Three Essays in Empirical Corporate Finance

Changjie Hu
Desautels Faculty of Management
McGill University
Montreal, Quebec
August 2020

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Table of Contents

Abstract

The core of the thesis includes three essays in empirical corporate finance. The first essay examines the relation between mandatory disclosure behavior and legal accountability. In this study, we treat the enactment of the Sarbanes-Oxley Act (SOX) in 2002 as a regulatory event that increases the legal accountability of top executives and compute the filing tones for a large sample of Forms 10-Q and 10-K filings between 1994 and 2017 using textual analysis. We document that the changes in filing tones contain substantial information that is reflected promptly in the capital market. We also show that a structural break exists in the distribution of filing tones around SOX. Firms use a more negative tone in their quarterly mandatory disclosure after SOX. Interestingly, investors exhibit a stronger reaction to per unit change of filing tones during the post-SOX era and we show that changes in investors' reactions are not merely driven by the systematic changes in tone distribution after SOX. We also document that filing tones are determined by common performance measures, but such relation is weakened after SOX.

The second essay studies the impact of the exit of Venture Capitalists (VCs) on innovation by comparing VC backed IPO firms with the non-VC backed. VCs play a significant role in bringing new ventures public by providing financing and consistent monitoring. Prior literature has established mostly a positive correlation between VCs and firm innovation because VCs may preselect more innovative firms to begin with. This study hopes to provide evidence on causal inference with reasonable assumptions from a "reverse treatment" perspective by examining the change in innovation when VCs exit. We treat the initial public offering (IPO) as a proxy for VC's exit since most VCs exit shortly after IPO due to their limited investment horizon. Using a difference-in-differences framework, we find that VC-backed firms experience a greater drop in

Research and Development (R&D) intensity after IPO-exits when compared to those non-VC backed.

The third essay revisits the long-debated relation between market competition and firm innovation. While traditionally competition is measured at the industry level with historical data, our study utilizes two new text-based measures of competitive threats developed by Hoberg *et al*. (2014) and Li *et al*. (2013) which are both firm-specific and forward-looking. We address the potential endogeneity concerns using instrumental variables along with the propensity score matching of firms that experience an exogenous shock from import competition with those that do not. Our results show that an increase in competition unambiguously promotes firm innovation.

Résumé

Le essentiel de la thèse comprend trois essais en finance d'entreprise empirique. Le premier essai examine la relation entre le comportement de divulgation obligatoire et la responsabilité juridique. Dans cette étude, nous traitons la promulgation de la loi Sarbanes-Oxley (SOX) en 2002 comme un événement réglementaire qui augmente la responsabilité juridique des cadres supérieurs et calcule les tonalités de dépôt pour un large échantillon de formulaires 10-Q et 10-K entre 1994 et 2017 à l'aide d'une analyse textuelle. Nous documentons que les changements dans les tonalités de dépôt contiennent des informations substantielles qui se reflètent rapidement sur le marché des capitaux. Nous montrons également qu'il existe une rupture structurelle dans la distribution des tonalités de dépôt autour de SOX. Les entreprises utilisent un ton plus négatif dans leur divulgation obligatoire trimestrielle après SOX. Il est intéressant de noter que les investisseurs présentent une réaction plus forte au changement unitaire des tonalités de dépôt au cours de l'ère post-SOX et nous montrons que les changements dans les réactions des investisseurs ne sont pas simplement du aux changements systématiques de la distribution des tons après SOX. Nous documentons également que les tonalités de classement sont déterminées par des mesures de performance courantes, mais cette relation est affaiblie après SOX.

Le deuxième essai étudie l'impact de la sortie de capital-risque (VC) sur l'innovation en comparant les entreprises introduites en bourse financées par capital-risque avec les sociétés non financées par capital-risque. Les VC jouent un rôle important en rendant publiques de nouvelles entreprises en fournissant un financement et un suivi cohérent. La littérature antérieure a établi principalement une corrélation positive entre les sociétés de capital-risque et l'innovation des entreprises, car les sociétés de capital-risque peuvent présélectionner des entreprises plus innovantes au départ. Cette étude espère fournir des preuves sur l'inférence causale avec des hypothèses raisonnables dans une

perspective de «traitement inverse» en examinant le changement dans l'innovation à la sortie de VC. Nous traitons l'offre publique initiale (IPO) comme un proxy pour la sortie de VC, car la plupart des VCs sortent peu de temps après l'introduction en bourse en raison de leur horizon d'investissement limité. En utilisant un cadre de différence dans les différences, nous constatons que les entreprises soutenues par VC connaissent une baisse plus forte de l'intensité de la recherche et du développement (R&D) après les introductions en bourse par rapport à celles qui ne sont pas soutenues par VC.

Le troisième essai revient sur la relation longuement débattue entre la concurrence sur le marché et l'innovation des entreprises. Bien que la concurrence soit traditionnellement mesurée au niveau de l'industrie avec des données historiques, notre étude utilise deux nouvelles mesures textuelles des menaces concurrentielles développées par Hoberg *et al.* (2014) et Li *et al*. (2013) qui sont à la fois spécifique à l'entreprise et tournées vers l'avenir. Nous abordons les problèmes potentiels d'endogénéité en utilisant des variables instrumentales ainsi que l'appariement des scores de propension des entreprises qui subissent un choc exogène de la concurrence des importations avec celles qui n'en subissent pas. Nos résultats montrent qu'une augmentation de la concurrence favorise sans ambiguïté l'innovation des entreprises.

Acknowledgments

First, I am especially grateful to my supervisors, Prof. Matthieu Bouvard and Prof. Adolfo de Motta, for their unwavering support and patient guidance during my doctoral study, and I would not have come this far without their supervision. I am also deeply indebted to my committee members, Prof. Sandra Betton, Prof. Alexandre Jeanneret, Prof. Jiro Kondo, and Prof. Jingjing Zhang for their insightful comments and kind encouragement. My sincerest gratitude also goes to Prof. Harjeet Bhabra, Prof. James Brander, Prof. Saraswata Chaudhuri, Prof. Alfred Davis, Prof. Thomas Rivera, Prof. David Schumacher, and Prof. Wei Wang for their invaluable comments as well as all my peers at Desautels who engage me in stimulating discussions concerning my research ideas and the challenges that I faced in my research from time to time. I also thank the discussant, Antoine Noel, and all the participants at the McGill-HEC Doctoral Workshop and the 2019 International Conference on Financial Development and Stability in Dynamic Global Economy. Special thanks go to McGill's accounting department for hosting a workshop on textual analysis that is especially useful for my theses. Last but not least, I would like to thank Desautels School of Management and McGill for providing continuous support and resources during my study. Without consistent help from the Ph.D. office, Prof. Samer Faraj, Stella Scalia, and Yanyan Shu, smooth transitions among various stages of my doctoral study will not be possible. All remaining mistakes in the thesis are my own.

Contribution of Authors

All three essays in this thesis are single-authored by Changjie Hu. They are completed under the guidance of Changjie’s supervisor, Prof. Adolfo de Motta, and ex-supervisor, Prof. Matthieu Bouvard. Changjie has also received feedback and comments from his committee members. Earlier versions of Chapter 3 were presented at the McGill-HEC Doctoral Workshop and the 2019 International Conference on Financial Development and Stability in Dynamic Global Economy.

Chapter 1. Introduction

Theoretical studies are fundamental to the understanding of any problems and relationships in the field of corporate finance research. Theories are built to predict and explain phenomena observed within firms and the economy based on reasonable assumptions. However, empirical research is equally pivotal in validating and complementing existing theoretical predictions. Today's economy constitutes firms of different sizes under the regulation and facilitation of various agencies. It is always important to understand empirically what is fueling the growth of these firms as well as to evaluate regulatory changes aimed at curbing undesirable behaviors that might undermine the foundation of our economy.

Governance and financial crises are often followed by policies aimed at improving information transparency. The unexpected accounting frauds that caused the bankruptcy of Enron and WorldCom back in the early 2000s inevitably led to the rapid enactment of the Sarbanes-Oxley Acts (SOX). An important requirement from SOX is that the top executives have to certify financial statements and become directly accountable for any misrepresentation of information. While one major objective of SOX is to improve the transparency and accuracy of firm disclosures, the effect of making top executives accountable for misreporting is not a priori unambiguous. Our first essay treats the enactment of SOX in 2002 as a regulatory event that increases the legal accountability of top executives across the board for all public firms in the US. Specifically, we use computational linguistics to capture managerial tones from Form 10-Q and 10-K filings and investigate whether there is any structural change in the informational content before and after SOX.

Our two subsequent essays revolve around the driving factors behind firm innovation. Innovation is deemed the key driver of growth and it is on-going at every stage of a firm's life cycle. In our second essay, we explore the relation between Venture Capitalists (VCs) and firm innovation. While prior studies have mainly documented a positive correlation when VCs enter, we examine their relation from a "reverse treatment" perspective (i.e., when VCs exit through IPOs). While the presence of VCs provides monitoring, experiences, and other useful resources that promote firm innovation, they eventually exit with all these positive influences.

Our third essay revisits the long-debated relation between market competition and firm innovation. As noted by prior literature, the finding of a true effect can be hampered by the complexity in market structures, characteristics of innovation, and dynamics of discovery.[1] Empirically, any effects found may also be further confounded by endogeneity concerns and measurement errors. In this essay, we utilize two new text-based measures of competitive threats developed by Hoberg *et al.* (2014) and Li *et al.* (2013), which are arguably more superior in capturing real forward-looking competitive threats faced by individual firms than traditional measures such as Herfindahl index (*HHI*) and Concentration Ratio (*CR*). We also use tariff rates and exchange rates as the instrumental variables to address potential endogeneity and measurement errors.

The abovementioned three essays constitute the next three chapters of this thesis respectively. Each of the three chapters is structured in a way self-contained with its own introduction, literature review, methodology, empirical findings, and conclusion. Chapter 5 concludes the entire thesis.

[1] See Kamien and Schwartz (1975) and Gilbert (2006).

Chapter 2. Accountability and the Disclosure of Soft Information: Evidence from the Sarbanes-Oxley Act

1. Introduction

Governance and financial crises are often followed by new legislation that aims to enhance transparency. While balance sheet numbers directly reflect the financial condition of a firm, qualitative statements also contain valuable soft information for the investors.[2] The strategic use of qualitative disclosure to convey information not directly apparent from the financial numbers have been prevalent in the forms of corporate filings, press conferences, and earnings announcements. However, it has been long discussed whether managers should be held legally liable for qualitative disclosure since qualitative statements may also be considered equally material (e.g., Brudney, 1989; O'Hare, 1998; Roussel, 1998; Huang, 2005; Hoffman, 2005; Padfield, 2007; Rogers *et al*., 2011). In this paper, we document how firms' disclosure behavior of soft information changes when the top executives became legally liable. Specifically, we treat the enactment of the Sarbanes-Oxley Act (SOX) in 2002 as a regulatory event that explicitly imposes legal penalties to the top executives (i.e., CEO and CFO) for inaccurate disclosure of financial reports and internal control structure (SOX, Section 302 and 906). The legal costs of making untruthful disclosure, therefore, increase after SOX. While prior studies have extensively explored the relation between voluntary disclosure and litigation risk, the conclusion is mixed at best. Therefore, the effect of making key personnel accountable for misrepresentation of

[2] In this paper we do not distinguish qualitative information with soft information and use them interchangeably.

information is a priori largely ambiguous. In this paper, we explore potential changes in firms' disclosure practices under heightened legal accountability.

The management has more flexibility over the choice of words in qualitative statements relative to financial statement numbers when it comes to filing. While explicit "management" of accounting numbers for any cause can easily violate security laws, the qualitative sections leave sufficient channels for the management to convey information that best suits their interests. In this article, we gauge managerial sentiments from firms' 10-K and 10-Q filings through computational linguistics and study the reporting behavior of the firms as well as the investors' reaction around the filing date during the pre-SOX and post-SOX periods. The enactment of SOX in 2002 aims to improve both the quality of financial reporting and investor confidence, which was deemed an inevitable remedy after a series of major corporate and accounting scandals (e.g., Enron and WorldCom) that fundamentally undermined the trust from the investors. The top executives face additional personal legal costs while making disclosure decisions after the enactment of SOX.[3] On one hand, a firm may reveal more truthfully to preempt lawsuits for hiding material information. On the other hand, it may also reduce overall disclosure to deter potential lawsuits and legal penalties. There exists a tension on how firms strategically modify their reporting behavior under greater regulatory scrutiny and heightened legal liability. It is also unclear how the investors update their beliefs accordingly to evaluate potentially subtle changes in reporting style or content. This paper first aims to shed some light on the way manager reports when the top executives face greater scrutiny and heightened legal consequences.

[3] Bamber et al. (2010) document that top executives influence the style of financial disclosure.

To systematically capture the soft information from 10-Q or 10-K filings, we use Loughran and McDonald (2011) word lists that share common sentiments (e.g., positive and negative) under business settings. More specifically, we tabulate the occurrence of sentimental words and derive various measures for filing tones. Using a large sample of 10-Q and 10-K filings between 1994 and 2017, we document a structural change in the way firms disclose: We observed an increasing trend in the negative tone which peaks and plateaus out after SOX. The positive tone seems to increase consistently over our full sample period but exhibits slightly more variation after SOX. Correspondingly, the net tone of an average firm exhibits a downward trend before SOX and remains relatively constant after SOX. Overall, the firms seem to be more conservative (negative) in their filings and show less variations after SOX when top executives face greater personal legal costs for misrepresentation. We then examine the information content of 10-Q or 10-K filings around the enactment of SOX. While negative and net tone predicts abnormal returns or future earnings, the positive tone does not.[4] We also observe that a positive change in net tone increases idiosyncratic return volatility, but such impact becomes negative in the post-SOX period. This result suggests that investors under a stricter regulatory environment perceive a positive change in net tone as a favorable disclosure that reduces uncertainty which in turn lower idiosyncratic return volatility. Similarly, a positive change in negative (positive) tone may be viewed as a negative (positive) disclosure that increases (decreases) return volatility during the post-SOX period. We also observe that tone changes have significant predictive effects in the subsequent earnings and such effect is stronger after SOX. The presence of significant alphas in the strategy of constructing zero-cost portfolios sorted based on changes in tone also confirmed our presumption that filing tones contain substantial forward-looking information. However, the alphas do not seem to differ

[4] It is challenging to differentiate whether the use of positive words is for good firm performance or merely to negate negative news. A positive correlation between negative and positive tones suggests the possibility of the latter.

between the pre- and post-SOX periods, suggesting no additional information being revealed in determining the relative performance ranking of the firms.

This paper studies how the investors respond before and after SOX as one might expect them to react correspondingly to changes in 10-K and 10-Q filings. More specifically, we examine the cumulative abnormal return (*CAR*) around individual filing dates. Our results show that investors react more vigorously to a per-unit change in tones after SOX. This is consistent with investors perceiving tone changes as having more information after the enactment of SOX.[5] The results are robust after controlling for important variables that are thought to affect post-filing abnormal returns such as accruals, earnings surprises, past returns, and readability. We also rule out the possibility that our results are driven by the potential "scaling effect" by using standardized tone variables as suggested in Tetlock *et al*. (2008).[6]

This paper also evaluates the determinants of filing tones for 10-K and 10-Q filings and studies how the passage of SOX affects the way firm managers report. We find that the level and the change of filing tone are closely related to the firm's current fiscal quarter performance. Profitability measures such as earnings are always the key focal points when managers draft quarterly reports. Earning volatility, Tobin's Q, leverage, and past returns are also important determinants of tones. This is not surprising since the goal of managers while constructing 10-K or10-Q reports is to convey truthful and accurate information on firms' operations and financial status to the stakeholders. Interestingly, we find a moderating effect on the relationship between

[5] Coates and Srinivasan (2014) caution that any effects documented under this setup may not have been caused by SOX per se due to the lack of a comparable unaffected control group. A potential argument for observed effects can be due to market discipline following a period where internal controls were deemed collapsed. We fully acknowledge the difficulty to arrive at causal interpretation.

