23)	(0.42)	(-3.66)
(5)	The baseline regression is run in the subsample during 1991-1995	-1.223***	7.374*	-0.029***
	# of observations = 1044	(-4.67)	(2.01)	(-4.73)
(6)	The baseline regression is run in the subsample during 1996-2000	-1.004	35.314***	-0.043***
	# of observations = 1425	(-1.07)	(3.21)	(-2.73)
(7)	The baseline regression is run in the subsample during 2001-2015	-6.793***	4.052	-0.058***
	# of observations = 790	(-2.67)	(1.29)	(-3.72)
	Controls	YES	YES	YES
	Industry-Year FE	YES	YES	YES

Cross-sectional Analysis of the Baseline

Having established the positive relationship between underpricing and expected liquidity, in this section I explore cross-sectional variations in this relation. VC funds typically have limited investment horizon (mostly up to ten years) and therefore often have stronger incentives to exit their investments compared to founders or employees after IPO. Better liquidity of the issuers' shares in the secondary market is particularly valuable for venture capitalists, because the unwinding of their positions would generate less price impact and higher payoffs. Hence, I expect a stronger relationship between expected liquidity and underpricing for VC-backed issuers than non-VC-backed issuers. I test this conjecture by adding the interaction term of the dummy variable of VC-backed and expected liquidity in the baseline regression.

I test the cross-sectional predictions in a regression by interacting expected liquidity with Top Underwriter and New Shares Ratio, respectively. I expect a positive coefficient on the interaction term of Top Underwriter and expected liquidity.

Table 2.5. Cross-sectional Analysis of the Baseline

The table reports coefficients and t-statistics in the parenthesis of OLS regressions, which are cross-sectional analysis of the baseline, by adding interaction terms of the dummy variable VC-backed, Top Underwriter, or New Shares Ratio with expected liquidity to Equation (1). Expected liquidity is Peer Spread, Peer Turnover, or Peer AIM. For the sake of brevity, estimation results for control variables including Integer Price, Sentiment, Log(Assets), Log(1+Age), and the intercept are not presented. Variable construction is described in Appendix. Industries are defined by the Fama-French 10-industry classification. The industry-year fixed effects control for time-varying industry effects. The t-statistics are computed based on standard errors adjusted for heteroskedasticity and industry-year clustering. Asterisks denote statistical significance at 1% (***), 5% (**), or 10% (*) level.

	Expected Liquidity Measures								
	Peer Spread			Peer Turnover			Peer AIM		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Liquidity	-0.354	3.205***	-3.768***	9.890**	-31.078***	45.650***	-0.034***	-0.029***	-0.124***
	(-0.89)	(3.26)	(-3.84)	(1.99)	(-5.60)	(4.08)	(-5.39)	(-5.71)	(-5.35)
VC	0.177***	0.040**	0.042**	-0.160***	0.029*	0.028	0.197***	0.051**	0.038*
	(3.83)	(2.09)	(2.19)	(-2.67)	(1.74)	(1.57)	(4.11)	(2.22)	(1.81)
Liquidity × VC	**-3.979*****			**24.426*****			**-0.087*****		
	(-3.70)			**(3.59)**			**(-4.32)**		
Underwriter	-0.007	0.033***	-0.008	-0.005	-0.047***	0.042***	0.012	0.174***	0.007
	(-1.42)	(4.04)	(-1.54)	(-1.07)	(-3.76)	(2.91)	(0.97)	(4.59)	(0.62)
Liquidity × Underwriter		**-0.875*****			**6.708*****			**-0.086*****	
		(-4.37)			**(5.28)**			**(-4.47)**	
New Shares	-0.156***	-0.119***	-0.404***	-0.116***	-0.110***	0.353***	-0.064**	-0.040	-0.400***

	(-4.09)	(-3.40)	(-3.67)	(-3.33)	(-3.38)	(2.64)	(-2.40)	(-1.44)	(-3.73)
Liquidity × New Shares			6.414***			-69.112***			0.168***
			(3.00)			**(-4.11)**			**(3.58)**
Industry-Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	3,649	3,649	3,649	3,764	3,764	3,764	3,764	3,764	3,764
Adjusted R2	0.345	0.349	0.338	0.373	0.373	0.315	0.306	0.302	0.296

The regression results are presented in Table 2.5. For the sake of brevity, regression results on control variables including the dummy variable of integer price, sentiment, asset size, age, and the constant term are not presented in the table. I continue to find strong support for hypothesis H1 across all three liquidity measures (Peer Spread, Peer Turnover, and Peer AIM). While the coefficient on the VC-backed dummy is positive when the interaction terms are not included, Columns (1), (4), and (7) show a significantly positive (negative) relationship between underpricing and the interaction term of VC-backed and liquidity (illiquidity). This implies that on average, VC-backed issuers leave more money on the table, but that positive relation mostly come from the subset of issuers whose expected liquidity value is high via IPO. Take Column (7) for example, for issuers with expected AIM around 0.36 (25 percentile in the sample), the combined coefficient on VC-backed is 0.166 (=0.197-0.087×0.36); for issuers with expected AIM around 2.89 (75 percentile in the sample), the combined coefficient on VC-backed is -0.054 (=0.197-0.087×2.89). This is consistent with my conjecture that VC funds are willing to leave more money on the table when IPOs are more beneficial in terms of improved liquidity, due to their needs to exit investments in time.

Columns (2), (5), and (8) show that the coefficients on Expected Liquidity×Top Underwriter are significantly positive (negative when measured by illiquidity). This finding indicates the relation between underpricing and liquidity is stronger when the underwriter has stronger bargaining power. When the underwriter's bargaining power is stronger, he can negotiate for more payoffs in the form of underpricing for IPO investors, per unit on liquidity value gain. In Columns (3), (6), and (9), I find that the coefficients of the interaction terms between liquidity (illiquidity) and New Shares Ratio are significantly negative (positive). The relation between underpricing and liquidity is stronger when fewer new shares are issued in an IPO.

Changes in the Order Handling Rules at Nasdaq

In the baseline, I show that underpricing is positively related to the issuer's expected liquidity, and I measure expected liquidity with the issuer's public peer firms' liquidity prior to the IPO time. With this measure, there is unlikely a reverse causality problem. I also include the industry-year fixed effects, to control for any unobserved time-varying industry effects driving both public peers' liquidity and underpricing. Still, this might not be a complete solution for the endogeneity concern. Also, there is a potential measurement error issue as the pre-IPO liquidity is not included in the baseline regression. In this section, I conduct a difference-in-difference test for hypothesis H1 with an exogenous shock to expected liquidity to a subset of issuers using an important regulation change, to formally address these concerns.

The regulation change involves the change to Order Handling Rules (OHR) on Nasdaq in 1997. Christie and Schultz (1994) first expose the lack of odd-eighth quotes on Nasdaq, which help reveal the scandal of Nasdaq dealers colluding to enhance the profitability of their market-making business. Specifically, some dealers did not include competitive limit orders from customers when these orders are better than their own quotes. By doing so, they managed to artificially maintain higher spread of stocks, which suppressed liquidity in the market. In the aftermath of the scandal, SEC enacted several major changes to OHR. First, the Limit Order Display Rule was phased in for all Nasdaq

National Market System issues from January 20, 1997 to October 13, 1997. The rule requires that limit orders should be displayed in the Nasdaq BBO (i.e., best bids and offers) when they are better than quotes posted by market makers. This allows the general public to compete directly with Nasdaq dealers. Second, the Quote Rule requires market makers to publicly display their most competitive quotes. Third, the Actual Size Rule reduces the minimum quote size of market makers from 1000 shares to 100 shares and thereby decreases dealers' market-making risk, and encourages them to maintain more competitive quotes. Lastly, the Excess Spread Rule is amended so that dealers' average spread during each month must be smaller than 150% of the average of the three narrowest spreads over the month. Prior to this, dealers face a similar requirement but on a continuous basis. Changing it to a monthly basis poses less restriction on dealers' ability to change their spreads. All these changes help improve the liquidity of stocks listed on Nasdaq (see Bessembinder 1999, and Barclay et al. 1999).

Using the changes of OHR on Nasdaq in 1997 as an exogenous shock to the expected liquidity of IPOs listed on Nasdaq, I can test the relation between expected liquidity and underpricing in a difference-in-difference framework. The first level of difference is the difference in the magnitude of underpricing before and after 1997. The second level of difference is the difference of the first-level difference among IPO deals listed on Nasdaq and non-Nasdaq exchanges. If I find that after 1997, Nasdaq deals tend to have more underpricing than before 1997, it could be due to a common time trend that is unrelated to the liquidity shock. Only if non-Nasdaq IPOs do not experience the same level of increase in underpricing after 1997, I can rule out the possibility of a common time trend.

By using Nasdaq deals as the treatment group and non-Nasdaq deals as the control group, I can identify the impact of expected liquidity on underpricing. I estimate the following regression equation.

$$\begin{aligned} Underpricing_{ijt} = \alpha_j + \beta_1 Nasdaq_i \times Post_t + \beta_2 Nasdaq_i \\ + \beta_3 Post_t + \gamma' X_{ijt} + \varepsilon_{ijt} \qquad (2) \end{aligned}$$

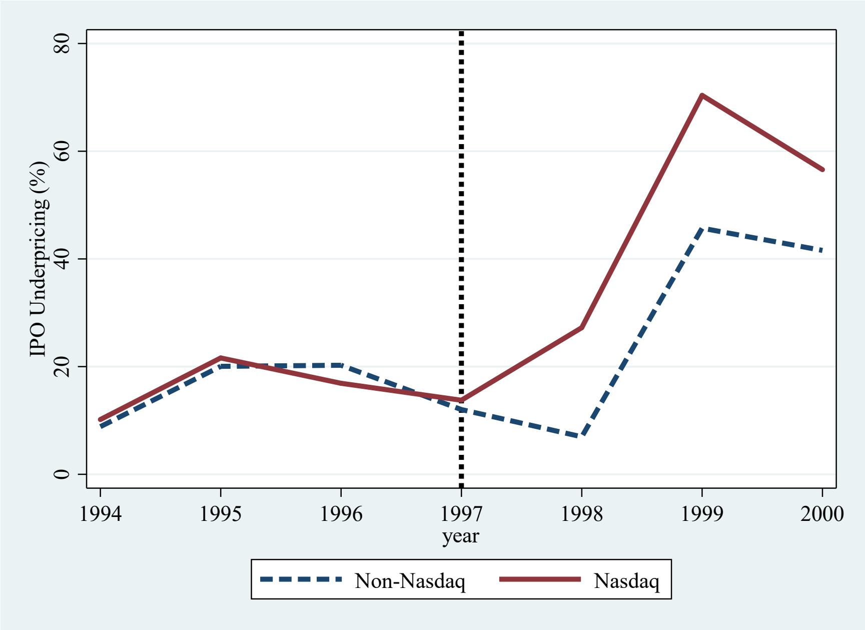
Where i, j, t index firms, industries, and years, respectively. Nasdaq is the dummy variable that is equal to one if the issuer is listed on Nasdaq and zero otherwise; Post is the dummy variable if IPO occurs after 1997 and zero otherwise. X_{ijt} is the same vector of control variables as in Equation (1), including Top Underwriter, New Shares Ratio, Integer Price, Sentiment, Log(Assets), Log(1+Age), and VC-backed. Standard errors are clustered at the level of industry-year groups. H1 predicts that $\beta_1 > 0$, that is, underpricing should increase more (or, decrease less) after 1997 for Nasdaq-listed deals than non-Nasdaq-listed deals, due to improved liquidity of Nasdaq shares after 1997. Because this is essentially an event study, I take the sample years of 1994-2000, which covers the six-year period before and after 1997. Issuers that went public in 1997 are excluded. I include only industry fixed effects but not year fixed effects, due to the relatively shorter time window compared to the full sample period, and the inclusion of Post, which is a dummy variable indicating years before and after the exogenous shock.[21]