[6] It is possible that firms systematically adjust reporting style (e.g. lower tone variances) after SOX without changing the actual amount of information for a given level of firm performance. Investors may offset such a change in scale by adjusting their beliefs accordingly.

profitability measures and managerial tones after SOX. The relation between quantifiable accounting information and filing tones are weakened after SOX. This is consistent with the evidence that there is a decrease in the dispersion of tones or changes of tones after SOX. In essence, we find that managers appear to be more conservative in discussing both good and bad news in the mandatory filings. The smoothing of filing tones may be deemed an effective way by the managers to deter unwanted scrutiny from the investors once a firm's top personnel becomes directly accountable for any misrepresentation of the firm's operational and financial circumstances after SOX.

The rest of this paper is organized as follows: Section 2 reviews the existing literature. Section 3 explains the process of sample construction and empirical methodology. Section 4 discusses the main results. Section 5 concludes.

2. Literature Review

Numerous studies have shown the inadequacy of quantitative information alone to explain returns (e.g., Shiller, 1981; Roll, 1988; Lev and Thiagarajan, 1993; Amir and Lev, 1996). Recent research has started to examine qualitative information as complementary factors in explaining returns. For example, Mayew and Venkatachalam (2012) extracted the audio from earnings conference calls to determine managerial affective states and study their correlation with the contemporaneous stock returns. Feldman *et al.* (2010) find that changes in filing tones provide incremental information on stock prices. From the literature of voluntary disclosure, the relationship between litigation risk and disclosure practice is still largely in a debate.[7] However, it is largely unclear how holding top executives legally liable might affect firms' disclosure choice on mandatory filings. While there is less flexibility on mandatory filings as compared to voluntary disclosure, it is undoubtedly a crucial channel for managers to convey information to all the stakeholders.

Related to our study, Arping and Sautner (2013) find that cross-listed firms in the U.S. became significantly more transparent following SOX than their European counterparts by looking at the analyst earnings forecasts. Another paper by Rice and Weber (2012), however, warns that SOX alone might not achieve its intended goal of identifying existing control weaknesses due to intricate incentives of the firms and external auditors.[8] Our study first contributes to the assessment of SOX's impact. More specifically, we examine whether firms strategically alter reporting behaviors under a heightened regulatory environment and how the investors perceive such changes.

While numerous papers look at the quality of accounting information after SOX, we focus on changes in the disclosure behavior of qualitative information. Tetlock (2007) is one of the most cited papers to use linguistic tools to quantify qualitative textual information and finds evidence that the pessimism in Wall Street Journal's daily news column exerts temporary but significant downward pressure on the market indices. Similarly, Dougal *et al.* (2012) establish that journalists associated with a more pessimistic column tone will lead to more negative market returns the next day. Several recent studies also document how firms may implicitly reveal important information through a subtle change in disclosure style. For example, Li (2008) is the first to study the relation

[7] Some studies (e.g., Kasznik and Lev, 1995; Skinner, 1997) document a positive correlation between litigation risk and voluntary disclosure but many others (e.g., Francis *et al*., 1994; Johnson *et al.*, 2001; Baginski *et al*., 2002; Houston *et al*., 2019) find otherwise.

[8] In general, there are numerous studies that evaluate the costs, benefits, and efficiency of SOX. However, there exists no consensus on the costs and benefits of SOX. A review paper by Coates and Srinivasan (2014) summarizes the findings of more than 120 papers that study the impact of SOX.

between 10-K readability and firm performance with a meaningfully large sample. He measures the readability using the Fog Index and concludes that firms with lower reported earnings produce reports that are more difficult to read. A more recent study by Loughran and McDonald (2014) constructed a new readability measure that outperforms the Fog Index for the business documents.[9] Other papers also explore the link between document readability and many important performance measures such as earnings, cash flows, earnings forecasts, and stock returns (e.g., Biddle *et al*., 2009; Lehavy, *et al*., 2011; Lawrence, 2013; Guay *et al*., 2016).

While readability focuses on the ability of receivers to understand the intended message, we adopt the "bag-of-words" approach by tracking the amount of positive and negative words used in corporate filings. For example, Feldman *et al.* (2009) measure tone changes of managers using pre-determined word lists that classify words into positive and negative categories. Similarly, Loughran, McDonald, and Yun (2009) target the frequency of the word "ethic" and its variants along with "corporate responsibility," "social responsibility," and "socially responsible" in 10-K filings to identify sin stocks and find the use of such words being related to low corporate governance and a higher probability of class action lawsuits. Price *et al*. (2012) measure the tone for quarterly earnings conference calls using Henry's (2008) word list and find that firms with a positive tone in the Q&A session experience significantly higher stock returns. This above approach allows us to measure managerial sentiments systematically in otherwise heterogeneous textual documents. While the aforementioned studies emphasize the important relation between

[9] Loughran and McDonald (2014) propose the use of the natural log of gross 10-K file size as a better proxy for readability and show that firms with larger 10-K file sizes are related to larger subsequent return volatility, analyst dispersion, and absolute earnings surprises. We include in this paper as a control for both readability and complexity of the business.

information and qualitative disclosure, our paper complements existing literature by examining how this relation changes under a heightened regulatory environment.

Our paper is related to Feldman *et al*. (2010) who investigate the market reactions after mandatory disclosures. They measure tones from the MD&A section of Forms 10-Q and 10-K and find that positive changes in tone are associated with immediate market returns after controlling for both earnings surprises and accruals. Another paper by Li (2010b) also studies the information content of corporate filings. He uses Naïve Bayesian machine learning algorithm to capture the disclosure tones and studies the determinants of such tones and demonstrates that firms with better current performance, lower accruals, lower market-to-book ratio, and lower return volatility inclined to produce more positive forward-looking statements in the MD&As section of corporate filings. Our paper complements the above studies but differs in several aspects. We explore how imposing legal accountability on top executives affect the disclosure of qualitative information, which is not the primary focus of Feldman *et al*. (2010) and Li (2010b).[10] Our results suggest that the top executives influence both the quality and the practice of a firm's disclosure under stricter regulatory scrutiny. While Li (2010b) does not yield conclusive evidence on the impact of SOX, we document some evidence of structural changes in the way firms report after SOX. Nonetheless, we aim to provide useful insights for the policymakers by assessing the potential consequences in firms' reporting practices after major regulatory changes.

[10] Firms' top personnel must certify financial information and be held legally liable should there be any misrepresentation of information disclosed to the public after SOX.

3. Data and Methodology

3.1 Sample Construction

The final sample used in the paper is constructed using a variety of sources. We obtained stock prices from the Center for Research in Security Prices (CRSP) and quarterly balance sheet variables from Compustat. We download all variants of 10-K and 10-Q filings from the SEC's Electronic Data Gathering, Analysis, and Retrieval (EDGAR) website from 1994 to 2017.[11] Following Loughran and McDonald (2011), we focus on the main text of the filing documents and remove irrelevant tags, spaces, tables, and exhibits.[12] The measure for sentiments is based on Loughran and McDonald's (2011) Master Dictionary.[13] After parsing all the 10-Q and 10-K filings, we combined the resulting tone variables with the Compustat data and excluded observations with a quarterly closing price of less than $5 or a book total asset value of less than $1 million. The final sample after the first stage of merging results in 391,881 firm-year observations. Sample size further declines to 253,772 after combining required stock price information from CRSP to calculate cumulative abnormal returns around filing date.

Prior studies have not fully arrived at a consensus on the word lists to measure textual sentiments. Four different word lists have been extensively used today: Harvard's GI, Diction, Henry's (2008), and Loughran and McDonald (2011). Henry's (2008) word list has an advantage over Harvard's GI and Diction because her word lists were created with a business context through the examination of earnings press releases for the telecommunications and computer services

[11] Bill MacDonald has created a software depository website for textual analysis related data and codes. For a more detailed description on parsing files from EDGAR, please refer to https://sraf.nd.edu/

[12] For more details please refer to https://sraf.nd.edu/data/stage-one-10-x-parse-data/

[13] The LM (2011) master dictionary has been updated on the June of 2017 and it can be found on https://sraf.nd.edu/textual-analysis/resources/#LM_10X_Summaries

industries. The LM word lists in Loughran and McDonald (2011) follow the spirit of Henry's which was created especially for use in a business context, but they include a much more comprehensive collection of tone related words. Motivated by Feldman *et al*. (2010), Loughran *et al*. (2016), and Cohen *et al*. (2020), we mainly explore the changes in filing tones. Looking at the changes in filing tone allows us to capture discernable differences in the expectation of managers over prior quarters as they tend to compare performance changes in the current filings. Our results pertain to the entire filings while many other studies look at Management Discussion and Analysis (MD&A).[14] Limiting to just MD&A sections would leave out substantial information related to the true competitiveness of the firm.

We subtract the tone with the same quarter a year ago which allows us to pool data from both 10-Q and 10-K forms. Since there exists a systematic difference in the required format as well as the way firm report, we included a control dummy indicating 10-K form in our main regression specifications.[15] There are several advantages associated with using changes in tone instead of the levels. First, there is a high autocorrelation in the levels of filing tone when compared to previous quarters since the management tends to compose a report based on the previous ones, leading to less variations.[16] Whereas using tone changes allows us to capture incremental information over the previous period. Second, although the LM word lists are specifically fine-tuned to a business context, the bag of words approach is certainly far from flawless to capture tones information across different industries due to heterogeneous industry-specific reporting styles. Focusing on tone changes will alleviate such concern since the use of tone words are usually stable over time

[14] Cohen *et al.* (2020) confirm that announcement effects associated with changes to sub-sections of filings are not statistically different. MD&As are certainly important to investors but there is no reason to solely restrict the focus to MD&As when market investors will likely read through other sections of a corporate filing as well.
[15] This dummy also acts as a partial control for the 4th quarter effect.
[16] See "Lazy Prices" by Cohen *et al.* (2020)

and differencing the tones eventually yield a heightened proxy for managerial sentiments. From our results, it seems that tones changes are more robust in capturing the variations of stock market returns around filing dates and are more consistent with the conclusions from the prior literature.

3.2 Research Designs

This paper examines changes in the disclosure behavior of firms when top executives face legal accountability. We treat the enactment of SOX in 2002 as an event that imposes such personal legal accountability as it is potentially less endogenous than the conventional way of defining firms with increasing litigation risks (e.g. firms in an increasingly litigious industry).[17] There are two key components of interest under this research question: 1) Will firms change reporting behaviors under a major change in legal accountability; 2) how will investors perceive it? The very first question to ask is whether filing tones contain additional material information that will affect stock price behavior. Information has always been an important aspect of research on disclosure, therefore we examine the information content of tone changes using the root mean square error of a market model regression (RMSE) as a proxy for idiosyncratic stock return volatility.[18] We also examine if such relationship changes after SOX. More specifically:

$$RMSE_i = \alpha + \beta_1 dTONE_{i,t} + \beta_2 POSTSOX + \beta_3 dTONE_{i,t} * POSTSOX + Controls + FEs + e_{i,t} \quad (1)$$

If there is material information when filing tone changes, we expect tone variables to show statistical significance in explaining the post-filing idiosyncratic volatility. Kothari *et al*. (2009)

[17] It affects all public firms in the US and its enactment is less associated with individual firm characteristics.

[18] This measure of return volatility takes into account market return volatility and is highly correlated with the simple standard deviation of raw stock return.

find that a positive disclosure by the firm leads to lower return volatility, cost of capital, and analysts, forecast error dispersions and the reverse is true for negative disclosures. Krishnaswami and Subramaniam (1999) in their paper use higher idiosyncratic return volatility as a proxy for less informative stock prices.

Following Li (2010b), we use future earnings as a way to examine the information content of tone variables. If there is any information about future profitability, we should expect a significant predictive relationship between the subsequent quarters' earnings and the tone variables. The empirical strategy is as follow:

$$EARN_{i,t+n} = \alpha + \beta_1 dTONE_{i,t} + \beta_2 POSTSOX + \beta_3 dTONE_{i,t} * POSTSOX$$
$$+ EARN_{i,t} + FEs + e_{i,t} \qquad where\ n = 1,2,3,4 \qquad (2)$$

dTone is a generic expression for *dNettone*, *dNegative*, and *dPositive*. To test if there are any systematic changes in the information content, we interact tone changes measures with *POSTSOX* dummy. Should there be any systematic difference, we expect to see statistical significance on the coefficients of the interaction terms. All the above specifications include fiscal year, industry, and firm fixed effects. Portfolio construction is another way to explore the information content of the changes in tone. The finding of significant alpha from forming a zero-cost portfolio will be consistent with the story that tone changes contain information.

Prior literature has examined managerial tones and their predictive correlations with performance measures and note the importance of accruals and earnings surprises in the return behavior of stock prices.[19] Our paper takes a step further to examine how such relations may change after the enactment of SOX. Our main determinants of filing tones consist of common proxies for firm profitability, performance, report readability, leverage, size, and other accounting fundamentals. The specific empirical setup is as follow:

$$\begin{aligned} NETTONE_{i,t} = \alpha + \beta_1 EARN_{i,t} + \beta_2 SIZE_{i,t} + \beta_3 SUE_{i,t} + \beta_4 ACCRUALS_{i,t} + \beta_5 TOBINSQ_{i,t} \\ + \beta_6 EARNVOL_{i,t} + \beta_7 DERATIO_{i,t} + \beta_8 FILESIZE_{i,t} + \beta_9 RET_{i,t} \\ + \beta_{10} POSTSOX + INTERACTIONS + FEs + e_{i,t} \qquad (3) \end{aligned}$$

The dependent variables include both *Nettone* and *dNettone* which are our main proxies for the managerial sentiment on firm performance in corporate filings. *EARN* is the quarterly earnings normalized by the book assets. *SIZE* is measured by the natural log of market value. *SUE* is the Compustat-based standardized unexpected earnings as in Livnat and Mendenhall (2006) to preserve sample observations. *ACCRUALS* is the quarterly earnings subtract cashflow from operations normalized by the book value of total assets. *TOBINSQ* is measured as the total market value plus the book value of debt normalized by the book value of total assets. *EARNVOL* is the earnings volatility of the past 5 years. *DERATIO* is the total value of liabilities divided by the market value of common equity. *FILESIZE* is the log of filing size in kilobytes. *RET* is the cumulative past two-week raw returns before filing date. *INTERACTIONS* denotes all the interaction terms of individual determinant with the *POSTSOX* dummy. The above regression also includes fiscal year, three-digits SIC, and firm fixed effects.

[19] See Sloan (1996), Collins and Hribar (2000), Livnat and Santicchia (2006), Feldman *et al*. (2010), Li (2010b), Miller (2010), Battalio *et al*. (2012), Huang *et al*. (2014).