---

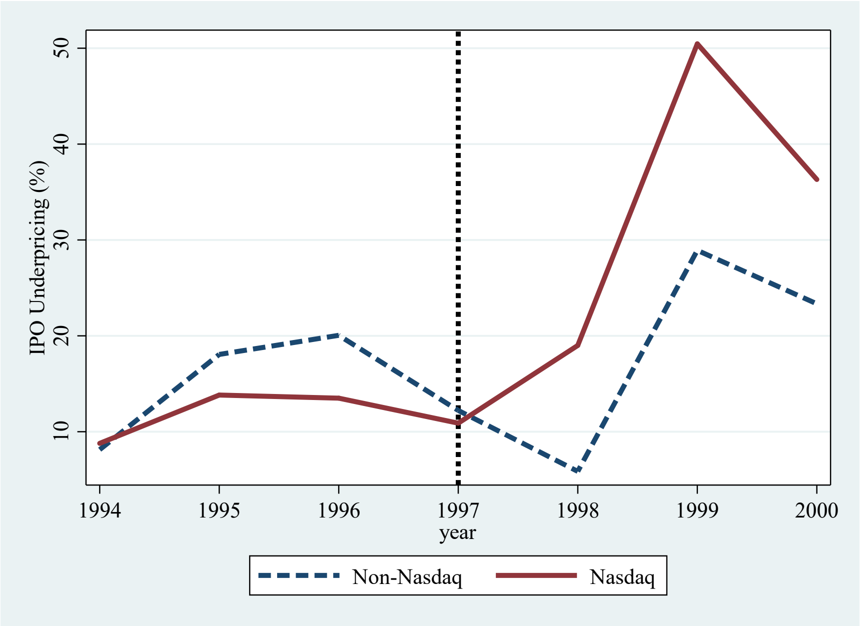
[21] As a robustness check, I replace the *Post* dummy variable with individual year fixed effects, and the results remain. Using the *Post* dummy variable enables direct comparison of average underpricing before and after the treatment event.

As a preliminary investigation, I plot the average underpricing of deals listed on Nasdaq and non-Nasdaq exchanges in each year from 1994 to 2000 in Panel A of Figure 2.2. The figure shows that both the time trend and levels of underpricing are extremely similar for Nasdaq and non-Nasdaq deals before 1997. After 1997 underpricing increases in both groups, but the increase is significantly higher for Nasdaq IPOs. In particular, from 1997 to 1998, shortly after the enactment of OHR changes, underpricing at Nasdaq increased tremendously from the previous year while underpricing at other exchanges actually decreased. This visual examination suggests that there is a real impact of OHR changes on underpricing.

Since the post-1997 era coincides with the tech stock bubble in 1998 and 1999, and many tech stocks go public on Nasdaq, I draw the same plot but excluding all tech stocks as defined in Loughran and Ritter (2004), in Panel B of Figure 2.2. The figure shows that Nasdaq IPOs have lower underpricing than non-Nasdaq deals before 1997, but higher underpricing than non-Nasdaq deals after 1997. The pattern that OHR affects underpricing is even more pronounced. This rules out the possibility that the result is driven by the tech bubble.



Panel A: The whole sample



Panel B: Excluding tech stocks

Figure 2.2. IPO Underpricing at Nasdaq and non-Nasdaq Exchanges

The figure shows the average IPO underpricing for deals listed on Nasdaq and non-Nasdaq exchanges in each year from 1994 to 2000. 1997 is the year when the SEC enacted major changes to Order Handling Rules of Nasdaq. Panel A includes all issuers during the period, while Panel B excludes tech stocks, as defined in Loughran and Ritter (2004).

Table 2.6. IPO Underpricing and Changes to Order Handling Rules at Nasdaq in 1997

The table reports coefficients and t-statistics in the parenthesis of OLS regressions of underpricing on the enactment of OHR at Nasdaq in 1997, and control variables, as shown in regression Equation (2).The sample period is 1994-2000 excluding the year of 1997, when changes of Order Handling Rules at Nasdaq were enacted. I run the regression in two periods: three years before and after the law passage (1994-2000), and two years before and after the law passage (1995-1999). In Columns (1) and (2), the regression is based on the full sample with all Nasdaq and non-Nasdaq IPOs. In Columns (3) and (4), the regression is based on a matched sample, where each Nasdaq IPO is matched with a non-Nasdaq IPO in the same industry (SIC two-digit code), with a similar size (market capitalization), and in the same IPO year. Each year I select all IPO deals and divide them into five quintiles according to their market capitalization in the first year, and conduct size matching by choosing firms in the same quintile. If there are multiple matches, I select the one with the smallest size difference. Only Nasdaq deals with a matched control deal are included in the sample. One non-Nasdaq deal can be shared by multiple Nasdaq deals as a matched control, so one non-Nasdaq deal can appear several times in the data and is counted each time as a separate observation. Nasdaq is a dummy variable that is equal to one if the issuer is listed on Nasdaq and zero otherwise. Post is a dummy variable that is equal to one if the issuer goes public after 1997 and zero otherwise. For the sake of brevity, estimation results for control variables including Top Underwriter, New Shares Ratio, Integer Price, Sentiment, Log(Assets), Log(1+Age), VC-backed and the intercept are not presented. Variable construction is described in Appendix. The industry fixed effects control for the issuer's industry, defined by the Fama-French 10-industry classification. The t-statistics are computed based on standard errors adjusted for heteroskedasticity and industry-year clustering. Asterisks denote statistical significance at 1% (***), 5% (**), or 10% (*) level.

	Full Sample		Excluding Tech deals	
	1994-2000	1995-1999	1994-2000	1995-1999
	(1)	(2)	(3)	(4)
Nasdaq × Post	**0.225*****	**0.253*****	**0.157*****	**0.161*****
	(6.13)	**(6.52)**	**(3.56)**	**(3.07)**
Nasdaq	-0.082***	-0.080**	-0.051**	-0.071**
	(-2.89)	(-2.50)	(-2.01)	(-2.04)
Post	-0.094**	-0.124**	-0.079**	-0.097**

	(-2.05)	(-2.69)	(-2.46)	(-2.21)
Industry FE	YES	YES	YES	YES
Observations	1,610	1,148	807	565
Adjusted R-squared	0.343	0.329	0.227	0.238

I then carry out the regression analysis of Equation (2) and present the results in Table 2.6. I compare underpricing in the periods of three years before and after 1997 (1994-2000), and two years before and after 1997 (1995-1999) in Column (1) and (2) respectively. There are 1,273 Nasdaq deals and 337 non-Nasdaq deals in Column (1), and 897 Nasdaq deals and 251 non-Nasdaq deals in Column (2). Consistent with the theory, the coefficient on Nasdaq×Post is significantly positive in both sample periods. This suggests that Nasdaq IPOs exhibit more underpricing post 1997 compared to non-Nasdaq IPOs. Combining the coefficient on Nasdaq×Post with the one on Nasdaq, I find that, Nasdaq IPOs experience less underpricing prior to 1997, but more underpricing after 1997. Taking Column (1) as an example, I can compute that average underpricing of Nasdaq IPOs is 5.6% less than that of non-Nasdaq IPOs before 1997 (based on the coefficient of -0.056 on Nasdaq), and is 15% higher after 1997 (=0.206-0.056, where 0.206 is the coefficient on Nasdaq×Post). Combining the coefficient on Nasdaq×Post with the one on Post also enables me to confirm the visual finding in Figure 2.2: average underpricing of Nasdaq deals increased significantly after 1997 (the coefficient on Nasdaq×Post is 0.206), while that of non-Nasdaq deals of the same time window stayed almost flat (the coefficient on Post is 0.049 and not significant), controlling for issuer and deal characteristics. I hence conclude that the economic impact of OHR changes on the underpricing of Nasdaq IPOs is large and significant. The results on control variables remain the same as in earlier analysis.

To address the concern that firms endogenously choose where to be listed and thus Nasdaq IPOs and non-Nasdaq IPOs can be fundamentally different, I run the same regression using a matched sample. I match each Nasdaq IPO with a non-Nasdaq deal from the same year in the same industry (SIC two-digit code) and with a similar size (market capitalization). For size matching, each year I select all IPO deals and divide them into five quintiles by ranking their market capitalization in the first year post IPO. If there are multiple matches, I select the one with the smallest size difference. Only Nasdaq deals with a matched control deal are included in the sample. There are 683 (531) Nasdaq deals and an equal number of matched non-Nasdaq deals in the period of 1994-2000 (1995-1999). Note that there are more non-Nasdaq deals in the matched sample compared to the full sample. This is because there are more Nasdaq IPOs than non-Nasdaq IPOs in the unmatched sample. In contrast, in the matched sample, one non-Nasdaq deal can be shared by multiple Nasdaq deals as a control, so it can appear multiple times in the data and is counted each time as a separate observation.

Regression results of the matched sample are presented in Columns (3) and (4) of Table 2.6, with the two sample periods of 1994-2000 and 1995-1999. The main finding of a positive and significant coefficient on Nasdaq×Post remains robust in both specifications. Of the two sample periods, the smaller estimate is β_1=0.177 in Column (3), which implies a relative increase of 17.7% in underpricing after 1997 for Nasdaq IPOs compared to non-Nasdaq IPOs. Based on the average IPO underpricing of 21.5%, the marginal effect of OHR measured by β_1 features an 82.3% (=17.7/21.5) increase of underpricing from its mean level. Combining the coefficient on Nasdaq×Post and the ones on Nasdaq or Post, I reach similar conclusions to what I have with the full sample, shown in the first two columns of Table 2.6.

Table 2.7. Placebo Tests in the Pre-OHR Period

The table reports coefficients and t-statistics in the parenthesis of OLS regressions of underpricing on a pseudo event that is assumed to only affect the liquidity of Nasdaq stocks in a particular year between 1981 and 1996, before OHR is enacted in 1997. The regression variables are shown in Equation (2). For brevity, only coefficients of key explanatory variables are shown. I take the pre-treatment period of the sample from 1981 to 1996 and select moving windows of seven years with a pseudo event occurring in the 4th year, The pseudo event years are the years between 1984 and 1993. Nasdaq is a dummy variable that is equal to one if the issuer is listed on Nasdaq and zero otherwise. Post is a dummy variable that is equal to one if the issuer goes public after the pseudo event year and zero otherwise. For the sake of brevity, estimation results for control variables including Top Underwriter, New Shares Ratio, Integer Price, Sentiment, Log(Assets), Log(1+Age), VC-backed and the intercept are not presented. Variable construction is described in Appendix. The industry fixed effects control for the issuer's industry, defined by the Fama-French 10-industry classification. The t-statistics are computed based on standard errors adjusted for heteroskedasticity and industry-year clustering. Asterisks denote statistical significance at 1% (***), 5% (**), or 10% (*) level.

Event Year	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
Nasdaq	**0.019**	**0.012**	**-0.105**	**0.042**	**0.019**	**0.014**	**0.008**	**-0.022**	**-0.008**	**-0.024**
× Post	**(0.25)**	**(0.16)**	**(-0.97)**	**(0.89)**	**(0.61)**	**(0.61)**	**(0.40)**	**(-0.55)**	**(-0.23)**	**(-0.86)**
Nasdaq	-0.022	-0.024	0.097	-0.001	0.000	-0.006	-0.010	0.029	0.020	0.011
	(-0.29)	(-0.33)	(0.88)	(-0.07)	(0.02)	(-0.49)	(-0.73)	(0.74)	(0.70)	(0.54)
Post	0.073*	0.055	-0.001	-0.044	0.022	0.009	0.021	0.024	0.054*	0.065**
	(1.97)	(1.02)	(-0.02)	(-0.87)	(0.46)	(0.46)	(1.18)	(0.78)	(1.75)	(2.50)
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	358	395	269	352	581	774	843	826	934	1,225
Adjusted R2	0.144	0.141	0.189	0.096	0.162	0.201	0.293	0.243	0.273	0.258

Next, I conduct placebo tests to show that the significant change in the trend and level of underpricing for Nasdaq and non-Nasdaq deals occurring in 1997 is not just a fluke of data or due to fundamental differences between Nasdaq and non-Nasdaq deals. I take the pre-treatment period of the sample from 1981 to 1996 and select moving windows of seven years with a pseudo event occurring in the 4th year, when only Nasdaq stocks are assumed to undergo an exogeneous shock in liquidity. There are in total ten such event windows, with the pseudo event years being each year between 1984 and 1993. I run the same regression Equation (2) for the full sample in these event windows with three years before and after the pseudo event year. I do not find any significant coefficients for the interaction term Nasdaq×Post. The results are presented in Table 2.7.

Another potential issue is that changes to OHR at Nasdaq, as a shock to the expected post-IPO liquidity, may create some selection bias because firms choose whether to go public strategically and such decisions may be affected by the liquidity value of the deal. However, I argue this selection channel should bias against finding a result. After the positive shock of OHR, the liquidity value of going public increases, which motivates some firms with low liquidity gain to go public who would otherwise choose to stay private. The entry of these marginal firms into the IPO market would contaminate the treated group and reduce the treatment effect. Hence, I believe this possible selection channel in fact makes my results stronger.

Table 2.8. Individual Peer liquidity and Order Handling Rules at Nasdaq in 1997

The table reports coefficients and t-statistics in the parenthesis of OLS regressions of individual peer liquidity on the enactment of OHR at Nasdaq in 1997, and control variables, as shown in regression Equation (3). The sample period is 1994-2000 excluding the year of 1997, when changes of Order Handling Rules at Nasdaq were enacted. I run the regression in two periods: three years before and after the law passage (1994-2000), and two years before and after the law passage (1995-1999). Nasdaq is a dummy variable that is equal to one if the issuer is listed on Nasdaq and zero otherwise. Post is a dummy variable that is equal to one if the issuer goes public after 1997 and zero otherwise. Individual peer spread (turnover, or AIM) are the monthly spread (turnover, or AIM) calculated as averages of daily data for each peer firm of the issuer. Peer firms are publicly traded companies in the same industry (SIC two-digit code), with a similar size (belonging to the same quintile in the year of IPO as the issuer when the COMPUSTAT-CRSP universe is sorted by market capitalization), and listed on the same exchange as the issuer. Log (1+ Peer Age), Log (Sales), Log (Market cap), and Number of shareholders are annual characteristics of the peer firm. Market return is the value-weighted monthly market return (NYSE/AMEX/NASDAQ/ARCA) reported in CRSP. Variance of market returns is the variance of daily market returns in a given month. Sentiment is the monthly market sentiment index downloaded from Prof. Jeff Wurgler's website. Interest rate is the monthly three-month T-bill rate downloaded from the Federal Reserve Bank's website. Additional variable construction is described in Appendix. The industry fixed effects control for the issuer's industry, defined by the Fama-French 10-industry classification. The t-statistics are computed based on standard errors adjusted for heteroskedasticity and firm-level clustering. Asterisks denote statistical significance at 1% (***), 5% (**), or 10% (*) level.

	Individual Peer Spread		Individual Peer Turnover		Individual Peer AIM	
	1994-2000	1995-1999	1994-2000	1995-1999	1994-2000	1995-1999
	(1)	(2)	(3)	(4)	(5)	(6)
Nasdaq × Post	**-0.033*****	**-0.029*****	**0.001*****	**0.001*****	**-0.237*****	**-0.190*****
	(-32.75)	**(-27.46)**	**(5.92)**	**(3.71)**	**(-8.78)**	**(-7.24)**
Nasdaq	0.003***	-0.001	0.003***	0.004***	0.182***	0.130***
	(2.60)	(-1.10)	(20.22)	(19.24)	(5.70)	(3.93)
Post	0.014***	0.010***	0.000	-0.000***	0.114***	0.077***
	(15.70)	(11.33)	(0.43)	(-3.28)	(5.96)	(4.10)
Log(1+ Peer Age)	0.001***	0.001***	-0.001***	-0.001***	0.124***	0.129***

	(3.51)	(4.20)	(-15.29)	(-16.52)	(10.88)	(10.56)
Log(Sales)	0.000	0.001**	-0.001***	-0.001***	0.032***	0.038***
	(1.59)	(2.07)	(-12.24)	(-10.27)	(3.81)	(4.06)
Log(Market cap)	-0.018***	-0.018***	0.002***	0.002***	-0.769***	-0.777***
	(-55.15)	(-49.91)	(33.56)	(28.70)	(-72.74)	(-67.07)
Number of shareholders	0.429***	0.430***	-0.035***	-0.034***	19.335***	20.102***
	(24.98)	(23.16)	(-7.88)	(-7.13)	(25.57)	(25.04)
Market return	0.014***	0.018***	0.006***	0.004***	-0.033	0.325***
	(15.52)	(15.48)	(23.35)	(12.27)	(-1.16)	(9.00)
Lagged market return	-0.004***	-0.002***	0.003***	0.004***	-0.576***	-0.307***
	(-4.90)	(-2.68)	(9.88)	(12.29)	(-18.01)	(-9.83)
Variance of market	19.732***	24.579***	2.381***	-2.582***	315.015***	814.810***
	(34.15)	(24.86)	(13.53)	(-10.82)	(18.00)	(26.54)
Sentiment	0.002***	0.001	-0.001***	-0.002***	0.101***	-0.009
	(7.80)	(1.51)	(-8.97)	(-15.00)	(13.04)	(-0.72)
Interest rate	-0.007	-0.103***	0.038***	-0.047***	-2.421***	5.706***
	(-0.41)	(-2.69)	(10.38)	(-5.91)	(-4.61)	(5.26)
Constant	0.125***	0.134***	-0.004***	0.001**	4.857***	4.447***
	(53.42)	(40.28)	(-10.00)	(1.97)	(61.10)	(44.52)
Industry FE	YES	YES	YES	YES	YES	YES
Observations	232,513	162,457	234,051	163,429	233,911	163,319
Adjusted R-squared	0.508	0.503	0.217	0.203	0.621	0.619

Lastly, I verify that changes to OHR at Nasdaq are indeed shocks to expected liquidity of issuers at Nasdaq and issuers at other exchanges are not subject to the same shock. I test the relation between the regulation change and expected liquidity directly. I use the monthly spread, turnover, and log AIM (averages of daily data) of each individual public peer as the dependent variable, instead of the average value across peer firms, for more accuracy of the test. To differentiate from the variables of Peer Spread, Peer Turnover, and Peer AIM defined earlier as the average liquidity across peer firms, I name them Individual Peer Spread, Individual Peer Turnover, and Individual Peer AIM. I run the following regression equation.[22]

$$\begin{aligned} Indiviual\ Peer\ Liquidity_{ijt} &= \alpha_j + \beta_1 Nasdaq_i \times Post_t + \beta_2 Nasdaq_i \\ &\quad + \beta_3 Post_t + \gamma' Z_{ijt} + \varepsilon_{ijt} \qquad (3) \end{aligned}$$

where liquidity is measured by spread, turnover, or log AIM, and i, j, t index firms, industries, and months, respectively. Z_{ijt} is a vector of control variables shown in previous studies that are related to firm liquidity. The literature documents commonality of liquidity, so I control for market level variables such as the market return, the lagged market return, the variance of daily market returns, market sentiment, and the interest rate measured by the three-month T-bill rate (see Huberman and Halka, 2001; Chordia, Roll, and Subrahmanyam, 2000, 2001, among others). All market level variables are of monthly frequency.[23] I also control for peer firms' characteristics such as the log of peer firm age as a public firm, the log of sales, the log of market capitalization, and the number of shareholders as a measure for ownership diversity. All firm characteristics are annual observations downloaded from COMPUSTAT. Industry fixed effects are included. Standard errors are clustered at the firm level as the dependent variable is a firm-specific liquidity measure. I also run the regression for two periods of 1994-2000 and 1995-1999, excluding the year of 1997.

---

[22] In Equation (3), some control variables have monthly frequencies, some firm-level characteristics have annual frequencies, and I run monthly regressions.

The results are presented in Table 2.8. Columns (1) and (2) show that in both periods, individual peer spread drops more for Nasdaq issuers than non-Nasdaq issuers and the difference is highly significant. Based on the results in Column (1), I document that peer firms of Nasdaq issuers experience a drop of 3.3% in quoted spread relative to those of non-Nasdaq issuers after 1997, in absolute terms. This is economically large as the average peer spread is 3.5%. Taking the coefficient on Nasdaq×Post and the one on Post in Column (1), I estimate that after the changes to OHR, average individual peer spread decreased 1.9% (= -0.033 + 0.014 = -0.019) for Nasdaq issuers, while average individual peer spread increased 1.4% (= 0.014) for non-Nasdaq issuers, in the period of 1994-2000. Column (2) shows similar patterns with slightly different magnitudes in the period of 1995-1999. I conclude that changes to OHR impact Nasdaq issuers' expected trading spread, in an economically significant way. This is consistent with the purpose of the regulation change, as changes to OHR are designed to reduce quoted spread on Nasdaq.

---

[23] The market return is the value-weighted NYSE/AMEX/NASDAQ/ARCA return reported by CRSP. The variance of market returns is the variance of daily market returns in a given month. Monthly market sentiment is downloaded from Prof. Jeff Wurgler's website. The monthly three-month T-bill rate is download from the Federal Reserve Bank's website.

Columns (3) and (4) report the results when liquidity is measured by turnover. The coefficient on Nasdaq×Post is statistically significant and positive in both periods. Using the coefficient estimates of Nasdaq×Post and Nasdaq in Column (3), I estimate that on average, average peer turnover of Nasdaq issuers is 0.3% higher than that of non-Nasdaq issuers before 1997, but even more so after 1997, when the difference becomes 0.4% (=0.001+0.003). Given that average peer turnover is 0.5%, the impact of changes to OHR at Nasdaq on turnover is also economically large.

Lastly, Columns (5) and (6) show that the price impact of trades becomes smaller for Nasdaq shares after 1997 compared to non-Nasdaq shares. Taking Column (5) for example, combining the coefficients on Nasdaq×Post and Nasdaq, I can see that the price impact measured by log AIM at Nasdaq was higher than that at NYSE and ASE by 0.182 before 1997, but became lower than the latter after 1997 by 0.055 (=-0.237+0.182). The coefficient on the Post dummy also shows that the price impact on NYSE and ASE increased after 1997 (their average log AIM increased by 0.114), but it decreased on Nasdaq (its average log AIM decreased by 0.123, which is -0.237+0.114). In summary, Table 2.8 shows that changes to OHR at Nasdaq indeed improves Nasdaq issuers' expected liquidity, measured by individual spread, turnover, and price impact of trades of its peer firms listed on Nasdaq. To make sure that the tech bubble is not driving these results, I drop all tech firms as defined in Loughran and Ritter (2004) and repeat the analysis in Table 2.6 and Table 2.8, and the results remain.

The National Securities Market Improvement Act (NSMIA)

The hypothesis H2 predicts a negative relation between underpricing and the issuer's pre-IPO liquidity. Intuitively, if the pre-IPO liquidity of issuers is better, the liquidity value gained via IPO is lower, reducing the surplus in the negotiation as well as the need for underpricing. However, testing H2 is more complicated than testing H1, because a secondary market for private shares does not widely exist. Most private firms have restrictions on shareholders selling their shares to a third party, for concerns of control rights. And such transactions are normally privately negotiated if they occur. As a result, I cannot construct liquidity measures for private shares the same way as I do for public shares such as spread, turnover, or AIM, and subsequently test a similar baseline for H2 to the one for H1 specified by Equation (1).