Looking from the investor perspective about information, the traditional accounting notion of value relevance comes into play. Value relevance relates to the ability of financial statement information to capture the firm value and whether information has been reflected in the stock price. Instead of gauging financial statement information, we hope to capture both the intended and unintended messages from the management through textual analysis. The idea is identical that material information released through corporate disclosure should be reflected in the stock price through investors' trading activities. Event study is an appropriate method to examine how investors perceive tone changes by the management and whether the investors perceive differently after SOX was in place. We use a standard event window of (-1,+1) to calculate *CAR* and to test for the information pricing process of tone changes around the immediate filing date. *CAR* is based on the market adjusted model and our result is largely consistent with the market model as well.[20]

While prior literature has focused on the correlation of tone levels with future accounting measures, our focus is on how investors perceive changes in tone after a major regulatory change that increases the legal accountability of top executives. More specifically, we conduct an event study to examine how changes in filing tone are reflected in the abnormal returns before and after the enactment of SOX. The baseline specification for event study is as follow:

$$CAR_{i,t} = \alpha + \beta_1 dTONE_{i,t} + \beta_2 POSTSOX + \beta_3 dTONE_{i,t} * POSTSOX + Controls + FEs + e_{i,t} \quad (4)$$

dTONE, again, can be *dNettone*, *dNegative*, or *dPositive* and our key interest is the coefficient, β_3, of the interaction terms. The changes in tones are measured by differencing contemporaneous filing

[20] Abnormal returns are defined as returns in excess of CRSP value-weighted market return. It implicitly assumes that a market beta of one. Although not shown in this paper, our results are largely consistent when calculating *CAR* using the market model instead.

tone with the tone from four quarters ago. Management should be aware of important changes that occurred during the current quarter from the previous. They will discuss them in the current and then leads to observable tone changes from the previous quarters. If we expect a systematic change in the way investors perceive information disclosure, β_3 should yield statistical significance. Should the implementation of SOX lead to more precise information disclosure from Forms 10K and 10Q, we expect a greater stock market reaction given the same magnitude of tone changes. Therefore, a positive β_3 for *dNettone* and *dPositive or* a negative β_3 for *dNegative* will be consistent with the explanation of better information quality per unit observed tone change. We included several important variables that are widely thought to affect filing returns as well as firm, three-digits SIC, and fiscal year fixed effects. To address the concern of our results driven by the scaling effect that the investors adjust their perception to offset a systematic shift managerial reporting style to the same extent, we standardized tone changes with the mean of the previous four quarters and divide by the standard deviation of the same period.[21] To further ensure the robustness of our results, we also conduct placebo tests of using pseudo dates for the implementation of SOX in the pre-SOX and post-SOX sub-periods.

[Figure 1 Here]

4. Results and Discussions

4.1 Descriptive Statistics

Figure 1a shows the mean level of net tone from 1994 to 2017. Net tone comprises both negative and positive components and the firms seem to be increasingly conservative in the overall choice

[21] The results are largely unchanged if we use eight previous quarters instead of four.

of words up to the enactment of SOX in 2002. The level of mean net tone appears to remain relatively stable after SOX. Figure 1b delves into the negative component of net tone and reveals an increasing trend in the use of negative words in the filings before SOX, but the trend plateaus out with relatively little variations after SOX. A plausible explanation is that the firm is learning the way to discuss imminent challenges as shown by the increasing use of negative words before SOX to avoid potential lawsuits.[22] Nonetheless, the unexpectedly stable pattern of negative tone level after SOX is intriguing. Figure 1c shows a similar trend for the use of positive words in filings before SOX. However, the trend seems to fluctuate to some extent with some variations after SOX. Overall, Figure 1b seems to be more suggestive of a structural change around SOX than Figure 1c and the pattern of net tone appears to be predominantly determined by the use of negative words.

[Figure 2 Here]

To further explorer the distribution of filing tones before and after SOX, we plot the histogram of tones distributions which are winsorized at 1%. Figure 2 suggests discernable changes to the tone distribution in the two sample periods. Overall, firms seem to become more conservative in filings by downplaying the level of positiveness (i.e., a leftward shift in net tone distribution and a rightward shift in negative tone distribution) in the post-SOX period. Interestingly, the use of positive words by firms exhibits a slight rightward shift after SOX, partially resembling the changing pattern of negative tone distribution. This observation seems to correspond to the argument that managers may use positive words to negate some of the negative expressions used in the filings, hence the positive correlation between the two. We also plot the distributions of the changes for our three tone measures and unanimously observe a lower dispersion for all after

[22] Firms were only required to file electronically in 1996. Many large firms begin to file electronically in 1994 and the number of total filings surged in 1996 when it became a mandatory requirement by the SEC.

SOX.[23] This is suggestive of a systematic change in the willingness of firms to disclose negative or positive news after SOX.

[Table 1 Here]

While Figure 1 reveals the time-series pattern of filing tones, Table 1 provides a detailed sample distribution of the tone, accounting, and return variables. On average, the use of negative words in a corporate filing (1.4704%) is of several magnitudes higher than that of positive words (0.5765%) and the mean of net tone is negative by construction. The changes in tone, however, has a much smaller mean. Intuitively, setting aside the economic cycle, idiosyncratic bad news for a firm will mean good news for its competitors within the same industry. The variables listed mostly follow reasonable distributions. While an average firm has a negative quarterly earning, a median firm does not. It is also interesting to note that the mean *CAR*(-1,+1) is negative which corresponds to an average negative change in the net tone. Table 2 reports results for the *t*-test of difference in means on each of the variables listed in Table 1. Almost all the variables show a statistically significant difference before and after SOX. The levels of both positive and negative tone seem to be higher after SOX, with a greater proportionate increase for the latter. Consistent with the picture depicted by Figure 1 and 2, the magnitude of changes for positive, negative, and net tones become smaller after SOX. The average file size of corporate filings, however, is growing over time. Interestingly, on average there is an increase in the abnormal return for various event windows whereas the firm's idiosyncratic stock volatility (i.e., RMSE) seems to decrease after SOX.

[Table 2 Here]

[23] Please see appendix Figure 1 for the scatter plots of absolute changes in tones, which are also suggestive of the increasing conservatism in disclosure from the managers.

There is a 23.67% correlation between the positive and negative tones which is consistent with the prior literature that firms often use positive words to negate negative words in the filings.[24] As expected, an increase in the negative tone (i.e., increasing *dNegative*) over the same quarter from the previous year is associated with the negative *CAR* around filing date and an increase in the net tone (i.e., increasing *dNettone*) with the positive. Negative tone and negative tone change are negatively correlated with contemporaneous earnings whereas positive changes in net tone are associated with positive earnings.

[Table 3 Here]

4.2 Information Content of 10-Q and 10-K filings

Table 3 reports the regression outputs of idiosyncratic return volatility on tone changes. The ability to explain significantly post-filing return volatility is itself indicative of the information content. We interact the *POSTSOX* dummy with the change in tones to examine potential effects from the enactment of SOX. Our results reveal that a positive change in net tone increases idiosyncratic return volatility, but such an effect becomes drastically more negative in the post-SOX period. Decomposing *dNettone*, we unsurprisingly observe the same for *dPositive* and the reverse for *dNegative*. Kothari *et al.* (2009) argue that greater uncertainty of cash flows from the firm leads to greater return volatility and that favorable disclosures will inform the market about higher than expected cash flow and hence reduce the uncertainty in the market. In line with their conclusion, we argue that investors under a stricter regulatory environment are more likely to perceive a positive change in net tone as a favorable disclosure that reduces uncertainty, hence the lower idiosyncratic return volatility in the post-SOX period. In other words, investors potentially deem

[24] A complete correlation table is provided in Appendix 2.

positive news to be more informative and reliable and contain less noise about the fundamental value of a firm. Whereas the disclosure of negative news seems to increase uncertainty and leads to higher post-filing return volatility after SOX. Nevertheless, evidence from Table 3 suggests that changes in tones contain substantial information that can explain the post-filing return volatility of firms.

[Table 4 Here]

Following a similar set up in Li (2010b), we empirically examine if changes in tone have any predicting power in the firm's prospective earnings and whether the underlying information has more predicting power after SOX. The independent variables are earnings in the subsequent four quarters and the main explanatory variables are the changes in tones. we have also included contemporaneous earnings as a control for earning persistence. The results are shown in Table 4. It is consistent with prior literature that filing tones have a significant correlation with the future performance of the firms even after controlling for current period earnings. We find some evidence that changes in net tone may have a positive predicting power in leading period earnings and the relationship is negative for negative changes in tone. Due to the underlying nature of positive words in a filing, it is perhaps not so surprising that positive tone changes do not yield any significant result. Following our previous specifications, we interact tone changes with the post-SOX dummy. Empirically it suggests that there is an increase in the information content of changes in negative and net tone after SOX. Changes in negative tone and net tone have a stronger association with earnings in the immediate quarter in the post-SOX period and the effect weakens in the following quarters. It interesting to note that the predictive power of changes in tone comes back for the same quarter in the following fiscal year, possibly because of the fiscal quarter cyclicity.

4.3 Determinants of Filing Tone

Investors are actively seeking information from corporate disclosures while making investment decisions. As for the managers, the goal of making disclosures is to inform and discuss the most important aspects of firms' actual performance. Using our model (3), we explore how contemporary firm performance might have a different degree of correlation with filing tones. The results are shown in Table 5.

We find that contemporary firm earning is positively and significantly correlated with both the level of net tone and the changes in net tone over prior periods.[25] Interestingly, the negative coefficient of the interaction of *EARN* and *POSTSOX* suggests a moderating effect on the direct association of earnings and tone changes. Coefficients of *SUE* and its interaction also suggest a similar trend in the post-SOX period. One possible explanation is that the firm during the post-SOX period is trying to moderate the tones of reporting and try to smooth tone across periods to prevent unwanted lawsuits. Managers possibly are less prone to the over-use of optimistic words or pessimistic words when the liability for misreporting increases after SOX.

[Table 5 Here]

Accruals is another important factor that reveals the underlying financial condition of a firm. Sloan (1996) documents that managers may manipulate the accrual components of earnings to meet short-term objectives and investors underreact to such manipulation. Our results also suggest that managers may indeed know the significance of accruals for future earnings as their tones seem to be significantly correlated with the accrual levels. Consistent with prior literature, we find that

[25] Note that the level of net tone can be relatively persistent within firm over time but changes in net tone are not and contain incremental information over the past period.

accruals are perceived negatively and are negatively correlated with *Nettone* or *dNettone* in the corporate filings but managers do not seem to systematically change the way of reporting based on accruals after SOX.

TOBINSQ measures market valuation with respect to its replacement value may indirectly indicate investor's confidence in the growth potential or future performance of the firm and hence the positive sign. However, we do not observe a significant change in the relation after SOX. *FILESIZE* not only is a control for firm-specific complexity since a more complicated business model usually results in a longer filing, but it is also a good proxy for document readability. Loughran and McDonald (2014) compares the FOG index with the log of file size and conclude that the latter is a better proxy for document readability. It seems that a more obscure document is associated with a less positive net tone and leads to less positive change in net tone from the prior period. While past earning volatility, Tobin's *q*, and filing file size seem to influence the absolute level of net tone differently in the post-SOX period, none of these seems to affect changes in net tone differently after SOX.

[Table 6 Here]

4.4 The Mapping of Rankings from Filing Tones

Another way to examine whether filing tones contain information about expected firm performance is via portfolio strategy. The mapping of filing tones to firm performance potentially may also have changed after SOX. Specifically, we sort portfolio into five equal quintiles based on *dNettone* and *dNegative*.[26] Portfolios are equally weighted and rebalanced on the first day of

[26] The results for positive tone changes are not shown as the strategy does not yield significant excess returns.

each month. The variable for sorting is the amount of tone changes from the most recent filings within the past 90 calendar days of each rebalancing. The results for changes in net tone and negative tone are shown in Table 6. The mean changes for net (negative) tone are -0.2389 (-0.5813) for quintile 1 and 0.2249 (0.6071) for quintile 5. A zero-cost strategy of buying the best quintile and shorting the worst quintile yield significant alphas for both tone changes.[27] We also examine the mean abnormal returns during the (-10,+10) filing window for the best quintile and the worst quintile.[28] The results are exhibited in Figure 3. A preliminary look at the figure reassures that a strategy of buying best and shorting the worst yield on average positive daily abnormal returns. Interestingly, from untabulated results, we do not find any noticeable difference between the pre-SOX and post-SOX periods for both the zero-cost strategy and the mean abnormal returns. Arguably, filing tones in both periods are equally well representative of actual firm performance relative to peer firms that those with the most positive (negative) filing tone remain the best (worst) performing ones. In other words, no additional information seems to be revealed with regard to relative expected performance after SOX.

[Figure 3 Here]

4.5 Reaction from the Capital Market

The goal of disclosure is ultimately to keep stakeholders informed of the firm's operations and financial condition. While some studies have look at the value relevance of tone changes in the

[27] In untabulated tables, including Fama-French 3 or 5 factors yield similar results.
[28] Quintiles are sorted based on tone changes of each fiscal quarters and we do not find significantly different patterns before and after SOX.

subsequent stock prices, we examine how the enactment of SOX may affect the reaction from the capital market.

Table 7 shows the results of the event study with filing date (i.e., $t = 0$) as the event date. Following the standard convention of short-window *CAR*, we set a 3-day event window (i.e., t-1 to t+1). The key variable of interest is the interaction term of tone changes and *POSTSOX* dummy, which is an indication of how much information being reflected in the stock price. Column (1) to (3) report the regressions of *CAR*(-1,+1) on changes in *Nettone*, standardized changes in *Nettone*, and standardized *Nettone* without firm-level controls but Column (4) to (6) do. All specifications include firm, industry, and year fixed effects with standard errors clustered at both firm and fiscal year level. The coefficients of *dNettone* are positively correlated with *CAR* the relationship is stronger in the post-SOX period as shown by the positive sign of the interaction terms. It is consistent with the argument that investors receive more information per unit tone change in the post-SOX period.[29]

[Table 7 Here]

An alternative explanation that can also rationalize the results is that investors adjust accordingly to the systematic change in the way managers report in 10-K or 10-Q. In this case, investors do not receive additional information per unit of tone changes but scale the informational content accordingly back to the level before SOX. As noted in Tetlock *et al*. (2008), if the tone variables calculated from 10-Q or 10-K filings are non-stationary when there exist regime changes in the distribution of sentimental words, standardization of those variables may be required. We

[29] Not shown in the paper, we also run identical tests for abnormal cumulative return of longer event windows (i.e., *CAR(-1,9)* and *CAR(-1,29)*). The results are less statistically significant or not significant at all. The investors seem to take into account the incremental information per unit of tone change very quickly after SOX within a short event window.

standardize *dNettone* and *Nettone* by subtracting their mean over the prior 4 fiscal quarters and divide by the standard deviations of the same period.[30] The interaction terms for both standardized tone measures are statistically significant and positive which potentially rules out the scaling effect explanation. We standardize *Nettone* as a robustness check and the results are largely consistent with our argument that per unit information content of corporate filings seems to increase after SOX.[31] If there is no significant systematic change affecting the regulatory or information environment after 2002, we argue that investors indeed receive more information per unit tone change in the post-SOX period.

As aforementioned, accruals and earnings surprises are two very important variables that could impact the stock returns. Consistent with prior studies, accruals appear to be negatively associated with abnormal returns and *SUE* positive. Abnormal returns in the past trading days are also included to account for the momentum effect. Earnings unsurprisingly yield a strong positive correlation with *CAR*. The coefficients of *POSTSOX* suggest that *CAR* around filing date on average is smaller in the post-SOX period. You and Zhang (2009) show that longer annual reports lead to a delayed market reaction from the investors up to a year and that 10-Ks with higher word counts reduces investor's ability to information quickly into the stock prices. Their findings bias against our results since the file size on average is bigger after SOX and we should expect market reaction to be smaller keeping all else the same, but we see heightened abnormal returns given per unit of tone change.

[30] We also repeat robustness checks by standardizing with the corresponding mean and standard deviation of the past 8 quarters. The results remain consistent.

[31] Standardizing *Nettone* resembles the logic of calculating *dNettone* in which we account for prior tone levels.