Meanwhile, the most common type of trading for private shares is by founders themselves when they raise capital from private share investors. Hence, I measure the pre-IPO liquidity of private firms by how easily founders can raise capital. When it is easier for founders to sell ownership in exchange for external equity capital, the private firm receives higher valuations, which is equivalent to the concept of smaller price impact for public shares. One important factor determining how easily private firms can raise capital is the size of the pool of potential investors available, which I use as the proxy for the liquidity of private shares in this section. Based on this measure, I test H2 by exploiting a law change that asymmetrically affects private firms' access to VC and PE investors and adopting a difference-in-difference approach. Even though the firm-specific pre-IPO liquidity of issuers is not directly observable, I can use the law change as an exogenous shock to pre-IPO liquidity and compare underpricing before and after the law change across deals. The law change examined is the National Securities Market Improvement Act (NSMIA), passed in October 1996. Ewens and Farre-Mensa (2017) provide an excellent and detailed description of the law. I describe and summarize the law as follows.

NSMIA brings two major changes to the issuance and trading of private securities. First, before the law change, a private firm seeking to raise capital needs to comply with state regulations known as blue-sky laws, in addition to federal regulations such as Regulation D. Since these state regulations are often complex and different from each other, any private firm raising capital from multiple states faces significant regulatory burdens. NSMIA creates certain federal provisions that exempt qualified private security issuers from having to comply with these blue-sky laws in each state where they raise capital. Specifically, securities sold under Rule 506 of Regulation D, which allows private firms to raise unlimited amount of capital when the investors are "accredited investors", are exempted.[24] This exemption also applies to the fundraising of many VCs and PEs.

Second, NSMIA affects VC and PE funds directly through changes to the Investment Company Act of 1940. The Act mandates that most investment advisors must register with the SEC, regularly disclose their investment positions, and limit their use of leverage. VC and PE funds have often relied on the Act's exemption to avoid having to comply with its costly registration and disclosure requirements. NSMIA expanded these exemptions and made it easier for VC and PE funds to satisfy the exemption criteria. The law effectively removes the 100-investor cap in private investment funds, allowing these funds to raise capital from a larger number of investors and prompting the rise of large VC and PE funds. This directly improves the liquidity of private securities by broadening the pool of potential buyers and increasing the amount of equity capital available for private firms. The market for private securities has also become more professionalized, with VC and PE funds and operating businesses all vying for opportunities to invest in private companies or to acquire them outright (see De Fontenay, 2017).

---

[24] "Accredited investors" are institutions, individuals with annual income above $200K ($300K for couples), or individuals and couples with net worth above $1 million excluding the primary residence.

Both features of NSMIA improve the liquidity environment of private firms. They not only make it easier for private firms to raise capital, but also directly expand the pool of potential investors in private firms. Even though the law impacts all private firms in the U.S. at the national level, I conjecture that the effect is more pronounced for private firms located in states with less local VC and PE funds. For example, consider a private firm located in San Francisco versus another one in North Dakota. For the firm in San Francisco, raising capital only within the state of California is likely to satisfy all of its capital needs, and thus the passage of NSMIA hardly makes any difference. While for the firm in North Dakota, due to the lack of in-state private capital and investors, it needs to face heavy compliance obstacles dealing with other state blue-sky laws prior to NSMIA, and the passage of NSMIA alleviates this burden substantially. Meanwhile, by removing the 100-investor cap, NSMIA should increase the amount of capital managed by VC and PE firms more in states with more such firms to begin with. As a result, NSMIA should have a larger impact on the North Dakota firm's pre-IPO liquidity than that of the San Francisco firm.

Motivated by the regulation's differential impact on firms located in different states, I adopt the difference-in-difference approach. I construct the treatment group as issuers located in states with few in-state VC and PE investors, and the control group as issuers in states with abundant private share investors. I can then compare the change of underpricing of the treatment group with that of the control group.

I collect the number of VC and PE firms by state and year from Thomson Reuters Eikon. Since NSMIA is enacted towards the end of 1996, I take 1996-1997 as the event time, and focus on the period of three years before and after the law change (1993-2000).[25] I then rank states in this period by the total number of these firms in the event period of 1993-2000. The ranking is shown in Table 2.9. Not surprisingly, I find that the number of VC and PE firms from the top four states of CA, NY, MA, and TX together account for 57.89% of all such firms in the entire country. Issuers located in these states are more likely to have larger number of potential investors within the state, and issuers located in other states have relatively fewer potential investors locally and therefore are more likely to raise capital from other states. I take issuers in the top four states as the control sample, and issuers outside these states as the treatment sample.

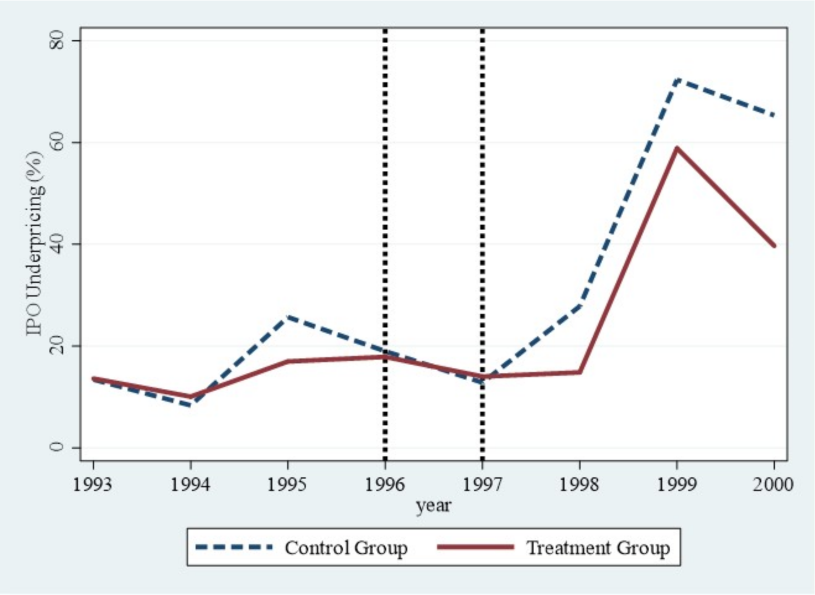
---

[25] Unlike changes to OHR at Nasdaq that affect the liquidity of public firm listed there immediately, it could take NSMIA longer time to impact the liquidity of private firms and the magnitude of IPO underpricing.

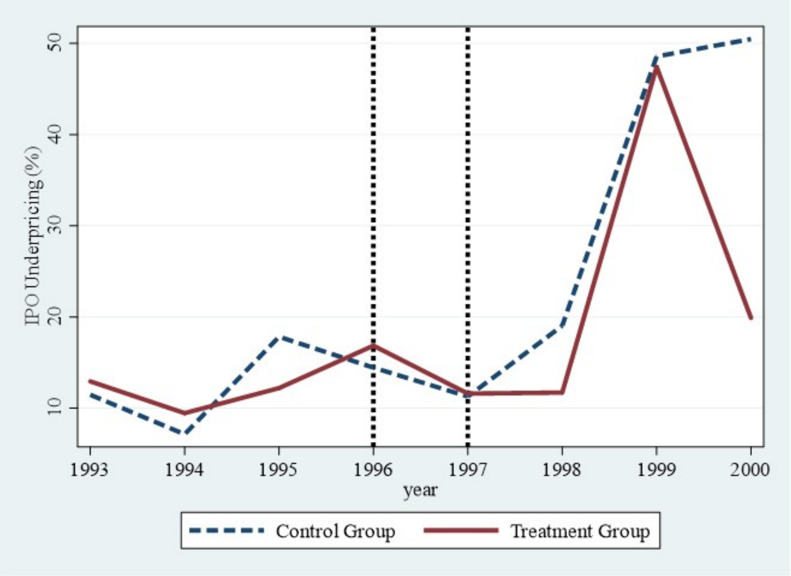
Table 2.9. Distribution of VC and PE Firms across States, 1993-2000

The table reports the number and percentage of VC and PE firms, in each of the 50 states and the District of Columbia in 1993-2000, ranked from the largest to the smallest. States (AK, ND) with zero VC and PE firms are not shown. This period covers three years before and after the passage of the National Securities Market Improvement Act (NSMIA), which is enacted in October 1996. Data source: Thomson Reuters Eikon.

Rank	State	Freq.	Percent	Rank	State	Freq.	Percent
1	CA	727	25.14	26	AZ	14	0.48
2	NY	519	17.95	27	AL	13	0.45
3	MA	246	8.51	28	RI	11	0.38
4	TX	182	6.29	29	LA	10	0.35
5	IL	136	4.70	30	NH	9	0.31
6	CT	128	4.43	31	DE	7	0.24
7	PA	100	3.46	31	NV	7	0.24
8	NJ	77	2.66	33	KS	6	0.21
9	WA	60	2.07	33	SC	6	0.21
10	MN	59	2.04	33	NM	6	0.21
11	CO	54	1.87	33	ME	6	0.21
12	GA	53	1.83	33	IA	6	0.21
12	MD	53	1.83	38	OK	5	0.17
14	FL	52	1.80	38	KY	5	0.17
15	OH	49	1.69	40	AR	3	0.10
16	NC	42	1.45	40	WY	3	0.10
17	VA	40	1.38	40	MT	3	0.10
18	DC	39	1.35	40	NE	3	0.10
19	MI	29	1.00	40	VT	3	0.10
20	TN	25	0.86	45	MS	2	0.07
21	MO	23	0.80	45	WV	2	0.07
22	IN	17	0.59	47	SD	1	0.03
22	WI	17	0.59	47	HI	1	0.03
24	OR	16	0.55	47	ID	1	0.03
24	UT	16	0.55				
					Total	2,892	100



Panel A: The whole sample



Panel B: Excluding tech stocks

Figure 2.3. IPO Underpricing in States with and without Large Numbers of VC and PE

The figure shows the average IPO underpricing for issuers headquartered in the top four states (CA, NY, MA, and TX) with largest number of VC and PE firms (the control sample) and issuers headquartered outside these states (the treatment sample). The National Security Market Improvement Act (NSMIA) was enacted in October 1996, and the two black dashed lines mark the event time. I plot average underpricing in each year in 1993-2000 of the two groups. Panel A includes all issuers, while Panel B excludes tech stocks, as defined in Loughran and Ritter (2004).

As a first look at the data, I plot the average underpricing of deals in these two subsamples in each year from 1993 to 2000 in Panel A of Figure 2.3. The figure shows that both the time trend and levels of underpricing are extremely similar in the two groups before 1996. After 1997 underpricing increases in both groups, but the increase is significantly lower for the treatment group, which is consistent with my conjecture. For concerns that the tech bubble period drives the results, I replot the figure excluding tech firms in Panel B of Figure 2.3, and find a similar pattern. The smaller magnitude of underpricing for the treatment sample is especially pronounced in 1998 and 2000. I hypothesize that IPO underpricing should decrease more (or, increase less) for issuers in the treatment sample after the passage of NSMIA, controlling for expected liquidity. I run the following difference-in-difference regression.

$$Underpricing_{ijt} = \alpha_j + \beta_1 Treated \times Post + \beta_2 Treated + \beta_3 Post$$
$$+\beta_4 Expected\ Liquidity_{ijt} + \gamma' X_{ijt} + \varepsilon_{ijt} \qquad (4)$$

where i, j, t index firms, industries, and years, respectively. Treated is the dummy variable that is equal to one if the issuer is headquartered outside of the control states (CA, NY, MA, and TX), and zero otherwise. I also explore alternative control samples with the top eight states (CA, NY, MA, TX, IL, CT, PA, and NJ), or the top two states (CA and NY). Post is the dummy variable if IPO occurs after 1997 and zero if it occurs before 1996. Standard errors are clustered at the level of industry-year groups.

Table 2.10. IPO Underpricing and the National Security Market Improvement Act: Diff-in-Diff

The table reports coefficients and t-statistics in the parenthesis of OLS regressions of IPO underpricing on the passage of the National Security Market Improvement Act (NSMIA) in October 1996, as shown in regression Equation (4). The sample period is 1993-2000 excluding the years of 1996 and 1997. I run the regression in two times periods: three years before and after the law passage (1993-2000) and two years before and after the law passage (1994-1999). The control sample includes issuers headquartered in states with the largest numbers of VC and PE firms (top eight, four, or two), and the treatment sample includes issuers headquartered outside of those states. Expected liquidity is measured by Peer Spread in Panel A, by Peer Turnover in Panel B, and by Peer AIM in Panel C. The three panels have otherwise identical columns. In Columns (1) and (2), the control sample includes issuers headquartered in CA, NY, MA, TX, IL, CT, PA, and NJ. In Columns (3) and (4), the control sample includes issuers headquartered in CA, NY, MA, and TX. In Columns (5) and (6), the control sample includes issuers headquartered in CA and NY. Treated is a dummy variable that is equal to one if the issuer is in the treatment sample and zero if the issuer is in the control sample. Post is a dummy variable that is equal to one if the issuer goes public after 1997 and zero if before 1996. For the sake of brevity, estimation results for control variables including Top Underwriter, New Shares Ratio, Integer Price, Sentiment, Log(Assets), Log(1+Age), VC-backed and the intercept are not presented. Variable construction is in Appendix. The industry fixed effects control for the issuer's industry, defined by the Fama-French 10-industry classification. The t-statistics are computed based on standard errors adjusted for heteroskedasticity and industry-year clustering. Asterisks denote statistical significance at 1% (***), 5% (**), or 10% (*) level.

Panel A: Controlling for peer spread

	Control Sample					
	CA, NY, MA, TX, IL, CT, PA, NJ		CA, NY, MA, TX		CA, NY	
	1993-2000	1994-1999	1993-2000	1994-1999	1993-2000	1994-1999
	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	**-0.127*** **	**-0.081***	**-0.134*** **	**-0.092***	**-0.087***	**-0.104*** **
	(-3.19)	**(-1.89)**	**(-3.38)**	**(-1.79)**	**(-1.76)**	**(-2.82)**
Treated	0.016	0.004	0.016	0.001	-0.000	-0.005
	(1.13)	(0.24)	(1.09)	(0.03)	(-0.01)	(-0.20)
Post	0.094**	0.046	0.113**	0.062	0.101**	0.083
	(2.12)	(1.06)	(2.52)	(1.28)	(2.01)	(1.66)
Peer Spread	-0.655	-1.644*	-0.705	-1.671**	-0.686	-1.673*
	(-1.01)	(-2.02)	(-1.09)	(-2.06)	(-1.05)	(-2.02)
Industry FE	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES
Observations	1,343	889	1,343	889	1,343	889
Adjusted R-squared	0.369	0.363	0.371	0.365	0.368	0.366

Panel B: Controlling for peer turnover

	Control Sample					
	CA, NY, MA, TX, IL, CT, PA, NJ		CA, NY, MA, TX		CA, NY	
	1993-2000 (1)	1994-1999 (2)	1993-2000 (3)	1994-1999 (4)	1993-2000 (5)	1994-1999 (6)
Treated × Post	**-0.108*****	**-0.052**	**-0.119*****	**-0.072***	**-0.073**	**-0.085****
	(-2.80)	**(-1.59)**	**(-3.34)**	**(-1.88)**	**(-1.60)**	**(-2.71)**
Treated	0.015	0.002	0.014	-0.004	0.005	-0.002
	(1.09)	(0.12)	(0.91)	(-0.22)	(0.22)	(-0.08)
Post	0.008	-0.009	0.028	0.008	0.014	0.027
	(0.15)	(-0.22)	(0.51)	(0.22)	(0.22)	(0.57)
Peer Turnover	32.759***	35.747***	32.815***	35.760***	32.852***	35.519***
	(3.49)	(2.75)	(3.53)	(2.77)	(3.51)	(2.73)
Industry FE	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES
Observations	1,343	889	1,343	889	1,343	889
Adjusted R-squared	0.409	0.411	0.411	0.412	0.408	0.413

Panel C: Controlling for peer AIM

	Control Sample					
	CA, NY, MA, TX, IL, CT, PA, NJ		CA, NY, MA, TX		CA, NY	
	1993-2000 (1)	1994-1999 (2)	1993-2000 (3)	1994-1999 (4)	1993-2000 (5)	1994-1999 (6)
Treated × Post	**-0.125*** **	**-0.073**	**-0.133*** **	**-0.086**	**-0.090** **	**-0.104*** **
	(-3.11)	**(-1.67)**	**(-3.36)**	**(-1.66)**	**(-1.84)**	**(-2.81)**
Treated	0.017	0.001	0.017	-0.003	0.004	-0.003
	(1.19)	(0.04)	(1.11)	(-0.15)	(0.20)	(-0.11)
Post	0.073	0.020	0.093**	0.036	0.083	0.059
	(1.61)	(0.43)	(2.02)	(0.71)	(1.60)	(1.14)
Peer AIM	-0.036**	-0.061***	-0.037**	-0.061***	-0.037**	-0.061***
	(-2.55)	(-3.41)	(-2.60)	(-3.43)	(-2.64)	(-3.39)
Industry FE	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES
Observations	1,343	889	1,343	889	1,343	889
Adjusted R-squared	0.372	0.369	0.374	0.371	0.371	0.372

H2 predicts that β_1<0. I explicitly control for the post-IPO expected liquidity as the shocks here are to pre-IPO liquidity. X_{ijt} is the same vector of control variables as in Equation (1), which are Top Underwriter, New Shares Ratio, Integer Price, Sentiment, Log(Assets), Log(1+Age), and VC-backed. For the same reasons explained in Equation (2), I do not include year fixed effects but just industry fixed effects in Equation (4). Replacing the Post dummy with year fixed effects does not change the key results. I investigate two event windows around the event: three years before and after the law change (1993-2000) and two years before and after the law change (1994-1999), both excluding 1996 and 1997.

The regression results are presented in Table 2.10. I control for Peer Spread, Peer Turnover, and Peer AIM in Panel A, B, and C, respectively. The three panels are otherwise identical except for the definition of expected liquidity. Columns (1) and (2) use issuers from the top eight states with the largest number of VC and PE firms (CA, NY, MA, TX, IL, CT, PA, and NJ) as the control firms, whereas Columns (3) and (4) use issuers from the top four states (CA, NY, MA, and TX) and Columns (5) and (6), the top two states (CA and NY). Across all 18 specifications (two sample periods × three measures of liquidity × three control samples), I find negative coefficient consistently on Treated×Post, statistically significant in all but two specifications.[26] Since all specifications generate qualitatively similar results, I only describe Panel A in detail below.

---

[26] The only exceptions are Columns (2) and (5) in Panel B, where the t-statistics of *Treated* × *Post* are -1.32 and -1.61, respectively.

Combining the coefficients on Treated×Post with the ones on Treated or Post, I reach two interesting conclusions. First, before the enactment of NSMIA, issuers located in states with less potential investors (the treatment group) experience larger underpricing than issuers located in states with more investors (the control group). This is reflected in the positive coefficient on Treated across all columns, which is also statistically significant in six out of 18 specifications. Based on these results, I estimate that underpricing for these issuers are about 1% to 4% higher. This itself is an intuitive finding, as it implies that IPO is especially important and beneficial for firms located in states without a large pool of potential investors before 1996, and these issuers are thus willing to leave more money on the table while going public. After the enactment of NSMIA, the pattern flipped, as issuers in the treatment group experience smaller underpricing than issuers in the control group. This is implied by the negative coefficient on Treated×Post and the positive coefficient on Treated, and the fact that the magnitude of the former is always larger than the magnitude of the latter. For example, Column (3) shows that the difference is -14% (=-0.178+0.041).

Second, the coefficient on the Post dummy is positive and statistically significant in 14 out of 18 specifications. This suggests that, during the sample period, underpricing increased for the control sample, but it increased less or even decreased for issuers in the treatment group. For example, the estimates in Column (3) of Panel A show that underpricing increased by 22.2% for issuers in the control group in 1993-2000, but only increased by 4.4% (=0.222-0.178) for issuers in the treatment group in the same period. Column (4) shows that while underpricing for the control group increased by 12.9% in 1994-1999, it actually decreased for the treatment group by 0.5% (=0.129-0.134). The coefficient estimates on control variables remain consistent with earlier analysis. Overall the evidence suggests that after the passage of NSMIA, the liquidity benefit provided by going public becomes smaller for issuers in the treatment group compared to issuers in the control group. And the compensation received by IPO investors from these issuers, which is measured by underpricing, becomes lower than that from the issuers in the control group. I conclude that NSMIA has significantly different economic impact on issuers in the treatment sample versus the control sample.

I also conduct placebo tests using the pre-treatment period before 1996. I use the top eight states (CA, NY, MA, TX, IL, CT, PA, NJ) as the control sample and the rest of the states as the treatment sample. I select moving windows of eight years with a pseudo event occurring in the 4th and 5th year (as the NSMIA event years include two years), when only private firms in treatment states are assumed to undergo an exogeneous shock in liquidity. There are in total nine such event windows, with the pseudo event years being every two years between 1984 and 1993. I run the same regression Equation (4) in these event windows with three years before and after the pseudo event years. I do not find any significant coefficients for the interaction term Treated×Post. I present the results when expected liquidity is proxied by Peer AIM for brevity in Table 2.11.[27]

---

[27] I find similar results with alternative specifications such as using top four or top two states instead of top eight as the control sample, restricting the event window to be two years instead of three years before and after pseudo event years, using *Peer Spread* or *Peer Turnover* as the proxy for expected liquidity.

Table 2.11. Placebo Tests in the pre-NSMIA Period

The table reports coefficients and t-statistics in the parenthesis of OLS regressions of underpricing on a pseudo event that is assumed to only affect the liquidity of private firms in treated states in a particular two-year period between 1981 and 1996, before NSMIA is enacted in October, 1996. The regression variables are shown in Equation (4). For brevity, only coefficients of key explanatory variables are shown. The control states are the top eight states with the highest number of PE and VC firms, i.e. CA, NY, MA, TX, IL, CT, PA, and NJ. The treated states include all other states. I take the pre-treatment period of the sample from 1981 to 1996 and select moving windows of eight years with a pseudo event occurring in the 4th and 5th year, The pseudo event years are the two consecutive years between 1984 and 1993. The event window includes three years before and after the pseudo event years. Treated is a dummy variable that is equal to one if the issuer located in the treated states. Post is a dummy variable that is equal to one if the issuer goes public after the pseudo event years and zero otherwise. Expected liquidity is proxied by Peer AIM. For the sake of brevity, estimation results for control variables including Peer AIM, Top Underwriter, New Shares Ratio, Integer Price, Sentiment, Log(Assets), Log(1+Age), VC-backed and the intercept are not presented. Variable construction is described in Appendix. The industry fixed effects control for the issuer's industry, defined by the Fama-French 10-industry classification. The t-statistics are computed based on standard errors adjusted for heteroskedasticity and industry-year clustering. Asterisks denote statistical significance at 1% (***), 5% (**), or 10% (*) level.