[Table 8 Here]

Table 8 reports similar results as in Table 7 for various forms of negative and positive tone measures with the same firm-level controls and fixed effects. *dNegative* shows an expected negative predictive relationship with the *CAR* whereas *dPositive* unsurprisingly yields no statistical significance at all. The interaction of *dNegative* with *POSTSOX* suggests some evidence of stronger investor reactions after SOX per unit of negative tone changes but not for positive changes. After standardizing *dNegative* and *dPositive*, we observe stronger results for the negative component of tone changes after SOX, but again, not for the positive. The results for standardized negative tone level (i.e., *Negative_S*) remain consistent with that of the other two measures. Inevitably, the results for positive tone measures are not significantly meaningful in explaining *CAR* around the filing dates in general. Investors seem to relatively decipher better the changes in the net tone and negative tone as compared to the positive. *dPositive*, irrespective of standardization, seems to be a relatively noisy proxy for positivity in managerial sentiment, at least from the perspective of the investors. However, we cannot explicitly reject the explanation that investors find positive disclosure from the managers uninformative. The interpretation here is limited by the fact that it is a joint test of the positive word list accurately capturing the positive sentimental changes of the managers and investors perceiving the use of positive words as a valid signal to be factored into the stock prices.

4.6 Placebo Tests

While Figure 1 exhibits trend discontinuity around the enactment of SOX, one might reasonably suspect that the observed change in the stock market reactions in the post-SOX period might not be a direct outcome of the legislation change. We further conduct two placebo tests in both the pre- and post-SOX subperiods assuming there is a pseudo-event occurring on the 31st of July in 1998 and 2010 respectively.[32] Results in Table 9 represent the same regression set-ups as in Tables 7 and 8 except that we redefine accordingly the post-event dummies, *POST(1998)* and *POST(2010),* which indicate the periods after 1998 and 2010 respectively. There is also no evidence of obvious structural change in the information content of tone measures in the two periods before and after SOX.

[Table 9 Here]

5. Conclusion

Governance and financial crises are often followed by regulation changes targeting at enhancing transparency through enforcing disclosure requirements or imposing penalties for misreporting. One important objective of SOX is to enhance transparency, but the effect of making top personnel accountable for misreporting is a priori ambiguous. This paper, therefore, explores the potential impact on disclosure behavior when firms' top executives become legally liable for the misrepresentation of information. We first document that a structural break exists in the distribution of filing tones around SOX and that firms appear to report more negatively and with less variations over time after SOX. We also find that changes in filing tones contain substantial information that is reflected promptly in the capital market. The investors exhibit a stronger reaction to per unit change of filing tones after SOX which are not driven entirely by the systematic changes in tone distribution after SOX. Our result also reveals a weaker correlation between filing tones and firms' contemporary quarter performance measures after the enactment of SOX. In essence, our evidence supports the view that holding top executives liable for misreporting

[32] 1998 and 2010 are approximately the mid-points of the two subperiods respectively.

potentially changes the way firms communicate to the capital market. Firms appear to adopt a more conservative strategy in the disclosure of soft information, which leads to a stronger reaction from the investors.

References

Amir, E., & Lev, B. (1996). Value-relevance of nonfinancial information: The wireless communications industry. *Journal of Accounting and Economics, 22*(1-3), 3-30.

Arping, S., & Sautner, Z. (2013). Did SOX Section 404 Make Firms Less Opaque? Evidence from Cross-Listed Firms. *Contemporary Accounting Research, 30*(3), 1133-1165.

Baginski, S. P., Hassell, J. M., & Kimbrough, M. D. (2002). The effect of legal environment on voluntary disclosure: Evidence from management earnings forecasts issued in US and Canadian markets. *The Accounting Review, 77*(1), 25-50.

Bamber, L. S., Jiang, J., & Wang, I. Y. (2010). What's my style? The influence of top managers on voluntary corporate financial disclosure. The accounting review, 85(4), 1131-1162.

Battalio, R. H., Lerman, A., Livnat, J., & Mendenhall, R. R. (2012). Who, if anyone, reacts to accrual information? *Journal of Accounting and Economics*, *53*(1-2), 205-224.

Biddle, G. C., Hilary, G., & Verdi, R. S. (2009). How does financial reporting quality relate to investment efficiency? *Journal of Accounting and Economics*, *48*(2-3), 112-131.

Brockman, P., & Yan, X. S. (2009). Block ownership and firm-specific information. *Journal of Banking & Finance*, *33*(2), 308-316.

Brudney, V. (1989). A note on materiality and soft information under the federal securities laws. *Virginia Law Review*, 723-760.

Coates, J. C., & Srinivasan, S. (2014). SOX after ten years: A multidisciplinary review. *Accounting Horizons*, *28*(3), 627-671.

Cohen, L., Malloy, C., & Nguyen, Q. (2020). *Lazy prices*. *The Journal of Finance*, *75*(3), 1371-1415.

Collins, D. W., & Hribar, P. (2000). Earnings-based and accrual-based market anomalies: one effect or two? *Journal of Accounting and Economics, 29*(1), 101-123.

Dougal, C., Engelberg, J., Garcia, D., & Parsons, C. A. (2012). Journalists and the stock market. *The Review of Financial Studies*, *25*(3), 639-679.

Feldman, R., Govindaraj, S., Livnat, J., & Segal, B. (2010). Management's tone change, post earnings announcement drift and accruals. *Review of Accounting Studies*, *15*(4), 915-953.

Francis, J., Philbrick, D., & Schipper, K. (1994). Shareholder litigation and corporate disclosures. *Journal of Accounting Research, 32*(2), 137-164.

Guay, W., Samuels, D., & Taylor, D. (2016). Guiding through the fog: Financial statement complexity and voluntary disclosure. *Journal of Accounting and Economics, 62*(2-3), 234-269.

Henry, E. (2008). Are investors influenced by how earnings press releases are written? *The Journal of Business Communication (1973)*, *45*(4), 363-407.

Hoffman, D. A. (2005). The best puffery article ever. *Iowa L. Rev.*, *91*, 1395.

Houston, J. F., Lin, C., Liu, S., & Wei, L. (2019). Litigation Risk and Voluntary Disclosure: Evidence from Legal Changes. *The Accounting Review*. doi:10.2308/accr-52355

Huang, A. H., Zang, A. Y., & Zheng, R. (2014). Evidence on the information content of text in analyst reports. *The Accounting Review*, *89*(6), 2151-2180.

Huang, P. H. (2005). Moody investing and the Supreme Court: Rethinking the materiality of information and the reasonableness of investors. *Sup. Ct. Econ. Rev.*, *13*, 99.

Huang, X., Teoh, S. H., & Zhang, Y. (2013). Tone management. *The Accounting Review*, *89*(3), 1083-1113.

Johnson, M. F., Kasznik, R., & Nelson, K. K. (2001). The impact of securities litigation reform on the disclosure of forward-looking information by high technology firms. *Journal of Accounting Research, 39*(2), 297-327.

Kasznik, R., & Lev, B. (1995). To warn or not to warn: Management disclosures in the face of an earnings surprise. *Accounting Review*, 113-134.

Kearney, C., & Liu, S. (2014). Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis, 33*, 171-185.

Kelly, P. J. (2014). Information efficiency and firm-specific return variation. *The Quarterly Journal of Finance, 4*(04), 1450018.

Kothari, S. P., Li, X., & Short, J. E. (2009). The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: A study using content analysis. *The Accounting Review, 84*(5), 1639-1670.

Krishnaswami, S., & Subramaniam, V. (1999). Information asymmetry, valuation, and the corporate spin-off decision. *Journal of Financial Economics, 53*(1), 73-112.

Lawrence, A. (2013). Individual investors and financial disclosure. *Journal of Accounting and Economics, 56*(1), 130-147.

Lee, D. W., & Liu, M. H. (2011). Does more information in stock price lead to greater or smaller idiosyncratic return volatility? *Journal of Banking & Finance, 35*(6), 1563-1580.

Lehavy, R., Li, F., & Merkley, K. (2011). The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review, 86*(3), 1087-1115.

Lev, B., & Thiagarajan, S. R. (1993). Fundamental information analysis. *Journal of Accounting Research, 31*(2), 190-215.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics, 45*(2-3), 221-247.

Li, F. (2010a). Survey of the Literature. *Journal of accounting literature, 29*, 143-165.

Li, F. (2010b). The information content of forward-looking statements in corporate filings—A naïve Bayesian machine learning approach. *Journal of Accounting Research, 48*(5), 1049-1102.

Livnat, J., & Mendenhall, R. R. (2006). Comparing the post-earnings announcement drift for surprises calculated from analyst and time series forecasts. *Journal of Accounting Research, 44*(1), 177-205.

Livnat, J., & Santicchia, M. (2006). Cash flows, accruals, and future returns. *Financial Analysts Journal, 62*(4), 48-61.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance, 66*(1), 35-65.

Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *The Journal of Finance, 69*(4), 1643-1671.

Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research, 54*(4), 1187-1230.

Loughran, T., McDonald, B., & Yun, H. (2009). A wolf in sheep's clothing: The use of ethics-related terms in 10-K reports. *Journal of Business Ethics, 89*(1), 39-49.

Mayew, W. J., & Venkatachalam, M. (2012). The power of voice: Managerial affective states and future firm performance. *The Journal of Finance, 67*(1), 1-43.

Morck, R., Yeung, B., & Yu, W. (2000). The information content of stock markets: why do emerging markets have synchronous stock price movements? *Journal of Financial Economics, 58*(1-2), 215-260.

O'Hare, J. (1998). The resurrection of the dodo: The unfortunate re-emergence of the puffery defense in private securities fraud actions. *Ohio St. LJ, 59*, 1697.

Padfield, S. J. (2007). Is puffery material to investors-maybe we should ask them. *U. Pa. J. Bus. & Emp. L., 10*, 339.

Price, S. M., Doran, J. S., Peterson, D. R., & Bliss, B. A. (2012). Earnings conference calls and stock returns: The incremental informativeness of textual tone. *Journal of Banking & Finance, 36*(4), 992-1011.

Rice, S. C., & Weber, D. P. (2012). How effective is internal control reporting under SOX 404? Determinants of the (non-) disclosure of existing material weaknesses. *Journal of Accounting Research, 50*(3), 811-843.

Rogers, J. L., Van Buskirk, A., & Zechman, S. L. (2011). Disclosure tone and shareholder litigation. *The Accounting Review, 86*(6), 2155-2183.

ROLL, R. (1988). R2. *The Journal of Finance, 43*(3), 541-566.

Roussel, R. G. (1998). Securities fraud or mere puffery: Refinement of the corporate puffery defense. *Vand. L. Rev., 51*, 1049.

Skinner, D. J. (1997). Earnings disclosures and stockholder lawsuits. *Journal of Accounting and Economics, 23*(3), 249-282.

Shiller, R. J. (1981). Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *The American Economic Review, 71*(3), 421-436.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Review*, 289-315.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance, 62*(3), 1139-1168.

Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance, 63*(3), 1437-1467.

West, K. D. (1988). Dividend innovations and stock price volatility. *Econometrica, 56*(1), 37-61.

You, H., & Zhang, X.-j. (2009). Financial reporting complexity and investor underreaction to 10-K information. *Review of Accounting Studies, 14*(4), 559-586.

Table 1. Summary Statistics

The table reports the summary statistics of the main variables used in this paper, including the number of observations, *N*, mean, stand deviations, and the value of the variable at various percentile. Data for tone variables were obtained from using Loughran and McDonald (2011). Control variables were calculated from Compustat and CRSP. For a detailed definition of variables, please refer to Appendix 1.

	N	Mean	S.D.	Min	.25	Mdn	.75	Max
Tone Variables								
Positive	391881	0.5765	0.2096	0.1273	0.4354	0.5541	0.6937	1.2499
Negative	391881	1.4704	0.5959	0.2896	1.0469	1.4173	1.8359	3.2300
Nettone	391689	-0.4086	0.2048	-0.8182	-0.5481	-0.4397	-0.3048	0.2708
dPositive	339575	0.0044	0.1561	-0.4722	-0.0753	0.0029	0.0832	0.4893
dNegative	339575	0.0278	0.4441	-1.3126	-0.1796	0.0150	0.2316	1.4543
dNettone	339318	-0.0076	0.1758	-0.5857	-0.0858	-0.0024	0.0774	0.5109
Positive_S	366844	0.0831	1.7047	-5.3235	-0.7769	-0.0905	0.8690	7.2997
dPositive_S	317412	0.0539	1.9931	-6.5953	-0.9621	0.0328	1.0097	7.3210
Negative_S	366850	0.1999	1.9239	-5.1272	-0.7915	-0.0726	0.9246	9.3011
dNegative_S	317420	0.1043	2.0943	-6.6569	-0.9421	0.0566	1.0128	8.5298
Nettone_S	366538	-0.0062	1.7592	-6.1654	-0.8690	-0.0646	0.8494	6.5608
dNettone_S	317138	0.0188	1.9675	-6.7363	-0.9815	0.0213	1.0138	6.8541
Filesize	391881	11.6641	0.9764	9.5027	10.9711	11.6685	12.3628	13.8271
Control Variables								
EARN	387632	-0.0005	0.0838	-3.5000	0.0008	0.0065	0.0181	0.1688
SIZE	351229	6.2401	1.8482	-0.9807	4.8899	6.1488	7.4762	10.6813
SUE	308165	-0.0009	0.0816	-1.2644	-0.0058	0.0009	0.0056	1.9459
ACCRUALS	356816	-0.0272	0.1036	-1.5201	-0.0617	-0.0213	0.0011	1.3631
TOBINSQ	350779	2.2944	5.1254	0.5140	1.0590	1.3759	2.1751	133.2000
EARNVOL	214828	0.0335	0.2712	0.0003	0.0046	0.0110	0.0254	9.7931
DERATIO	350777	2.0294	3.7609	0.0047	0.1962	0.5936	1.8463	47.9809
Return Variables								
CAR(-1,+1)	253772	-0.0007	0.0598	-1.4823	-0.0235	-0.0013	0.0214	4.8358
CAR(-10,-6)	253772	0.0008	0.0640	-1.0412	-0.0264	-0.0007	0.0262	3.5295
CAR(-10,-2)	253772	0.0012	0.0834	-1.2875	-0.0358	-0.0006	0.0362	3.7662
RET(-10,-1)	253772	0.0072	0.0933	-1.2413	-0.0335	0.0058	0.0474	3.5779
RMSE(0,22)	253677	0.0226	0.0173	0.0000	0.0120	0.0182	0.0280	1.4645
RMSE(6,28)	253364	0.0218	0.0169	0.0000	0.0115	0.0173	0.0268	1.3578

Table 2. Pre-SOX versus Post-SOX

This table reports the means for pre-SOX and post-SOX periods. A *t-test* of difference in mean between the two periods was reported in the last column and coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For a detailed definition of variables, please refer to Appendix 1.