Event window	1981-1988	1982-1989	1983-1990	1984-1991	1985-1992	1986-1993	1987-1994	1988-1995	1989-1996
Treated × Post	-0.070	-0.077	0.085	0.013	-0.005	-0.023	0.004	-0.083	-0.042
	(-1.13)	(-1.67)	(0.93)	(0.32)	(-0.19)	(-1.45)	(0.21)	(-1.56)	(-0.98)
Treated	0.095	0.082*	0.000	0.021	0.012	0.019*	-0.002	0.083	0.034
	(1.58)	(1.86)	(0.00)	(1.14)	(0.96)	(1.77)	(-0.14)	(1.63)	(0.86)
Post	0.167**	0.173***	0.010	-0.029	0.011	0.054***	0.036**	0.127***	0.111***
	(2.56)	(2.95)	(0.12)	(-0.51)	(0.51)	(3.25)	(2.02)	(3.67)	(2.91)
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	323	217	144	366	633	846	829	753	963
Adjusted R-squared	0.074	0.164	0.066	0.080	0.105	0.142	0.130	0.193	0.154

One might be concerned with the endogenous choice of locations by private firms, and argue that the treatment effect is not randomly assigned among issuers. In particular, private firms may deliberately choose to be close to potential investors and be around where the capital is. While I completely agree with this point, I argue that the self-selection problem is only likely to attenuate my results. Location choices of a start-up company could be a rational decision determined by many factors such as the hometown of the founders, the proximity to valuable human capital, the friendliness of the business environment, besides the proximity to capital. For firms that choose to be physically close to potential investors, capital must be one of the most important factors. Hence, one can infer that raising capital should be more crucial for issuers from San Francisco, who choose to be in the same city as numerous VC and PE funds than issuers from North Dakota, who have to cross state borders to reach a large private investor. Therefore, NSMIA, which is designed to facilitate capital raising by private firms, should have a larger impact on issuers from San Francisco than issuers from North Dakota, if these two firms were randomly assigned to the same state. Self-selection thus should weaken my results on Treated × Post. Hence the main finding in this section is unlikely due to self-selection issues, but directly supports H2.[28]

---

[28] When firms select their locations for factors unrelated to capital, for example, the founder's birthplace or the proximity to human capital, there are no issues with the estimation, as these factors are uncorrelated with the different impact of NSMIA on firms in different locations.

Table 2.12. IPO Underpricing and the National Security Market Improvement Act

The table reports coefficients and t-statistics in the parenthesis of OLS regressions of IPO underpricing on the passage of the National Security Market Improvement Act (NSMIA) in October 1996, as shown in regression Equation (5). The sample period is 1993-2000 excluding the years of 1996 and 1997. Expected liquidity is measured by Peer Spread in Panel A, by Peer Turnover in Panel B, and by Peer AIM in Panel C. I run the regression in two times periods: in Columns (1) and (2), the sample is from three years before and after the law passage (1993-2000); and in Columns (3) and (4), the sample is from two years before and after the law passage (1994-1999). I interact the rank (Rank) or the percentage (Percentage) of VC and PE firms of each state shown in Table 2.7 with the dummy variable Post, which is equal to one if the issuer goes public after 1997, and zero if before 1996. For the sake of brevity, estimation results for control variables including Top Underwriter, New Shares Ratio, Integer Price, Sentiment, Log(Assets), Log(1+Age), VC-backed and the intercept are not presented. Variable construction is described in Appendix. The industry fixed effects control for the issuer's industry, defined by the Fama-French 10-industry classification. The t-statistics are computed based on standard errors adjusted for heteroskedasticity and industry-year clustering. Asterisks denote statistical significance at 1% (***), 5% (**), or 10% (*) level.

Panel A: Controlling for peer spread

	1993-2000		1994-1999	
	(1)	(2)	(3)	(4)
Rank × Post	**-0.009*****		**-0.006*****	
	(-4.97)		**(-3.09)**	
Rank	0.002**		0.002**	
	(2.50)		(2.22)	
Percentage × Post		**0.007*****		**0.006****
		(2.80)		**(2.62)**
Percentage		-0.002		-0.001
		(-1.46)		(-0.75)
Post	0.206***	0.066	0.114	-0.005
	(4.21)	(1.31)	(1.65)	(-0.08)
Peer Spread	-3.515***	-3.520***	-3.892***	-3.923***
	(-3.85)	(-3.84)	(-2.87)	(-2.86)
Industry FE	YES	YES	YES	YES
Observations	1,343	1,343	889	889
Adjusted R-squared	0.256	0.256	0.276	0.279

Panel B: Controlling for peer turnover

	1993-2000		1994-1999	
	(1)	(2)	(3)	(4)
Rank × Post	**-0.007*****		**-0.004***	
	(-3.25)		**(-1.97)**	
Rank	0.002***		0.002**	
	(2.81)		(2.18)	
Percentage × Post		**0.005****		**0.005****
		(2.51)		**(2.51)**
Percentage		-0.001		-0.001
		(-1.47)		(-0.62)
Post	0.066	-0.042	0.039	-0.043
	(1.25)	(-0.98)	(1.10)	(-1.16)
Peer Turnover	44.996***	44.985***	45.139***	44.866***
	(4.22)	(4.25)	(3.00)	(3.01)
Industry FE	YES	YES	YES	YES
Observations	1,343	1,343	889	889
Adjusted R-squared	0.327	0.327	0.345	0.347

Panel C: Controlling for peer AIM

	1993-2000		1994-1999	
	(1)	(2)	(3)	(4)
Rank × Post	**-0.009*****		**-0.006*****	
	(-5.21)		**(-3.17)**	
Rank	0.002**		0.002**	
	(2.40)		(2.06)	
Percentage × Post		**0.007*****		**0.006****
		(2.84)		**(2.64)**
Percentage		-0.002		-0.001
		(-1.43)		(-0.63)
Post	0.183***	0.043	0.089	-0.029
	(3.61)	(0.85)	(1.25)	(-0.51)
Peer AIM	-0.092***	-0.092***	-0.102***	-0.102***
	(-4.52)	(-4.55)	(-3.42)	(-3.44)
Industry FE	YES	YES	YES	YES
Observations	1,343	1,343	889	889
Adjusted R-squared	0.269	0.268	0.292	0.295

Another potential selection bias related to the passage of NSMIA is similar to the one I discussed about OHR. NSMIA provides private firms with better access to capital and expands the pool of potential investors, which may affect these firms' incentives to go public. I argue this selection effect should bias against finding a result. NSMIA improves pre-IPO liquidity, thereby reducing the liquidity value of going public. This effect in turn encourages some firms with low liquidity value gain to stay private who would otherwise go public. The exit of these marginal firms from the IPO market improves the quality of the remaining treated firms, creating a positive bias for the liquidity value in the treated IPO sample against finding a negative β_1. Hence, I believe this possible selection effect in fact makes my results stronger.

Lastly, instead of dividing the deals into the control sample and the treatment sample explicitly, I test whether issuers incorporated in states with less number or lower percentage of VC and PE firms experience less underpricing after 1997. I use Rank to denote the rank of the states, and Percentage to denote the percentage of VC and PE firms of the states in US, as presented in Table 2.9. For example, for state NJ, Rank=8 and Percentage=2.66. The regression equation is as follows,

$$\begin{aligned} Underpricing_{ijt} = \alpha_j &+ \beta_1 Rank\ (or\ Percentage) \times Post \\ &+ \beta_2 Rank (or\ Percentage) + \beta_3 Post \\ &+ \beta_4 Expected\ Liquidity_{ijt} + \gamma' X_{ijt} + \varepsilon_{ijt} \qquad (5) \end{aligned}$$

where i, j, t index firms, industries, and years, respectively. Standard errors are clustered at the level of industry-year groups. The coefficient Rank×Post is expected to be negative, and the coefficient on Percentage×Post is expected to be positive. Table 2.12

summarizes the results. In the three panels, I control for expected liquidity with Peer Spread, Peer Turnover, or Peer AIM. The finding is consistent with my prediction. The coefficient on Rank×Post is negative and significant in five out of six specifications, and the coefficient on Percentage×Post is positive and significant in all six specifications. The effect is also economically large. For example, when Rank is increased by one, underpricing post-1997 is lowered by 0.7% to 0.9% in the period of 1993-2000. For every 1% decrease in Percentage, underpricing post-1997 is lowered by 0.5% to 0.6%.

For robustness, I repeat my analysis in Table 2.10 and Table 2.12 by dropping the state of California, which has the highest number of VC and PE firms in the sample period, to rule out the possibility that the results are driven by one particular state. I also repeat the analysis excluding all tech firms to address the concern that results are driven by the tech bubble. The results remain with both exercises.

Discussion: Changes to OHR on Nasdaq and the Passage of NSMIA

I use changes to OHR on Nasdaq in 1997, and the passage of NSMIA in October 1996, as two separate exogenous shocks to issuer's expected liquidity with the difference-in-difference approach. Since these two regulations are close to each other in time, regression Equation (3) and Equation (4) can share the same dummy variable of Post. If the dummy variables indicating the two treatment deals Nasdaq and Treated (NSMIA) are also highly correlated, then empirically these two experiments could be just the same one in nature. For example, if all Nasdaq deals post 1997 are also issuers headquartered in CA and NY, Section 5.4 is just presenting the same set of results with a different name of variable.

To rule out this possibility, and since the variable of interests are Nasdaq×Post and Treated (NSMIA)×Post, I investigate the distribution of the two dummy variables Nasdaq and Treated among observations used in the two regressions of Equations (3) and (4), after 1997 (Post = 1), in the years of 1998, 1999, and 2000. I construct a 2×2 matrix, showing the number of observations in four groups: Nasdaq=1 and Treated (NSMIA)=1, Nasdaq=1 and Treated (NSMIA)=0, Nasdaq=0 and Treated (NSMIA)=1, and Nasdaq=0 and Treated (NSMIA)=0. Since the Treated(NSMIA) dummy changes when I use three alternative control samples: deals in top eight states (CA, NY, MA, TX, IL, CT, PA, and NJ), deals in top four states (CA, NY, MA, and TX), and deals in top two states (CA and NY), I present three panels of this matrix, in Table 2.13.

These matrices show that there is no strong correlation between Nasdaq and Treated (NSMIA). There are more observations with Treated (NSMIA)=1 (177 obs) than Treated (NSMIA)=0 (412 obs) when Nasdaq=1 in Panel A, but the pattern is flipped in Panel C (338 vs. 251 obs). Panel B shows a very balanced distribution of the two variables (247 vs. 342 obs). Hence I conclude that these two regulation changes represent two independent events that I can reply on for separate difference-in-difference analysis.

Table 2.13. Changes to OHR on Nasdaq and the Passage of NSMIA

The table presents the distribution of the two dummy variables Nasdaq and Treated (NSMIA) among observations used in the two regressions of Equations (3) and (4), after 1997 (Post = 1), in the years of 1998, 1999, and 2000. I construct a 2×2 matrix, showing the number of observations in four groups: Nasdaq=1 and Treated (NSMIA)=1, Nasdaq=1 and Treated (NSMIA)=0, Nasdaq=0 and Treated (NSMIA)=1, and Nasdaq=0 and Treated (NSMIA)=0. Panel A defines Treated (NSMIA) in Equation (4) with the control group consisting of deals in CA, NY, MA, TX, IL, CT, PA, and NJ. Panel B defines Treated (NSMIA) in Equation (4) with the control group consisting of deals in CA, NY, MA, and TX. Panel C defines Treated (NSMIA) in Equation (4) with the control group consisting of deals in CA and NY.

Panel A: Control group of NSMIA in top eight states

	Nasdaq=1	*Nasdaq=0*
Treated (NSMIA)=1	177	43
Treated (NSMIA)=0	412	85

Panel B: Control group of NSMIA in top four states firms

	Nasdaq=1	*Nasdaq=0*
Treated (NSMIA)=1	247	56
Treated (NSMIA)=0	342	72

Panel C: Control group of NSMIA in top two states firms

	Nasdaq=1	*Nasdaq=0*
Treated (NSMIA)=1	338	86
Treated (NSMIA)=0	251	42

Conclusion

Traditionally, IPO underpricing has been explained by theories based on asymmetric information. In these theories, underpricing is the compensation to IPO investors for their information disadvantage or a tool of signaling by high-quality firms. In this chapter, I argue that, IPOs enable issuers to access the secondary market via public-listing and enhance firm value, which I call liquidity value.

I conjecture that underpricing is positively related to the expected post-IPO liquidity of the issuer, and negatively related to the issuer's pre-IPO liquidity. I first test a baseline specification investigating the relation between underpricing and expected liquidity. Consistent with the conjecture, I find a positive and significant coefficient when regressing underpricing on expected liquidity. I conduct cross-sectional analysis for the baseline regression, and find that the relation is stronger for issuers with VC investors involved, and when the underwriter has more bargaining power and the fraction of new issuance is smaller.

I then exploit two regulation changes as exogenous shocks to the expected post-IPO liquidity and the pre-IPO liquidity, respectively. The first one is the changes to OHR at Nasdaq in 1997, which improves the liquidity of Nasdaq-listed stocks but not for stocks on other exchanges. With a difference-in-difference approach, I use Nasdaq IPOs as the treatment sample and non-Nasdaq IPOs as the control sample. The result shows that Nasdaq IPOs exhibit more underpricing after the regulation change. The second one is the enactment of NSMIA, which removes compliance burdens for private firms raising capital from different states. I argue that this is an exogenous positive shock to pre-IPO liquidity.

I further conjecture that the law has more impact on issuers located in states where private equity capital is more scarce, proxied by fewer in-state VC and PE firms. Using the issuers in states with less private equity investors as the treatment sample and the ones in states with the more private equity investors as the control sample, I find that the treatment sample shows less underpricing than the control sample post NSMIA.

REFERENCES CITED

Aggarwal, R.K., Krigman, L., and Womack, K.L., 2002. Strategic IPO underpricing, information momentum and lockup expiration selling. Journal of Financial Economics 66, 105-137.

Aggarwal, R.K., Prabhala, N.R., and Puri, M., 2002. Institutional allocation in initial public offerings: empirical evidence. Journal of Finance 57, 1421-1442.

Albuquerque, A., 2009. Peer firms in relative performance evaluation. Journal of Accounting and Economics 48, 69-89.

Allen, F., and Faulhaber, G.R., 1989. Signaling by underpricing in the IPO market. Journal of Financial Economics 23, 303-323.

Amihud, Y., and Mendelson, H., 1986. Asset pricing and the bid-ask spread. Journal of Financial Economics 17, 223–249.

Amihud, Y., & Mendelson, H. 1988. Liquidity and asset prices: Financial management implications. Financial Management, 5-15.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5, 31-56.

Asparouhova, E., Bossaerts, P., Eguia, J. and Zame, W., 2015. Asset pricing and asymmetric reasoning. Journal of Political Economy, 123(1), 66-122.

Asparouhova, E., Bossaerts, P., Roy, N. and Zame, W., 2016. “Lucas” in the Laboratory. Journal of Finance, 71(6), 2727-2780.

Asparouhova, E., Hertzel, M. and Lemmon, M., 2009. Inference from streaks in random outcomes: Experimental evidence on beliefs in regime shifting and the law of small numbers. Management Science, 55(11), 1766-1782.

Bacchetta, P. and Van Wincoop, E., 2005. Rational inattention: A solution to the forward discount puzzle (No. w11633). National Bureau of Economic Research.

Baker, M., and Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. Journal of Finance 61, 1645-1680.

Balakrishnan, K., Billings, M.B., Kelly, B., and Ljungqvist, A., 2014. Shaping liquidity: On the causal effects of voluntary disclosure. Journal of Finance 69, 2237-2278.

Barber, B.M., Lee, Y.T., Liu, Y.J. and Odean, T., 2008. Just how much do individual investors lose by trading?. Review of Financial Studies, 22(2), 609-632.

Barberis, N., Shleifer, A. and Vishny, R., 1998. A model of investor sentiment. Journal of financial economics, 49(3), 307-343.

Barclay, M.J., Christie, W.G., Harris, J.H., Kandel, E., and Schultz, P.H., 1999. Effects of market reform on the trading costs and depth of NASDAQ stocks. Journal of Finance 54, 1–34.

Barclay, M.J. and Litzenberger, R.H., 1988. Announcement effects of new equity issues and the use of intraday price data. Journal of Financial Economics, 21(1), 71-99.

Beatty, R.P., and Welch, I., 1996. Issuer expenses and legal liability in initial public offerings. Journal of Law and Economics 39, 545-602.

Bechara, A. and Damasio, H., 2002. Decision-making and addiction (part I): impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future consequences. Neuropsychologia, 40(10), 1675-1689.

Belt, D., 2018. Pre-IPO liquidity for late stage start-ups. Fenwick & West LLP.

Benveniste, L.M., and Spindt, P., 1989. How investment bankers determine the offer price and allocation of new issues. Journal of Financial Economics 24, 343-361.

Bessembinder, H., 1999. Trade execution costs on NASDAQ and the NYSE: A post-reform comparison. Journal of Financial and Quantitative Analysis 34, 387– 407.

Bessembinder, H., and Kaufman, H.M., 1997. A comparison of trade execution costs for NYSE and NASDAQ-listed stocks. Journal of Financial and Quantitative Analysis 32, 287-310.

Black, F., 1986. Noise. Journal of Finance, 41(3), 528-543.

Black, B.S., and Gilson, R.J., 1998. Venture capital and the structure of capital markets: Banks versus stock markets. Journal of Financial Economics 47, 243-277.

Booth, J.R., and Chua, L., 1996. Ownership dispersion, costly information, and IPO underpricing. Journal of Financial Economics 41, 291-310.

Bradley, D.J., Cooney, J.W., Jordan, B.D., and Singh A.K., 2004. Negotiation and the IPO offer price: A comparison of integer vs. non-integer IPOs. Journal of Financial and Quantitative Analysis 39, 517–540.

Bradley, D.J., and Jordan, B.D., 2002. Partial adjustment to public information and IPO underpricing. Journal of Financial and Quantitative Analysis 37, 595-616.

Busse, J.A. and Green, T.C., 2002. Market efficiency in real time. Journal of Financial Economics, 65(3), 415-437.

Calderón-Garcidueñas, L., Mora-Tiscareño, A., Ontiveros, E., Gómez-Garza, G., Barragán-Mejía, G., Broadway, J., Chapman, S., Valencia-Salazar, G., Jewells, V., Maronpot, R.R. and Henríquez-Roldán, C., 2008. Air pollution, cognitive deficits and brain abnormalities: a pilot study with children and dogs. Brain and cognition, 68(2), 117-127.

Carter, R.B., and Manaster, S., 1990. Initial public offerings and underwriter reputation. Journal of Finance 45, 1045-1068.

Chang, S.C., Chen, S.S., Chou, R.K. and Lin, Y.H., 2008. Weather and intraday patterns in stock returns and trading activity. Journal of Banking & Finance, 32(9), 1754-1766.

Chemmanur, T.J., 1993. The pricing of initial public offerings: a dynamic model with information production. Journal of Finance 48, 285-304.

Chemmanur, T.J., and Fulghieri, P., 1999. A theory of the going-public decision. Review of Financial Studies 12, 249-279.

Chordia, T., Roll, R., and Subrahmanyam, A., 2000. Commonality in liquidity. Journal of Financial Economics 56, 3–28.

Chordia, T., Roll, R., and Subrahmanyam, A., 2001. Market liquidity and trading activity. Journal of Finance 56, 501–530.

Christie, W.G., and Schultz, P.H., 1994. Why do Nasdaq market makers avoid odd-eighth quotes? Journal of Finance 49, 1813-1840.

Cornelli, F., and Goldreich, D., 2002. Bookbuilding and strategic allocation. Journal of Finance 56, 2337-2369.

Coval, J.D. and Moskowitz, T.J., 1999. Home bias at home: Local equity preference in domestic portfolios. Journal of Finance, 54(6), 2045-2073.

De Fontenay, E., 2017. The Deregulation of Private Capital and the Decline of the Public Company. Hastings Law Journal 68, 445-502.

Dorn, D. and Sengmueller, P., 2009. Trading as entertainment?. Management Science, 55(4), 591-603.

Dorn, A.J., Dorn, D. and Sengmueller, P., 2014. Trading as gambling. Management Science, 61(10), 2376-2393.

Dunbar C.G., and Foerster, S.R., 2008. Second time lucky? Withdrawn IPOs that return to the market. Journal of Financial Economics 87, 610-635.

Edmans, A., Garcia, D. and Norli, Ø., 2007. Sports sentiment and stock returns. Journal of Finance, 62(4), 1967-1998.

Ellul, A., and Pagano, M., 2006. IPO underpricing and after-market liquidity. Review of Financial Studies 19, 381–421.

Erber, R. and Tesser, A., 1992. Task effort and the regulation of mood: The absorption hypothesis. Journal of Experimental Social Psychology, 28(4), 339-359.

Ewens, M., and Farre-Mensa, J., 2017. The evolution of the private equity market and the decline in IPOs. Working Paper.

Feng, L. and Seasholes, M.S., 2004. Correlated trading and location. Journal of Finance, 59(5), 2117-2144.

Field, L.C., and Sheehan, D.P., 2004. IPO underpricing and outside blockholdings. Journal of Corporate Finance 10, 263-280.

Fonken, L.K., Xu, X., Weil, Z.M., Chen, G., Sun, Q., Rajagopalan, S. and Nelson, R.J., 2011. Air pollution impairs cognition, provokes depressive-like behaviors and alters hippocampal cytokine expression and morphology. Molecular psychiatry, 16(10), p.987.

Gao, X. and Lin, T.C., 2014. Do individual investors treat trading as a fun and exciting gambling activity? Evidence from repeated natural experiments. Review of Financial Studies, 28(7), 2128-2166.

Gargano, A. and Rossi, A.G., 2018. Does it pay to pay attention?. Review of Financial Studies, 31(12), 4595-4649.

Grinblatt, M. and Keloharju, M., 2009. Sensation seeking, overconfidence, and trading activity. Journal of Finance, 64(2), 549-578.

Habib, M., and Ljungqvist, A., 2001. Underpricing and entrepreneurial wealth losses: Theory and Evidence. Review of Financial Studies 14, 433-458.

Hanley, K.W., 1993. The underpricing of initial public offerings and the partial adjustment phenomenon. Journal of Financial Economics 34: 231-250.

Hirshleifer, D. and Shumway, T., 2003. Good day sunshine: Stock returns and the weather. Journal of Finance, 58(3), 1009-1032.

Hong, H., Torous, W. and Valkanov, R., 2007. Do industries lead stock markets?. Journal of Financial Economics, 83(2), 367-396.

Hu, X., Li, O.Z. and Lin, Y., 2014. Particles, pollutions and prices. Available at SSRN 2525980.

Huang, L. and Liu, H., 2007. Rational inattention and portfolio selection. Journal of Finance, 62(4), 1999-2040.

Huang, R.D., and Stoll, H.R., 1996. Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE. Journal of Financial Economics 41, 313-357.

Huberman, G., and Halka, D., 2001. Systematic liquidity, Journal of Financial Research 24, 161–178.

Huberman, G., 2001. Familiarity breeds investment. Review of Financial Studies, 14(3), 659-680.

Hughes, P.J., and Thakor, A.V., 1992. Litigation risk, intermediation, and the underpricing of initial public offerings, Review of Financial Studies 5, 709-742.

Kamstra, M.J., Kramer, L.A. and Levi, M.D., 2003. Winter blues: A SAD stock market cycle. American Economic Review, 93(1), 324-343.

Kamstra, M.J., Kramer, L.A. and Levi, M.D., 2015. Seasonal variation in Treasury returns. Critical Finance Review. 4(1), 45-115

Kamstra, M.J., Kramer, L.A., Levi, M.D. and Wermers, R., 2017. Seasonal asset allocation: Evidence from mutual fund flows. Journal of Financial and Quantitative Analysis, 52(1), 71-109.