	Pre-SOX					**Post-SOX**					
	N	***Mean***	**.25**	**Mdn**	**.75**	***N***	***Mean***	**.25**	**Mdn**	**.75**	***Pre - Post***
Positive	150918	0.557	0.399	0.531	0.686	240963	0.589	0.456	0.566	0.698	-0.031***
Negative	150918	1.270	0.807	1.195	1.640	240963	1.596	1.194	1.531	1.923	-0.326***
Nettone	150740	-0.349	-0.531	-0.387	-0.200	240949	-0.446	-0.555	-0.462	-0.352	0.098***
dPositive	115111	0.008	-0.103	0.005	0.117	224464	0.003	-0.065	0.002	0.071	0.006***
dNegative	115111	0.057	-0.232	0.041	0.345	224464	0.013	-0.162	0.007	0.187	0.044***
dNettone	114889	-0.018	-0.151	-0.010	0.122	224429	-0.002	-0.065	0.000	0.064	-0.015***
Positive_S	133352	0.146	-0.803	-0.051	0.920	233492	0.047	-0.764	-0.112	0.840	0.099***
dPositive_S	99449	0.052	-0.978	0.027	1.020	217963	0.055	-0.954	0.035	1.006	-0.002
Negative_S	133358	0.332	-0.768	-0.016	1.022	233492	0.124	-0.804	-0.107	0.873	0.208***
dNegative_S	99457	0.164	-0.928	0.093	1.088	217963	0.077	-0.949	0.040	0.980	0.087***
Nettone_S	133080	-0.020	-0.929	-0.104	0.830	233458	0.002	-0.836	-0.042	0.859	-0.021***
dNettone_S	99212	-0.011	-1.027	0.004	1.002	217926	0.032	-0.963	0.030	1.019	-0.043***
FILESIZE	150918	11.148	10.330	10.907	11.928	240963	11.988	11.415	11.905	12.534	-0.840***
EARN	149017	-0.003	0.001	0.007	0.019	238615	0.001	0.001	0.006	0.018	-0.005***
SIZE	133676	5.687	4.402	5.530	6.799	217553	6.580	5.299	6.518	7.803	-0.893***
SUE	103182	-0.002	-0.007	0.001	0.005	204983	0.000	-0.005	0.001	0.006	-0.002***
ACCRUALS	124121	-0.022	-0.063	-0.021	0.012	232695	-0.030	-0.061	-0.022	-0.001	0.008***
TOBINSQ	133350	2.403	1.063	1.383	2.303	217429	2.228	1.057	1.372	2.109	0.175***
EARNVOL	29836	0.024	0.005	0.011	0.024	184992	0.035	0.005	0.011	0.026	-0.011***
DERATIO	133347	1.886	0.170	0.585	1.824	217430	2.118	0.213	0.598	1.861	-0.232***
CAR(-1,+1)	86570	-0.002	-0.026	-0.003	0.022	167202	0.000	-0.022	-0.001	0.021	-0.001***
CAR(-10,-6)	86570	0.002	-0.032	-0.001	0.032	167202	0.000	-0.024	-0.001	0.024	0.002***
CAR(-10,-2)	86570	0.004	-0.040	0.000	0.044	167202	0.000	-0.034	-0.001	0.033	0.004***
RET(-10,-1)	86570	0.011	-0.033	0.007	0.052	167202	0.005	-0.034	0.005	0.045	0.006***
RMSE(0,22)	86541	0.028	0.015	0.023	0.035	167136	0.020	0.011	0.016	0.024	0.008***
RMSE(6,28)	86401	0.028	0.015	0.024	0.035	166963	0.019	0.010	0.015	0.022	0.010***

Table 3. Idiosyncratic Return Volatility

This table reports the regression outputs of idiosyncratic return volatility on changes in tones. Firm-level controls include firm size, 10K dummy, earnings, accruals, Tobin's Q, leverage, earnings surprise, document size, and past returns. Standard errors were clustered by firms and reported in the parentheses. Fiscal year, 3-digit SIC, and firm fixed effects are included in all specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For a detailed definition of variables, please refer to Appendix 1.

	(1) RMSE(0,22)	**(2)** RMSE(0,22)	**(3)** RMSE(0,22)	**(4)** RMSE(6,28)	**(5)** RMSE(6,28)	**(6)** RMSE(6,28)
dNettone	0.0008***			0.0008***		
	(0.000)			(0.000)		
dNettone * POSTSOX	-0.0012***			-0.0014***		
	(0.000)			(0.000)		
dNegative		-0.0001			0.0000	
		(0.000)			(0.000)	
dNegative * POSTSOX		0.0004***			0.0004***	
		(0.000)			(0.000)	
dPositive			0.0012***			0.0013***
			(0.000)			(0.000)
dPositive * POSTSOX			-0.0011***			-0.0012***
			(0.000)			(0.000)
POSTSOX	-0.0009***	-0.0009***	-0.0008***	-0.0021***	-0.0020***	-0.0019***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.0265***	0.0270***	0.0266***	0.0259***	0.0264***	0.0259***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Firm-Level Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.45	0.45	0.45	0.46	0.46	0.46
Observations	235562	235724	235724	235271	235433	235433

Table 4. Predicting Future Earnings

This table reports the output for the predictive regression of future quarterly earnings on tone changes. The interaction terms of the tone variables with the *POSTSOX* dummy were also included in all specifications. Standard errors were clustered by firms and reported in the parentheses. Fiscal year, 3-digit SIC, and firm fixed effects are included in all specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For a detailed definition of variables, please refer to Appendix 1.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	EARN(t+1)	*EARN(t+2)*	*EARN(t+3)*	*EARN(t+4)*	*EARN(t+1)*	*EARN(t+2)*	*EARN(t+3)*	*EARN(t+4)*	*EARN(t+1)*	*EARN(t+2)*	*EARN(t+3)*	*EARN(t+4)*
dNettone	0.0027***	0.0024***	0.0026***	0.0008								
	(0.001)	(0.001)	(0.001)	(0.001)								
dNettone * POSTSOX	0.0045***	0.0023*	0.0016	0.0024*								
	(0.001)	(0.001)	(0.001)	(0.001)								
dNegative					-0.0021***	-0.0017***	-0.0014***	-0.0006				
					(0.000)	(0.000)	(0.000)	(0.000)				
dNegative * POSTSOX					-0.0011**	-0.0007	-0.0007	-0.0011*				
					(0.001)	(0.001)	(0.001)	(0.001)				
dPositive									-0.0012	-0.0006	0.0001	-0.0010
									(0.001)	(0.001)	(0.001)	(0.001)
dPositive * POSTSOX									0.0023	0.0003	-0.0007	-0.0002
									(0.001)	(0.002)	(0.002)	(0.002)
POSTSOX	-0.0000	0.0038***	0.0079***	0.0011	-0.0002	0.0037***	0.0079***	0.0011	-0.0003	0.0036***	0.0078***	0.0009
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
EARN	0.2087***	0.1362***	0.1210***	0.1537***	0.2084***	0.1360***	0.1208***	0.1536***	0.2092***	0.1366***	0.1214***	0.1540***
	(0.026)	(0.023)	(0.024)	(0.027)	(0.026)	(0.023)	(0.023)	(0.027)	(0.026)	(0.023)	(0.023)	(0.027)
Constant	0.0005	-0.0028***	-0.0061***	-0.0022***	0.0007	-0.0027***	-0.0060***	-0.0022***	0.0007	-0.0027***	-0.0060***	-0.0021***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.50	0.49	0.49	0.47	0.51	0.49	0.49	0.47	0.50	0.49	0.49	0.47
Observations	323030	314102	305709	298126	323268	314336	305938	298358	323268	314336	305938	298358

Table 5. Determinants of Filing Tones

This table reports the regression results of tones (tone changes) on common determinants. Column (1), and (3) report the coefficients of the determinants without interaction terms. Column (2) and (4) include the interaction terms of determinants with the POSTSOX dummy. Standard errors were clustered by firm and fiscal year and reported in the parentheses. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For a detailed definition of variables, please refer to Appendix 1

	(1)	**(2)**	**(3)**	**(4)**
	Nettone	***Nettone***	***dNettone***	***dNettone***
EARN	0.3056***	0.4993***	0.1345***	0.2456***
	(0.029)	(0.067)	(0.018)	(0.056)
SIZE	0.0200***	0.0145***	-0.0025	-0.0046*
	(0.003)	(0.004)	(0.002)	(0.003)
SUE	0.0083	0.0750***	0.0923***	0.2964***
	(0.008)	(0.017)	(0.016)	(0.046)
ACCRUALS	-0.1610***	-0.1748***	-0.0730***	-0.0851**
	(0.015)	(0.037)	(0.012)	(0.034)
TOBINSQ	0.0016**	0.0051***	0.0014***	0.0035**
	(0.001)	(0.002)	(0.000)	(0.002)
EARNVOL	-0.0161	-0.4379***	0.0141**	0.1776*
	(0.012)	(0.127)	(0.006)	(0.102)
DERATIO	-0.0017**	-0.0036**	-0.0014***	0.0006
	(0.001)	(0.001)	(0.000)	(0.001)
FILESIZE	-0.0191***	-0.0402***	-0.0296***	-0.0358***
	(0.004)	(0.006)	(0.004)	(0.011)
RET(-10,-1)	0.0154**	-0.0035	0.0203**	0.0143
	(0.006)	(0.020)	(0.008)	(0.032)
POSTSOX		-0.3458***		-0.0682
		(0.074)		(0.129)
EARN * POSTSOX		-0.2320***		-0.1299**
		(0.067)		(0.054)
SIZE * POSTSOX		0.0068**		0.0032
		(0.003)		(0.002)
SUE * POSTSOX		-0.0709***		-0.2219***
		(0.018)		(0.047)
ACCRUALS * POSTSOX		0.0219		0.0167
		(0.037)		(0.035)
TOBINSQ * POSTSOX		-0.0037*		-0.0025
		(0.002)		(0.002)
EARNVOL * POSTSOX		0.4261***		-0.1654
		(0.126)		(0.102)
DERATIO * POSTSOX		0.0020		-0.0021***
		(0.001)		(0.001)
FILESIZE * POSTSOX		0.0270***		0.0073
		(0.007)		(0.012)
RET(-10,-1) * POSTSOX		0.0216		0.0073
		(0.020)		(0.032)

Table 5 Continued

Constant	-0.3472***	-0.0785	0.3609***	0.4116***
	(0.063)	(0.063)	(0.053)	(0.119)
Industry FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes
Adjusted R^2	0.46	0.46	0.04	0.04
Observations	152350	152350	152335	152335

Table 6. Portfolio Return

This table shows the mean of change in tones and quintile returns for the top and bottom quintile portfolios. Stocks were sorted into five equal quintiles based on *dNettone* and *dNegative*. Portfolios are equal-weighted and rebalanced monthly on the first day of each month. The firms included in the sorting process must be the most recent filings within the past 90 calendar days of each rebalancing date. It also reports the returns for the zero-cost strategy of longing the best quintile and shorting the worst. Standard errors were reported int he parentheses and *** indicates significance at 1% level.

		Quintile		**Difference**
dNettone		**Q1**	**Q5**	**Q5 - Q1**
	Quintile Return	0.0095*** (0.004)	0.0127*** (0.004)	0.0032*** (0.001)
	Quintile Mean	-0.2389	0.2249	0.4638
dNegative		**Q1**	**Q5**	**Q1 - Q5**
	Quintile Return	0.0127*** (0.004)	0.0104*** (0.004)	0.0023*** (0.001)
	Quintile Mean	-0.5813	0.6071	-1.1884

Table 7. Post-SOX Value Relevance of Net Tones Changes

This table reports the results under an event study setting. Cumulative abnormal returns (CARs) were calculated based on the market adjusted model. Standard errors were clustered by firm and fiscal year and reported in the parentheses. Fiscal year, 3-digit SIC, and firm fixed effects are included in all specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For a detailed definition of variables, please refer to Appendix 1.

CAR(-1,+1)	(1)	(2)	(3)	(4)	(5)	(6)
dNettone	0.0023***			0.0015**		
	(0.001)			(0.001)		
dNettone * POSTSOX	0.0034**			0.0033*		
	(0.002)			(0.002)		
dNettone_S		0.0000			-0.0001	
		(0.000)			(0.000)	
dNettone_S * POSTSOX		0.0003***			0.0003**	
		(0.000)			(0.000)	
Nettone_S			0.0001*			0.0001
			(0.000)			(0.000)
Nettone_S * POSTSOX			0.0004***			0.0003**
			(0.000)			(0.000)
10K				0.0003	0.0007	0.0002
				(0.001)	(0.001)	(0.001)
POSTSOX				-0.0046***	-0.0049***	-0.0047***
				(0.001)	(0.001)	(0.001)
SIZE				-0.0062***	-0.0065***	-0.0062***
				(0.001)	(0.001)	(0.001)
EARN				0.0631***	0.0749***	0.0630***
				(0.014)	(0.013)	(0.014)
ACCRUALS				-0.0198***	-0.0205***	-0.0197***
				(0.004)	(0.005)	(0.004)
TOBINSQ				0.0000	0.0000	0.0000
				(0.000)	(0.000)	(0.000)
DERATIO				-0.0001	-0.0001	-0.0001
				(0.000)	(0.000)	(0.000)
SUE				0.0316***	0.0317***	0.0317***
				(0.005)	(0.005)	(0.005)
FILESIZE				0.0002	0.0000	0.0001
				(0.000)	(0.000)	(0.000)
CAR(-10,-2)				-0.0398***	-0.0392***	-0.0398***
				(0.005)	(0.005)	(0.005)
CAR(-10,-6)				0.0213***	0.0213***	0.0212***
				(0.004)	(0.005)	(0.004)
Constant	-0.0006***	-0.0006***	-0.0007***	0.0409***	0.0445***	0.0410***
	(0.000)	(0.000)	(0.000)	(0.005)	(0.005)	(0.005)
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.02	0.02	0.02	0.03	0.03	0.03
Observations	236618	224662	246405	219136	209036	219193

Table 8. Post-SOX Value Relevance of Tone Decomposition

This table reports the results for different forms of negative and positive tone measures. Cumulative abnormal returns (CARs) were calculated based on the market adjusted model. The included firm-level controls were the same as in Table 6. Standard errors were clustered by firm and fiscal year and reported in the parentheses. Fiscal year, 3-digit SIC, and firm fixed effects are included in all specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For a detailed definition of variables, please refer to Appendix 1.

CAR(-1,+1)	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
dNegative	-0.0006**					
	(0.000)					
dNegative * POSTSOX	-0.0009*					
	(0.000)					
dNegative_S		-0.0000				
		(0.000)				
dNegative_S * POSTSOX		-0.0002**				
		(0.000)				
Negative_S			0.0000			
			(0.000)			
Negative_S * POSTSOX			-0.0005***			
			(0.000)			
dPositive				0.0003		
				(0.001)		
dPositive * POSTSOX				0.0012		
				(0.002)		
dPositive_S					-0.0001	
					(0.000)	
dPositive_S * POSTSOX					0.0002	
					(0.000)	
Positive_S						-0.0000
						(0.000)
Positive_S * POSTSOX						0.0001
						(0.000)
POSTSOX	-0.0047***	-0.0049***	-0.0047***	-0.0047***	-0.0049***	-0.0047***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.0407***	0.0441***	0.0413***	0.0430***	0.0452***	0.0431***
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Firm-Level Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.02	0.02	0.02	0.03	0.03	0.03
Observations	236618	224662	246405	219136	209036	219193

Table 9. Placebo Tests

This table reports the results under a placebo event setup, assuming the event occurs in 1998 and 2010 correspondingly. Cumulative abnormal returns (*CAR*) were calculated based on the market adjusted model. Firm-level controls were also included in all specifications and were the same as in Table 6. Standard errors were clustered by firm and fiscal year and reported in the parentheses. Fiscal year, 3-digit SIC, and firm fixed effects are included in all specifications. Column (1) to (3) includes all samples from the year 1993 to 2001 (i.e., pre-SOX period) and column (4) to (6) include samples from 2003 to 2017 (i.e., post-SOX period). Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For a detailed definition of variables, please refer to Appendix 1.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	CAR(-1,+1)	*CAR(-1,+1)*	*CAR(-1,+1)*	*CAR(-1,+1)*	*CAR(-1,+1)*	*CAR(-1,+1)*
	a = 1998	*a = 1998*	*a = 1998*	*a = 2010*	*a = 2010*	*a = 2010*
dNettone	0.0016			0.0032		
	(0.001)			(0.002)		
dNettone * POST(a)	0.0011			0.0032		
	(0.001)			(0.003)		
dNegative		-0.0008			-0.0012**	
		(0.000)			(0.000)	
dNegative * POST(a)		-0.0005			-0.0002	
		(0.001)			(0.001)	
dPositive			0.0008			0.0010
			(0.001)			(0.002)
dPositive * POST(a)			-0.0003			0.0016
			(0.002)			(0.003)
POST(a)	-0.0025***	-0.0026***	-0.0026***	0.0006	0.0006	0.0006
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Firm-level Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.04	0.04	0.04	0.03	0.03	0.03
Observations	56300	56427	56427	154436	154450	154450