Karlsson, N., Loewenstein, G. and Seppi, D., 2009. The ostrich effect: Selective attention to information. Journal of Risk and uncertainty, 38(2), 95-115.

Koudijs, P., 2016. The boats that did not sail: Asset price volatility in a natural experiment. Journal of Finance, 71(3), 1185-1226.

Kramer, L.A. and Weber, J.M., 2012. This is your portfolio on winter: Seasonal affective disorder and risk aversion in financial decision making. Social Psychological and Personality Science, 3(2), 193-199.

Krigman, L., Shaw, W.H., and Womack, K.L., 2001. Why do firms switch underwriters? Journal of Financial Economics 60, 245-284.

Kuhnen, C.M. and Knutson, B., 2011. The influence of affect on beliefs, preferences, and financial decisions. Journal of Financial and Quantitative Analysis, 46(3), 605-626.

Kumar, A., 2009. Who gambles in the stock market?. Journal of Finance, 64(4), 1889-1933.

Kyle, A., 1985. Continuous auctions and insider trading. Econometrica 53, 1315-3-1335.

Lee, C.M., Shleifer, A., and Thaler, R.H., 1991. Investor sentiment and the closed end fund puzzle. Journal of Finance 46, 75–109.

Lee, P.M., and Wahal, S., 2004. Grandstanding, certification and the underpricing of venture capital backed IPOs. Journal of Financial Economics 73, 375-407.

Li, Q. and Peng, C.H., 2016. The stock market effect of air pollution: evidence from China. Applied Economics, 48(36), 3442-3461.

Liu, J., Stambaugh, R.F. and Yuan, Y., 2019. Size and value in China. Journal of Financial Economics.

Liu, X., and Ritter, J.R., 2011. Local underwriter oligopolies and IPO underpricing. Journal of Financial Economics 102, 579-601.

Ljungqvist, A., Nanda, V., and Singh, R., 2006. Hot markets, investor sentiment, and IPO pricing. Journal of Business 79, 1667–1703.

Logue, D.E., Rogalski, R., Seward, J., and Foster-Johnson, L., 2002. What's special about the role of underwriter reputation and market activities in IPOs? Journal of Business 75, 213-243.

Loughran, T., and Ritter, J.R., 2002. Why don't issuers get upset about leaving money on the table in IPOs? Review of Financial Studies 15, 413-443.

Loughran, T. and Schultz, P., 2004. Weather, stock returns, and the impact of localized trading behavior. Journal of Financial and Quantitative Analysis, 39(2), 343-364.

Loughran, T., and Ritter, J.R., 2004. Why has IPO underpricing changed over time? Financial management 33, 5-37.

Lowry, M., and Schwert, G.W., 2004. Is the IPO pricing process efficient? Journal of Financial Economics 71, 3-26.

Lucas, D., and McDonald, R., 1990. Equity issues and stock price dynamics. Journal of Finance 45, 1019-1043.

Maksimovic, V., and Pichler, P., 2001. Technological innovation and initial public offerings. Review of Financial Studies 14, 459-494.

Milgrom, P., and Stokey, N., 1982. Information, trade and common knowledge. Journal of Economic Theory 26, 17-27.

Newey, W. and West, K., 1987. A simple positive definite, heteroscedasticity and autocorrelation consistent covariance matrix estimator. Econometrica, 55, p.103080.

Newey, W. and West, K., 1994. Automatic lag selection in covariance matrix estimation. The Review of Economic Studies, 61(4), 631-653.

Payzan-LeNestour, E. and Woodford, M., 2018. 'Outlier Blindness': Efficient Coding Generates an Inability to Represent Extreme Values. Working Paper.

Peng, L. and Xiong, W., 2006. Investor attention, overconfidence and category learning. Journal of Financial Economics, 80(3), 563-602.

Pinegar, J.M., 2002. Losing sleep at the market: Comment. American Economic Review, 92(4), 1251-1256.

Pirinsky, C. and Wang, Q., 2006. Does corporate headquarters location matter for stock returns?. Journal of Finance, 61(4), 1991-2015.

Rabin, M., 2002. Inference by believers in the law of small numbers. The Quarterly Journal of Economics, 117(3), 775-816.

Ritter, J.R., 1987. The costs of going public. Journal of Financial Economics 19, 269-281.

Ritter, J.R., and Welch, I., 2002. A review of IPO activity, pricing, and allocations, Journal of Finance 57, 1795-1828.

Rock, K., 1986. Why new issues are underpriced. Journal of Financial Economics 15, 187-212.

Saunders, E.M., 1993. Stock prices and Wall Street weather. The American Economic Review, 83(5), 1337-1345.

Seasholes, M.S. and Zhu, N., 2010. Individual investors and local bias. Journal of Finance, 65(5), 1987-2010.

Sicherman, N., Loewenstein, G., Seppi, D.J. and Utkus, S.P., 2015. Financial attention. Review of Financial Studies, 29(4), 863-897.

Suglia, S.F., Gryparis, A., Schwartz, J. and Wright, R.J., 2007. Black carbon associated with cognition among children in a prospective birth cohort study. Epidemiology, 18(5), p.S163.

Welch, I., 1989. Seasoned offerings, imitation costs, and the underpricing of initial public offerings. Journal of Finance 44, 421-449.

Zingales, L., 1995. Insider Ownership and the decision to go public. Review of Financial studies 62, 425-448.

APPENDIX A

CHAPTER 1 VARIABLE DEFINITION

Variables	Definition	Source
Log(Volume_t)	Log of hourly trading volume of the portfolio.	WIND
Log(Volume_(t-1, t-55))	Log of the average of past 55 hourly trading volume of the portfolio. There are five trading hours each day. This variable measures the average trading volume of the past 11 days.	WIND
Smog_t	Hourly smog at hour t. It is measured by PM2.5 concentration value during an hour.	US Embassy in China.
Smog_t-1	Hourly smog at hout t-1.	US Embassy in China.
Smog_t-2	Houlry smog at hour t-2.	US Embassy in China.
Smog_t-3	Houlry smog at hour t-3.	US Embassy in China.
Very Unhealthy_t	Dummy variable take 1 if PM2.5 concentration value is > 150.5, otherwise 0.	
Hour9	Dummy variable takes 1 if it is 9:00, otherwise 0.	
Hour10	Dummy variable takes 1 if it is 10:00, otherwise 0.	
Hour11	Dummy variable takes 1 if it is 11:00, otherwise 0.	
Hour13	Dummy variable takes 1 if it is 13:00, othewrise 0.	
Hour14	Dummy variable takes 1 if it is 14:00, otherwise 0.	
MON	Dummy variable takes 1 if it is Monday, otherwise 0.	
TUE	Dummy variable takes 1 if it is Tuesday, otherwise 0.	
WED	Dummy variable takes 1 if it is Wednesday, otherwise 0.	
THR	Dummy variable takes 1 if it is Thursday, otherwise 0.	

APPENDIX B

CHAPTER 1 AIR QUALITY INDEX (AQI) AND PM2.5 VALUE

Category	AQI	PM2.5	PM2.5 Health Effects Statement	PM2.5 Cautionary Statement
Good	0-50	0.0-12.0	PM2.5 air pollution poses little or no risk.	None
Moderate	51-100	12.1-35.4	Unusually sensitive individuals may experience respiratory symptoms.	Unusually sensitive people should consider limiting prolonged outdoor exertion.
Unhealthy for Sensitive Groups	101-150	35.5-55.4	Increasing likelihood of respiratory symptoms in sensitive individuals, aggravation of heart or lung disease and premature mortality in persons with cardiopulmonary disease and the elderly.	Active children and adults, and people with respiratory disease, such as asthma, should limit prolonged outdoor exertion.
Unhealthy	151-200	55.5-150.4	Increased aggravation of heart or lung disease and premature mortality in persons with cardiopulmonary disease and the elderly; increased respiratory effects in general population.	Active children and adults, and people with respiratory disease, such as asthma, should avoid prolonged outdoor exertion; everyone else, especially children, should limit prolonged outdoor exertion.
Very Unhealthy	201-300	150.5-250.4	Significant aggravation of heart or lung disease and premature mortality in persons with cardiopulmonary disease and the elderly; significant increase in respiratory effects in general population.	Active children and adults, and people with respiratory disease, such as asthma, should avoid all outdoor exertion; everyone else, especially children, should limit outdoor exertion.
Hazardous	301-500	250.5-350.4	Serious aggravation of heart or lung disease and premature mortality in persons with cardiopulmonary disease and the elderly; serious risk of respiratory effects in general population.	Everyone should avoid all outdoor exertion.
Beyond	>500	350.5-500	Extremely high levels of PM2.5	

APPENDIX C

CHAPTER 2 VARIABLE DEFINITION

Variables	Definition	Source
Offer Price	IPO Offer Price	SDC
Underpricing	Percentage change from the offer price to the first-day closing price	SDC & CRSP
Peer Spread	I select the issuer's peer firms as the publicly traded ones with the same industry (SIC 2-digit code), similar size, and listed on the same exchange. Daily spreads of each peer firm is calculated in the 12 months preceding the IPO time. I then take an average of the daily spreads, and average across peer firms to construct peer spread.	CRSP
Peer Turnover	I select the issuer's peer firms as the publicly traded ones with the same industry (SIC 2-digit code), similar size, and listed on the same exchange. Daily turnover of each peer firm is calculated in the 12 months preceding the IPO time. I then take an average of the daily turnover, and average across peer firms to construct peer turnover.	CRSP
Peer AIM	I select the issuer's peer firms as the publicly traded ones with the same industry (SIC 2-digit code), similar size, and listed on the same exchange. Following Amihud (2002), I use daily CRSP data (CRSP variables *ret, prc,* and *vol*) to calculate the ratio of absolute stock return to dollar volume [10,000,000 × | *ret* | ÷ (*prc* × *vol*)] for each day in the 12-month period before the IPO for each peer firm. I then average over the period and average across peer firms, and the final measure is the natural log of one plus the average peer AIM.	CRSP
Top Underwriter	A dummy variable that is equal to one if the lead underwriter has an updated Carter and Manaster's (1990) rank of eight or more, and zero otherwise.	Jay Ritter's website
Integer Price	A dummy variable that is equal to one if the offer price is an integer and zero otherwise.	SDC

New Shares Ratio	It is the fraction of the ownership the issuer sells during the IPO. It is calculated as IPO proceeds / (IPO price × number of shares outstanding).	SDC & CRSP
Sentiment	It's the monthly market sentiment index based on the closed-end fund discount, the NYSE share turnover, the number of IPOs, the share of equity issuance in total equity and debt issuance, and the dividend premium, constructed in Equation (2) of Baker and Wurgler (2006).	Prof. Jeffrey Wurgler's website
Assets	Firm’s pre-issue book value of assets, in millions of dollars.	SDC
Age	Calendar year of offering minus the calendar year of founding.	Jay Ritter's website
VC-backed	Equals one (zero otherwise) if the IPO was backed by venture capital.	SDC
Peer Age	Number of years that the peer firm has been public.	CRSP
Sales	Peer firm's annual sales, in millions.	COMPUSTAT
Market Cap	I first obtain daily market capitalization, which equals daily price times number of shares outstanding. I then take the average of daily values within a given month to reach monthly value.	CRSP
Shareholders	Number of shareholders, in millions	COMPUSTAT
Market Return	NYSE/AMEX/NASDAQ/ARCA monthly market return.	CRSP
Lagged Market	Market return of the same month in the previous year.	CRSP
Variance of Market	Square of the standard deviation of daily market return within a given month.	CRSP
Interest Rate	Monthly three-month treasury bill rate.	The Federal Reserve