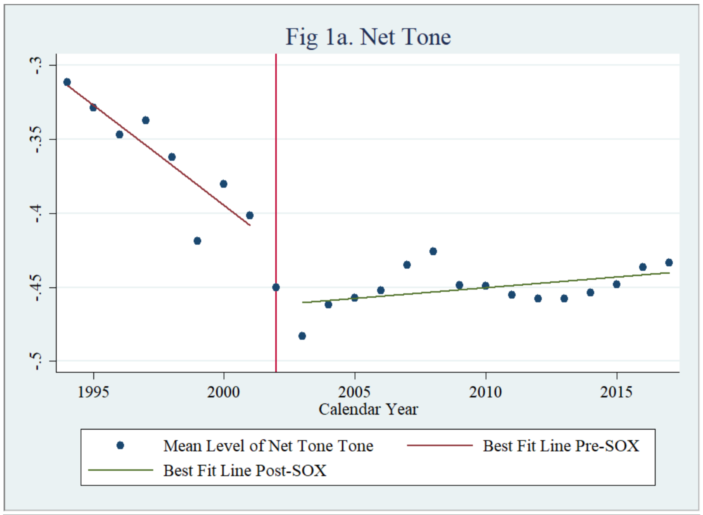


Fig 1a. Net Tone
-.3
-.35
-.4
-.45
-.5
1995
2000
2005
2010
2015
Calendar Year
Mean Level of Net Tone Tone
Best Fit Line Pre-SOX
Best Fit Line Post-SOX

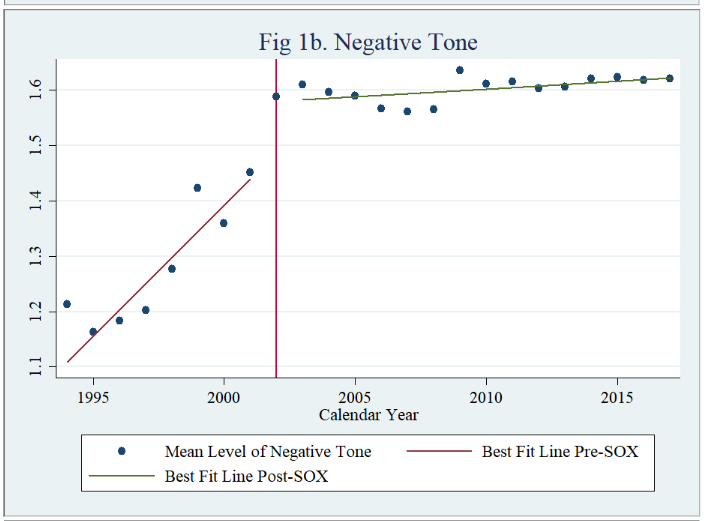


Fig 1b. Negative Tone
1.6
1.5
1.4
1.3
1.2
1.1
1995
2000
2005
2010
2015
Calendar Year
Mean Level of Negative Tone
Best Fit Line Pre-SOX
Best Fit Line Post-SOX

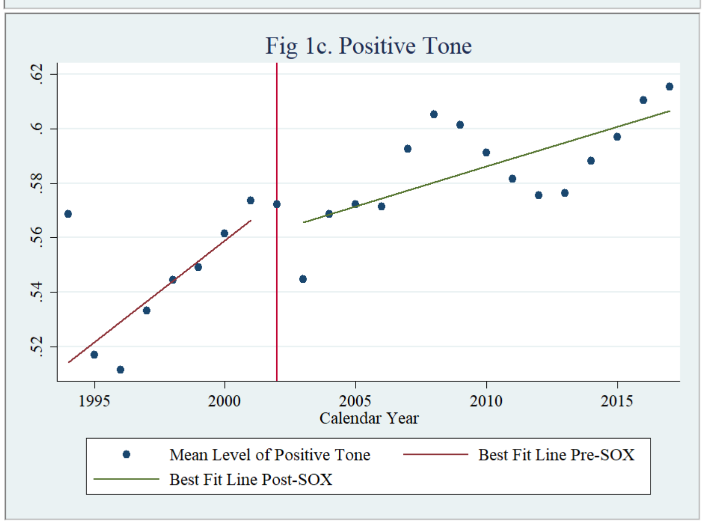


Fig 1c. Positive Tone
.62
.6
.58
.56
.54
.52
1995
2000
2005
2010
2015
Calendar Year
Mean Level of Positive Tone
Best Fit Line Pre-SOX
Best Fit Line Post-SOX

Figure 2. Tone Distributions

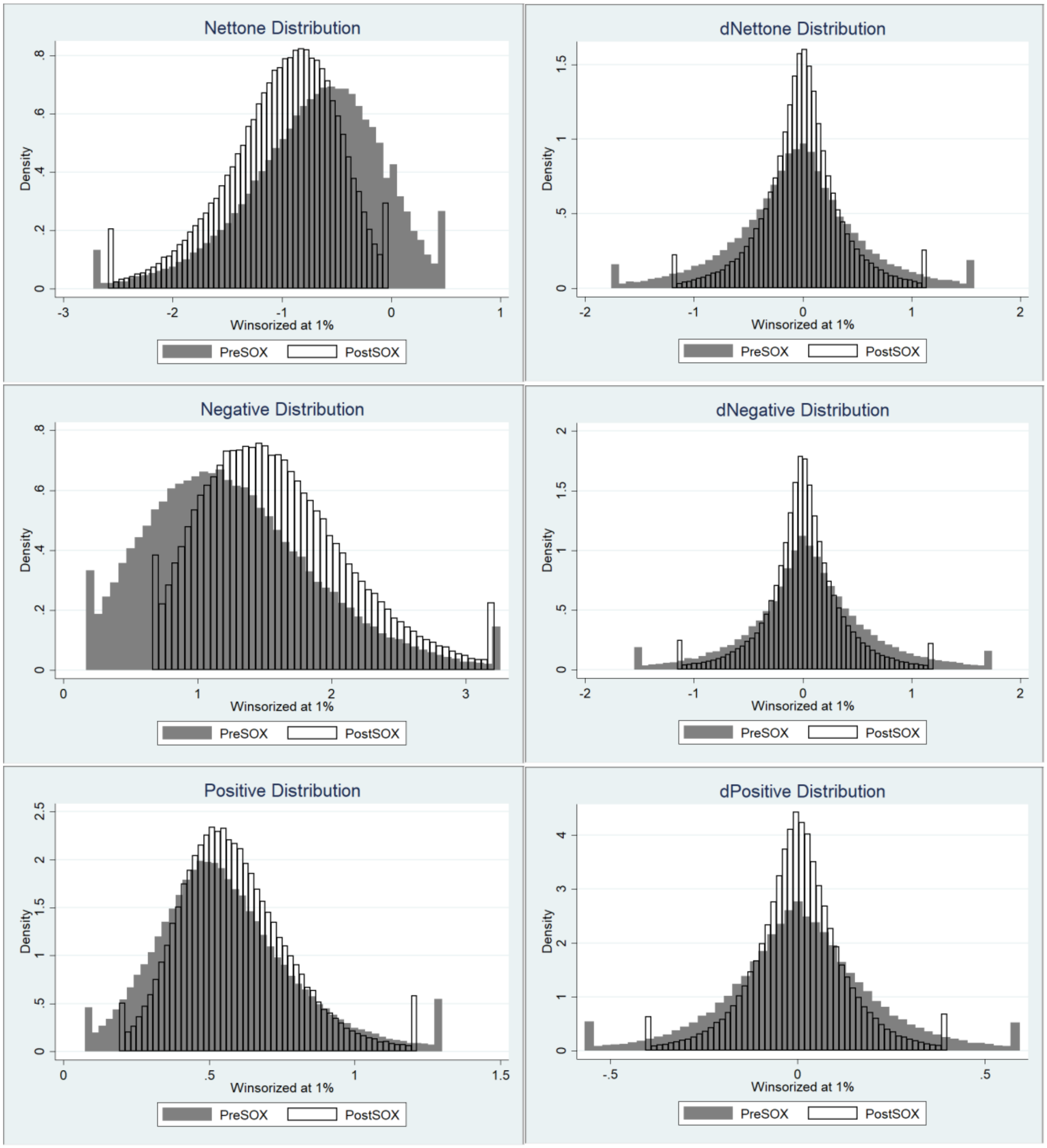
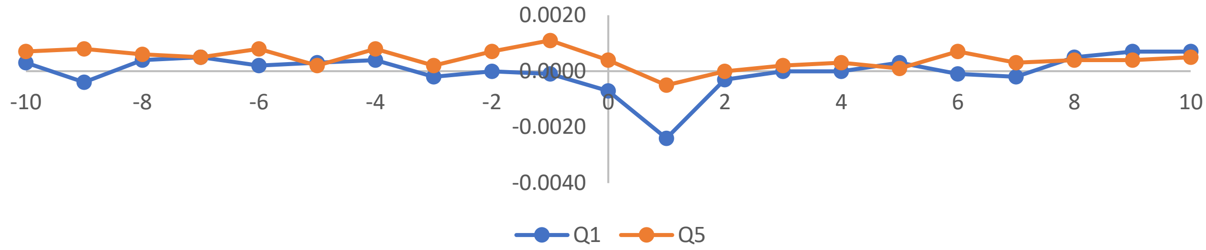


Figure 3a. Mean Daily CAR Based on dNettone Quintile



Q5 - Q1

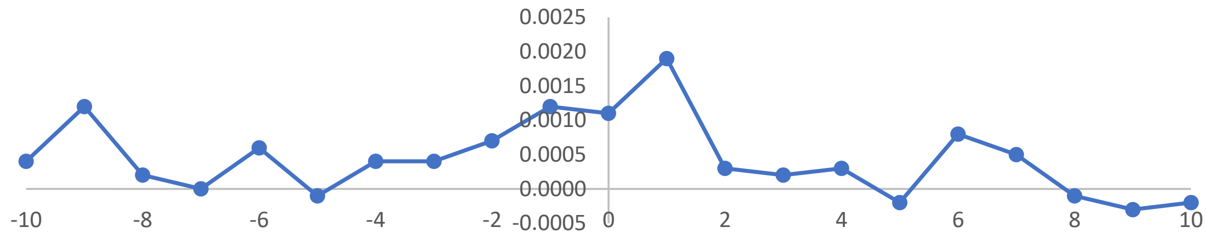
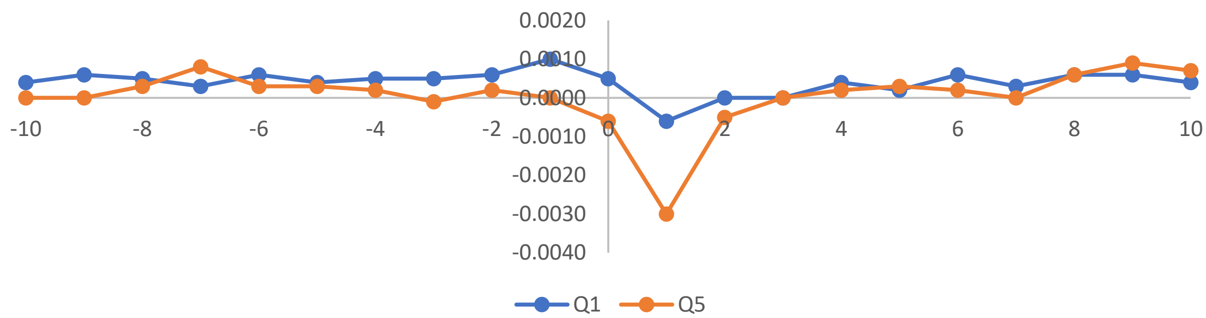
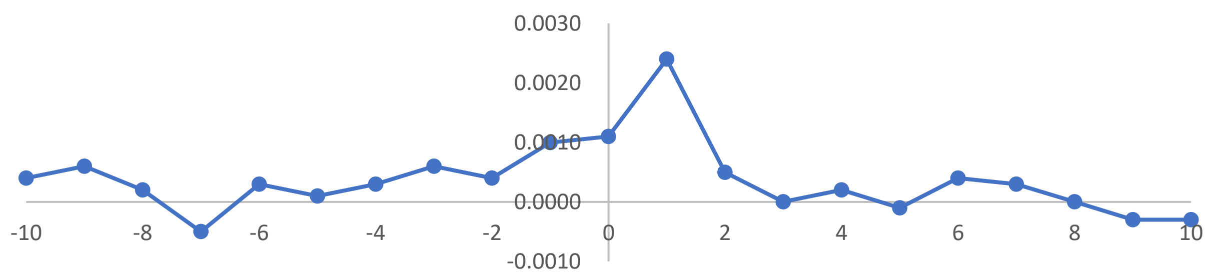


Figure 2b. Mean Daily CAR Based on dNegative Quintile



Q1 - Q5



Appendix 1. Variable Definitions

Variable Name	Description
Positive	The total number of positive words normalized by total words in the documents x 100 using word lists from Loughran and McDonald (2011).
Negative	The total number of negative words normalized by total words in the documents x 100 using word lists from Loughran and McDonald (2011).
Nettone	The total number of positive words subtract negative words normalized by total number of positive and negative words in the documents x 100 using word lists from Loughran and McDonald (2011).
dPositive	$Positive_t – Positive_{t\text{-}4}$
dNegative	$Negative_t – Negative_{t\text{-}4}$
dNettone	$Nettone_t – Nettone_{t\text{-}4}$
Positive_S	*Positive* subtracts the mean of the previous four quarters and divides by the standard deviation of the same period.
Negative _S	*Negative* subtracts the mean of the previous four quarters and divides by the standard deviation of the same period.
Nettone _S	*Nettone* subtracts the mean of past four quarters and divides by the standard deviation of the same period.
dPositive_S	*dPositive* subtracts the mean of past four quarters and divides by the standard deviation of the same period.
dNegative _S	*dNegative* subtracts the mean of past four quarters and divides by the standard deviation of the same period.
dNettone_S	*dNettone* subtracts the mean of past four quarters and divides by the standard deviation of the same period.
FILESIZE	Natural log of formatted 10-Q and 10-K file size in kilobytes as in Loughran and McDonald (2011).
10K	A dummy equals to 1 if it is a 10-K filing.
EARN	The quarterly earnings normalized by the book assets.
SIZE	The natural log of market value

Appendix 1 Continued

SUE	Compustat-based standardized unexpected earnings as in Livnat and Mendenhall (2006) where EPS assumes to follow a seasonal random walk and the best expectation of the EPS is the EPS 4 quarters ago.
ACCRUALS	The quarterly earnings subtract cashflow from operations normalized by the book value of total assets
TOBINSQ	The total market value plus the book value of debt normalized by the book value of total assets.
EARNVOL	The standard deviation of quarterly *EARN* of the past 5 years
DERATIO	The total value of liabilities divided by market value of common equity
POSTSOX	A dummy equals to 1 if the filing date is after 31 July 2002.
CAR(X,Y)	Cumulative abnormal return using the market-adjusted model between day *X* and day *Y* around the filing date.
RET(X,Y)	Raw return between day *X* and day *Y* around the filing date.
RMSE(X,Y)	The root mean square error of a market model return regression between day *X* and day *Y* after the filing date.

Appendix 2. Pair-wise Correlations Table

This table reports the full correlations of tone variables with firm characteristics and returns related measures. Correlations with * indicate significance at 5% level. For a detailed definition of variables, please refer to Appendix 1.

		(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	***Positive***	1.0000						
(2)	***Negative***	0.2367*	1.0000					
(3)	***Nettone***	0.5389*	-0.6373*	1.0000				
(4)	***dPositive***	0.3484*	0.0215*	0.2456*	1.0000			
(5)	***dNegative***	0.0216*	0.3348*	-0.2593*	0.0681*	1.0000		
(6)	***dNettone***	0.2162*	-0.2101*	0.3790*	0.6143*	-0.6537*	1.0000	
(7)	***Positive_S***	0.4538*	0.0695*	0.2727*	0.6048*	0.0404*	0.3696*	1.0000
(8)	***Negative_S***	0.0739*	0.3639*	-0.2476*	0.0370*	0.6104*	-0.4073*	0.1128*
(9)	***Nettone_S***	0.2752*	-0.2363*	0.4240*	0.4079*	-0.4216*	0.6046*	0.5097*
(10)	***dPositive_S***	0.1881*	0.0194*	0.1270*	0.6685*	0.0626*	0.4087*	0.5581*
(11)	***dNegative_S***	0.0156*	0.1685*	-0.1290*	0.0538*	0.6482*	-0.4295*	0.0422*
(12)	***dNettone_S***	0.1129*	-0.1053*	0.1837*	0.4386*	-0.4268*	0.6308*	0.3471*
(13)	***Filesize***	0.2639*	0.3797*	-0.1539*	-0.0183*	0.0985*	-0.0815*	0.1729*
(14)	***EARN***	-0.0642*	-0.1270*	0.0693*	0.0136*	-0.0487*	0.0413*	0.0131*
(15)	***SIZE***	0.1892*	0.1178*	0.0462*	0.0020	-0.0143*	0.0126*	0.0071*
(16)	***SUE***	0.0098*	-0.0063*	0.0129*	0.0208*	-0.0538*	0.0517*	0.0133*
(17)	***ACCRUALS***	-0.0386*	0.0233*	-0.0449*	-0.0008	-0.0196*	0.0148*	-0.1020*
(18)	***TOBINSQ***	0.0864*	-0.0128*	0.0765*	0.0039	-0.0137*	0.0153*	-0.0037
(19)	***EARNVOL***	0.0128*	0.0233*	-0.0112*	-0.0024	-0.0071*	0.0027	-0.0043
(20)	***DERATIO***	-0.0769*	0.1155*	-0.1647*	-0.0028	0.0335*	-0.0204*	-0.0030
(21)	***CAR(-1,+1)***	0.0055*	-0.0033	0.0074*	0.0043	-0.0118*	0.0126*	0.0070*
(22)	***CAR(-10,-6)***	-0.0007	-0.0111*	0.0114*	0.0053*	-0.0070*	0.0086*	0.0042
(23)	***CAR(-10,-2)***	0.0011	-0.0163*	0.0173*	0.0080*	-0.0133*	0.0160*	0.0099*
(24)	***RET(-10,-1)***	0.0015	-0.0080*	0.0099*	0.0031	-0.0105*	0.0096*	0.0065*
(25)	***RMSE(0,22)***	-0.0108*	0.0138*	-0.0183*	0.0059*	0.0361*	-0.0186*	0.0005
(26)	***RMSE(6,28)***	-0.0073*	0.0104*	-0.0123*	0.0065*	0.0404*	-0.0221*	0.0064*

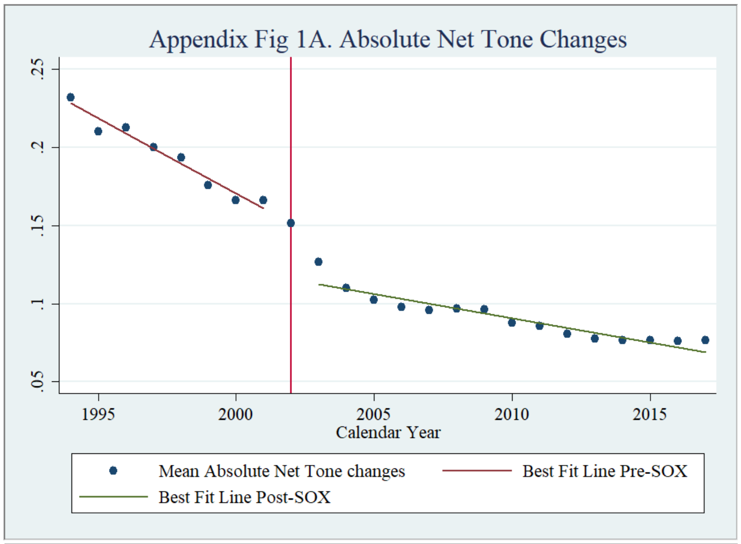
		(8)	(9)	(10)	(11)	(12)	(13)	(14)
(8)	***Negative_S***	1.0000						
(9)	***Nettone_S***	-0.4898*	1.0000					
(10)	***dPositive_S***	0.0454*	0.3693*	1.0000				
(11)	***dNegative_S***	0.5821*	-0.3814*	0.0634*	1.0000			
(12)	***dNettone_S***	-0.3526*	0.5835*	0.5557*	-0.5365*	1.0000		
(13)	***Filesize***	0.1681*	-0.0113*	-0.0128*	0.0742*	-0.0699*	1.0000	

Appendix 2 continued

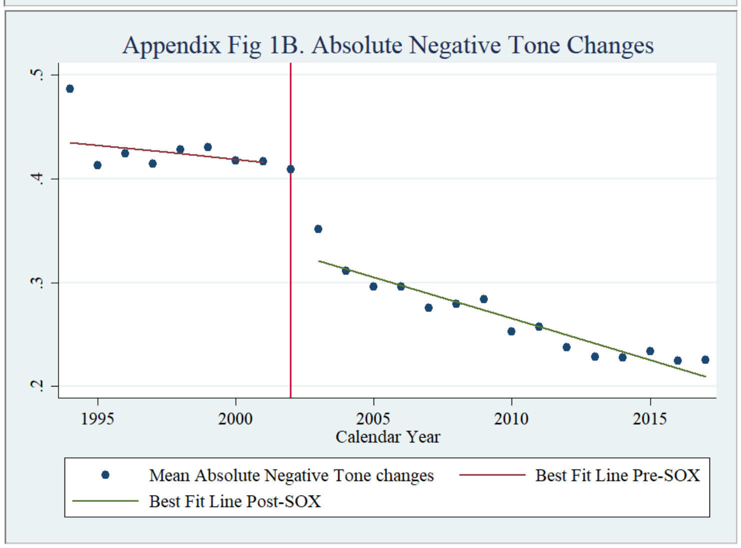
(14)	***EARN***	-0.0360*	0.0420*	0.0093*	-0.0145*	0.0186*	-0.0727*	1.0000
(15)	***SIZE***	-0.0176*	0.0118*	-0.0020	-0.0076*	-0.0042	0.3108*	0.1256*
(16)	***SUE***	-0.0444*	0.0410*	0.0186*	-0.0467*	0.0475*	-0.0005	0.2691*
(17)	***ACCRUALS***	-0.0486*	-0.0334*	-0.0006	-0.0145*	0.0118*	-0.1266*	-0.0352*
(18)	***TOBINSQ***	-0.0109*	0.0035	-0.0013	0.0005	-0.0019	-0.0484*	-0.0416*
(19)	***EARNVOL***	-0.0088*	0.0014	0.0013	-0.0013	0.0006	0.0037	-0.1563*
(20)	***DERATIO***	0.0253*	-0.0229*	0.0046	0.0053*	0.0029	0.0880*	-0.0744*
(21)	***CAR(-1,+1)***	-0.0103*	0.0145*	0.0029	-0.0101*	0.0097*	0.0056*	0.0556*
(22)	***CAR(-10,-6)***	-0.0058*	0.0063*	0.0066*	-0.0054*	0.0117*	-0.0150*	0.0340*
(23)	***CAR(-10,-2)***	-0.0079*	0.0129*	0.0074*	-0.0091*	0.0152*	-0.0178*	0.0498*
(24)	***RET(-10,-1)***	-0.0036	0.0068*	0.0035	-0.0073*	0.0115*	-0.0104*	0.0500*
(25)	***RMSE(0,22)***	0.0305*	-0.0184*	-0.0007	0.0144*	-0.0086*	-0.1006*	-0.1628*
(26)	***RMSE(6,28)***	0.0374*	-0.0200*	0.0002	0.0138*	-0.0076*	-0.1102*	-0.1587*

		(15)	**(16)**	**(17)**	**(18)**	**(19)**	**(20)**	**(21)**
(15)	***SIZE***	1.0000						
(16)	***SUE***	0.0145*	1.0000					
(17)	***ACCRUALS***	-0.0962*	0.1191*	1.0000				
(18)	***TOBINSQ***	0.1448*	0.0120*	-0.0804*	1.0000			
(19)	***EARNVOL***	-0.0350*	0.0038	0.0255*	0.1291*	1.0000		
(20)	***DERATIO***	-0.1996*	-0.0397*	0.1216*	-0.2154*	-0.0496*	1.0000	
(21)	***CAR(-1,+1)***	0.0050	0.0479*	-0.0179*	-0.0103*	-0.0058*	0.0080*	1.0000
(22)	***CAR(-10,-6)***	-0.0016	0.0274*	-0.0089*	-0.0102*	-0.0088*	0.0016	-0.0096*
(23)	***CAR(-10,-2)***	-0.0055*	0.0361*	-0.0102*	-0.0129*	-0.0083*	0.0073*	-0.0300*
(24)	***RET(-10,-1)***	0.0062*	0.0339*	-0.0150*	-0.0139*	-0.0094*	0.0091*	0.1508*
(25)	***RMSE(0,22)***	-0.3339*	-0.0673*	0.0322*	0.0649*	0.0739*	0.0516*	-0.0195*
(26)	***RMSE(6,28)***	-0.3356*	-0.0671*	0.0313*	0.0631*	0.0715*	0.0667*	-0.0470*

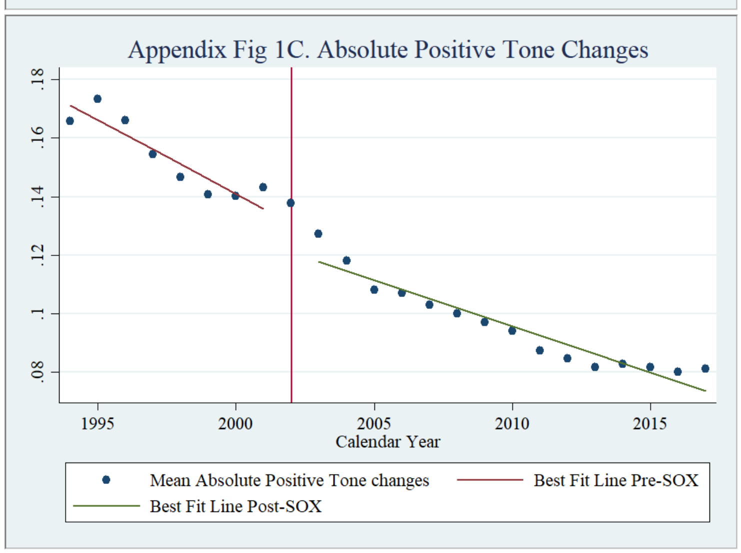
		(22)	**(23)**	**(24)**	**(25)**	**(26)**
(22)	***CAR(-10,-6)***	1.0000				
(23)	***CAR(-10,-2)***	0.7231*	1.0000			
(24)	***RET(-10,-1)***	0.6281*	0.8673*	1.0000		
(25)	***RMSE(0,22)***	-0.0380*	-0.0504*	-0.0689*	1.0000	
(26)	***RMSE(6,28)***	-0.0318*	-0.0393*	-0.0530*	0.8560*	1.0000



Appendix Fig 1A. Absolute Net Tone Changes
.05
.1
.15
.2
.25
1995
2000
2005
2010
2015
Calendar Year
Mean Absolute Net Tone changes
Best Fit Line Pre-SOX
Best Fit Line Post-SOX



Appendix Fig 1B. Absolute Negative Tone Changes
.2
.3
.4
.5
1995
2000
2005
2010
2015
Calendar Year
Mean Absolute Negative Tone changes
Best Fit Line Pre-SOX
Best Fit Line Post-SOX



Appendix Fig 1C. Absolute Positive Tone Changes
.08
.1
.12
.14
.16
.18
1995
2000
2005
2010
2015
Calendar Year
Mean Absolute Positive Tone changes
Best Fit Line Pre-SOX
Best Fit Line Post-SOX

Chapter 3. Does the Exit of Venture Capitalists Affect Innovation? Evidence from IPOs

1. Introduction

Venture Capitalists (VCs) play a significant role in funding many new ventures today. Startup firms often reach out to VCs due to difficulty in securing loans from banks at the early stage. According to the National Venture Capital Association (2016), VC backed IPOs account for a significant 32.15% of the total IPOs in the US between 1995 and 2015 and their average exit time is 5.72 years. A paper by Gompers and Lerner (1999) further brings out the importance of VC to finance entrepreneurs' early-stage ideas from a demand and supply perspective. Cumming and Johan (2012) also reiterate how VCs can provide essential help to the entrepreneurial firms prosper by lending financial, strategic, marketing, legal, administrative, or human resource advice and business network. The infusion of VCs into a startup naturally increases the financial support and strategic guidance to a firm's innovation process, hence promoting greater investment in Research and Development (R&D). This is aligned with many studies that document empirical evidence of a positive correlation between innovation and VCs (e.g., Kortum and Lerner, 2000; Hirukawa and Ueda, 2011; Cumming and Li, 2013).

One of the biggest challenges, however, in extant literature concerns the preselection biases where VCs may select firms based on unobserved characteristics that account for the disparity in innovation level between VC and non-VC backed firms. While various studies have tried to address the endogeneity issues through selection models and matching, it is challenging to find a clean experimental setup due to the lack of counterfactuals should the VCs not invest in these firms. In this paper, we treat the initial public offerings (IPO) as a proxy for VC's exits since most VCs exit after IPO due to their limited investment horizon. Using a difference-in-differences (DiD) approach, we investigate the relation between VCs and innovation from a unique "reverse treatment" perspective by comparing the drop in innovation intensity between VC backed firms and the non-VC backed after IPOs. Intuitively, should VC contribute significantly to a firm's innovation, it is also likely that we will observe a bigger fall in innovation intensity after it leaves unless a firm learns from the experiences, guidance, and influence brought about by the VC.

We first document that VC backed firms are more innovative both before and after the IPO. Under our DiD framework, the VC backed firms on average experience a greater fall in the R&D intensity after IPO as compared to their non-VC backed counterparts. Similar results are also observed when we instead use innovation outputs such as patent counts and forward citations. Arguably, this withdrawal effect in innovation is ascribed to the exiting of VCs after we control for firm characteristics and fixed effects. This main finding is further confirmed with the propensity score matching of VC backed firms with similar non-VC backed firms in our attempt to address the potential endogeneity concerns.[33] Our argument is that VC helps R&D investments by providing expertise and resources to firms in both financial and operational aspects. As such, firms may experience a greater withdrawal effect in R&D intensity after VCs' exit.

Upon documenting a greater fall in R&D intensity for the VC backed firms after IPO, this paper further investigates whether VC involvement shapes the way firm innovates after VC leaves. Potentially, firms learn and inherit from VCs' guidance to retain propensity to innovate. If the involvement from VCs in any way shapes the innovation behavior of the firms, we might expect a

[33] We fully acknowledge the limitation of matching in addressing endogeneity since we will only be able to match on the observed characteristics. Following prior literature, our matching covariates always include the States of incorporation and industry which are pre-determined before receiving investments from the VCs and are unchanged afterwards.

smaller drop in R&D intensity in VC-backed firms with greater pre-IPO VC involvement holding other factors equal. Otherwise, we might also observe a greater withdrawal effect if we expect that a higher VC involvement is associated with closely monitoring and greater incentives for pre-IPO innovation. The two abovementioned effects are non-mutually exclusive and the net effect really depends on whether the withdrawal effect from VC's exit overshadows any positive influence inherited from the VCs with respects to innovation. While VC firms on average experienced a greater dip in R&D intensity after IPO, we find that such drop is more pronounced for firms with greater pre-IPO VC involvement. Our empirical evidence hence suggests that the withdrawal effect is more dominant for firms with a higher level of pre-IPO VC involvement. This is also consistent with our main findings that VCs positively affect innovation.

A plausible explanation from Darrough and Rangan (2005) that VCs no longer have incentives to upkeep the long-term performance of the firms when their exit is imminent does not contradict our results. If it is indeed a change in the incentives of the major stakeholders after IPO that causes less motivation for R&D investment, we should see such a drop in R&D intensity partially alleviated in the presence of new investors who are capable of replacing the role of the VCs. Cadman and Sunder (2014) suggest that the institutional investors who have long-term horizon incentives and possess a large enough ownership will counter the short-term incentives of some pre-IPO stakeholders (i.e., VCs). Consistent with their argument, we find evidence that the presence of high institutional holdings does moderate such a negative impact in R&D intensity when VCs leave, though not entirely. Institutional investors with their longer investment horizon arguably have sufficient motivations to replace the monitoring role of VCs and induce firms to innovate for prospective gains.

The background of our study is related to Bernstein (2015) who suggests that firms going IPOs experience a drop in innovation quality. Nonetheless, ours differs in two important aspects: First, we focus on the comparison of VC backed and non-VC backed firms using IPO as an exit event whereas Bernstein (2015) is comparing IPO firms with those that withdrew from IPO and remained private for reasons unrelated to innovation. Second, we focus on R&D expenditure to gauge the level of innovation which is consider an ‘input’ to innovation whereas Bernstein focuses on patents related measures which are considered the ‘output’ from innovation. The firm-specific efficiency in the translation process from input to output as well as the lag between developing, filing, and issuing of a patent can be difficult to account for. Nonetheless, we include the number of patents filed and citations as additional checks. From the accounting literature, some studies examine a firm’s manipulative behaviors when short-term investors exit. For instance, Ertimur *et al*. (2014) find evidence that managers delay the disclosure of bad news to uphold stock prices when insiders sell during an IPO. Darrough and Rangan (2005) also show that the insiders of firms manipulate short-term earnings when there are insider sellings during an IPO. Our paper differs by shedding light not only on the motivations of different stakeholders in pursuing innovation around IPO but also whether VC’s pre-IPO positive impact on innovation will persist after IPO-exit. While Darrough and Rangan (2005) conclude a negative effect between insider selling and R&D expenditure from a cross-sectional test, we also aim to provide reasonable time-series variations to study the impact of VC’s exit on innovation.

Our paper first complements the literature by documenting evidence on the monitoring role of VC as well as its influence on the firm’s innovation around IPO-exit. We pay particular attention to the IPO-exit since it is arguably the most important event in the life cycle of venture capitals. Under a “reverse treatment” setup, we show that firms experience a greater drop in R&D intensity when VC's exit becomes imminent, which indirectly lends support to the argument that the presence of VC encourages innovation. We also fill the void in extant literature by examining whether VCs create persistency in firms' innovation and whether the presence of more sophisticated institutional investors with long-run interest in the firm might take on the role of VCs to continue propelling innovation after IPO. Given the importance of VC financing in spurring economic growth, this paper seeks to evaluate the role of VC in firm innovation at different stages of a firm's life cycle. Therefore, we hope to provide some empirical insights for the policymakers about how the turnover of different stakeholders may impact firms' R&D decisions and hence the competitiveness of the local economy.

The rest of our paper proceeds as follows: Section 2 discusses the contemporary literature of the relation between VCs and innovation as well as the formulation of our hypotheses based on extant research. Section 3 outlines the sample construction procedures and describes the preliminary results. Section 4 delves into the empirical strategies to examine the effects of VC's exit. Section 5 discusses the results and Section 6 concludes.

2. Theory and Hypotheses

The literature about the impact of VCs on innovation is well established although with some mixed evidence. Our paper contributes by providing direct evidence on the role of VC on promoting innovation from IPO-exits. Many document a positive association between VCs and innovation, as measured by patents (e.g., Kortum and Lerner, 2000; Hirukawa and Ueda, 2011; Cumming and Li, 2013). Dushnitsky and Lenox (2005) also suggest a positive correlation between the presence of VCs and the quality of patents. Hellmann and Puri (2000) use a manual collection of Silicon Valley firms and show that innovators are more likely to obtain VCs than imitators and that VC backed firms are quicker in bringing products to market. Using samples from Germany and Italy respectively, Engel and Keilbach (2007) and Caselli *et al*. (2009), however, find that the positive correlation between innovation and VCs might be due to a pre-investment selection process (i.e., endogeneity). While prior literature has tried to tackle this endogeneity issue through selection models and matching of similar non-VC backed firms, it is seemingly impossible to fully resolve the preselection bias of the VCs at ventures' initial stage due to missing counterfactuals. While ample empirical evidence has seemingly justified the beneficial role of VCs in propelling innovation, none has explicitly studied this relation from the perspective of exit events. If the presence of VCs indeed stimulates firms' innovation, we should observe at least some withdrawal effect in innovation when they exit when compared to similar firms without VCs' backing, unless a firm successfully inherits the beneficial characteristics of VCs before they leave.

Our paper fits under the bigger umbrella of the relations among VC, investors, and the entrepreneurial firm, with an emphasis on the IPO-exit.[34] The theme of our paper is related to extant literature that documents the use of staged financing by the VCs to time their exits (e.g., Gompers, 1995; Bergemann and Hege, 1998; Cornelli and Yosha, 2003). VCs have a significant influence on the major operational decisions of the firms and they use contractual agreements that grant them intervention rights related to their subsequent exit routes (e.g., Gompers, 1999; Kaplan and Stromberg, 2003). As noted by Lerner (1994), VCs are indeed a group of sophisticated investors with a clear investment goal who enter and exit at the most appropriate time to maximize their returns. Schwienbacher and Giot (2007) study the dynamics of exit options for venture capitals and find that a large syndicate size accelerates the exit of all types (E.g., IPO-exit, trade sale, or liquidation) but the greatest impact is for IPOs. Another recent study by Puri and Zarutskie

[34] A survey study conducted by Da Rin *et al*. (2013) summarizes the past and contemporaries research on VCs.

(2012) find that VC backed firms are less likely to fail and more likely to go public and be acquired relative to the non-VC backed firms. Although many VCs exit through trade sales, Cumming and Johan (2012) asserts that going public is typically viewed as the best exit outcome for VCs. Therefore, our paper focuses centrally on the IPO-exits, which is deemed the most successful way of exiting an investment in terms of returns.

From the perspective of an exit, Schwienbacher (2008) shows that more innovative and profitable ventures are more likely to go public than those more imitative or with derivative projects and that the uncertainty associated with the exit stage of VC will also induce the entrepreneur to adopt strategies that increase the likelihood of an IPO, leading to greater innovation. However, VC's effort in promoting innovation should subdue upon the exit of VCs and we should subsequently expect a withdrawal effect in R&D intensity.[35] Therefore, using a sample of the U.S. IPO firms, we argue that:

H1: The exit of venture capitalists will negatively impact a firm's innovation.

VCs will benefit the entrepreneurs by providing them with strategic advice and coaching in addition to providing financing, which is consistent with the findings of Baker and Gompers (2003) that the outperformance of VC backed firms is ascribed to the certification and coaching effects. We expect the positive influences from VCs to dissipate once their exits become imminent. Cumming (2008) suggests that VC control rights are relevant in the exit decisions and that VCs

[35] Darrough and Rangan (2005) investigate a subsample of IPOs from 1986 to 1990 and find that firm's insiders might manipulate stock performance by cutting R&D expenditures to increase current reported earnings when they sold their shares at the time IPO. Their paper focuses solely on earnings manipulation at the point of IPO and does not emphasize the general level of R&D for the many years after IPO. Their findings do not contradict our argument that there is a change in incentives to do innovation after IPO for the VCs and such effect may not only be attributed to the potential earnings manipulation, but also to the withdrawal effect from high pre-IPO investment in R&D in the first place which may explains why this negative effect exists even in a prolonged period after IPO.

are less likely to gain utility from the private benefits of the entrepreneurs (e.g., the reputation associated with being a public firm CEO) and therefore they will be less actively involved after IPOs.[36]

Our paper also studies the relation between VCs and innovation by examining the degree of pre-IPO VC involvement. Prior literature suggests that funds with larger portfolios, in terms of the number of firms financed per manager, will likely spend less effort to monitor and advise their investee firms (e.g., Kanniainen and Keuschnigg, 2003; Keuschnigg, 2004; Kanniainen and Keuschnigg, 2004; Cumming, 2006). Hence, we expect that the degree of VC involvement will set their investee firms' innovation strategies apart. It is not unreasonable to assume that firms learn and inherit from VCs' coaching and retain the high propensity to innovate. If higher involvement from VCs in any way shapes better the innovation behavior of the firms, we might expect a smaller drop in R&D intensity after IPO-exit in VC-backed firms. However, should we also expect that a higher pre-IPO VC involvement is associated with greater VC participation to push for innovation, we might observe a stronger withdrawal effect when VCs exit. The above two effects are certainly not mutually exclusive and to study which of the two effects plays a dominant role in affecting a firm's R&D around IPO, we formulate the following pair of hypotheses:

H2a: The negative impact on a firm's innovation during the exit of venture capitalists will be greater for firms with a high level of pre-IPO involvement.

[36] Many studies argue that the private benefits of the entrepreneurs are higher after IPOs as compared to after trade sales (e.g., Black and Gilson, 1998; Bascha and Walz, 2001; Hellmann, 2006).

H2b: The negative impact on a firm's innovation during the exit of venture capitalists will be smaller for firms with a high level of pre-IPO involvement.

The intuition behind *H2a* is that the reversion in R&D intensity after VCs exit will dominate any positive traits inherited from VCs' involvement. The reverse is true for *H2b*. There exists a tension since both effects can come into play simultaneously. Our empirical results may shed some light on which one dominates.

Bolton *et al*. (2006) provide a theoretical framework explaining how managerial short-termist behavior can be a direct outcome from the speculative motive of the firm's existing controlling shareholders. While VC is usually the dominant controlling party in a VC backed IPO, the controlling interests may be greatly reshuffled after IPO. Similar to the VCs, institutional investors who own a significant stake in the equity will directly take on a monitoring role in mitigating potential agency problems within firms (Hartzell and Starks, 2003). Bushee (1998) finds that firms are less likely to reduce R&D spending in an attempt to reverse an earnings decline in the presence of high institutional holdings because institutional investors' monitoring reduces pressures for potential myopic behavior. This finding is consistent with Cadman and Sunder (2014) who suggest that institutional investors with significant ownership can counter short-termist behavior in a firm. Therefore, we hypothesize that the continuous participation of sophisticated institutional investors can partially take on the role of the VCs and potentially limit potential myopic behaviors that overweight short-term performance over long-run prospects. We expect the negative impact on innovation when VC exits will be countered partially, if not all, in the presence of high institutional monitoring. This then leads to our third hypothesis:

H3: *The negative impact on a firm's innovation during the exit of venture capitalists will be smaller with the presence of high institutional monitoring.*

Institutional investors with longer investment horizons do not gain from the myopic decisions made by the managers to temporarily uphold stock performance to please the short-term investors at the cost of long-term prospects. The intuition behind *H3* is that the presence of institutional investors will continue to encourage and monitor the R&D investment and ensures that short-term investors do not exploit the firm at the cost of long-term performance. Arguably, existing stakeholders and the new post-IPO investors who plan to stay with the firm for longer terms may want to strike a balance between growth prospects and short-term stock performance.

3. Sample Construction and Summary Statistics

Our initial sample contains both VC and non-VC backed firms that go public for the first time between 1972 and 2018. The firms are identified through Thomson Financial's Global New Issues database from the Securities Data Company (SDC). We exclude ADRs, REITs, spinoffs, carve-outs, closed-end funds or trusts, and firms that go public for a second time or with total proceeds of lower than $5 million which is consistent with prior literature on IPOs to exclude noisy observations. The sample size is further trimmed down by excluding financial firms with a SIC code that begins with 6 and in the regulated utility industry (i.e., SIC 4800-4999). There are a few main reasons for such exclusion. First, firms that begin with the SIC code of 6 are mainly banks, financial institutions, or insurance companies that rarely spend on R&D, and their R&D expenditure data from the Compustat database is usually missing or zero. The structure of such firms also significantly differs from the rest. Second, such firms are mainly non-VC backed and only a very few VC backed IPOs are identified in this group each year. Firms in the regulated industry will also differ significantly in terms of firm operations than a regular firm. We decided to exclude these special categories of IPOs to avoid potential bias.

The initial sample is then merged with the Compustat database on WRDS for balance sheet variables. Accounting data of those firms on Compustat are usually available at least 2-3 years before an IPO due to disclosure requirements.[37] The number of firms is further trimmed down to 3,412 with 37,884 firm-year observations.[38] We then merged with the NBER patent database which includes patents filed and citations counts for the empirical tests on innovation quality.[39] We obtain data on institutional investors from 13F filings and calculate the average quarterly institutional holdings of a particular year as the concentration of institutional ownership. We then indicate a firm as highly monitored by institutional investors (i.e., *HIGHINST = 1*) when the institutional ownership of a firm is above the median ownership within its respective 2-digit SIC industry of that year.

[Table 1 Here]

Panel A of Table 1 shows the distribution of IPOs under Fama-French 30 Industry Classification and Panel B illustrates the level of mean R&D intensity between VC backed and non-VC backed firms within a broader Fama-French 12 Industry Classification. A significant proportion of IPOs are from the Healthcare, Services, Business equipment, and Retail industries. Firms in different industries seem to differ systematically in terms of innovation and VC backed firms are on average more innovative than their non-VC counterparts. Therefore, we include an even more refined

[37] It is interesting to note that some firms have more than 3 years of accounting data available before IPO. However, not all firms contain complete accounting information for our regressions and are dropped accordingly depending on our model specifications.
[38] The final sample size used in our regressions depends on the specifications of our setups.
[39] See https://sites.google.com/site/patentdataproject/

Fama-French 48-industry fixed effect in most of our regressions to control for the systematic difference in the way firms invest in R&D across different industries.

[Fig 1 & 2 Here]

Figure 1 delineates a simple picture of the change of R&D patterns before and after IPOs. Indeed, VC backed firms are more innovative than non-VC backed firms based on the level of R&D investment in both periods on average. VC firms experienced a drop of about 53% in R&D intensity on average whereas that of the non-VC backed is only about 36%. Our empirical setups will take into account the systematic differences in the characteristics across the two groups via firm-level controls and various fixed effects. Figure 2 shows the time-series R&D intensity of all sample firms around the IPO.VC backed firms on average are more innovative both before and after IPO but they encountered a greater drop in R&D intensity after IPO. On average, the R&D intensity of a firm seems to peak 2 to 3 years before the IPO.

[Table 2 Here]

Table 2 exhibits the summary statistics of firm characteristics. Panel A shows the distributions of important firm-level characteristics at various quantiles for the entire sample. On average, firms spent an amount roughly equal to 14.1% of their total assets on R&D and about 6.4% on capital expenditures each year. An average firm seems to experience a negative return on ROA and receiving negative cash flow from operation whereas a median firm does not. Patent and citations counts are strongly skewed to the right tail of the distribution. Panel B shows the mean of all the above firm characteristics grouped by whether a firm is VC backed. We also perform a *t*-test on the difference of mean between VC and non-VC backed firms as shown in the last column. A quick snapshot of the summary statistics reveals several interesting facts: the non-VC backed firms are on average larger than those VC backed. The non-VC backed also seems to be more profitable on average. It is also unsurprising that the VC backed firms are statistically different in many firm characteristics such as leverage, size, and profitability when compared to the non-VC backed. As aforementioned, a common criticism for the comparison of the two groups is selection bias. The systematic differences between the control and treatment groups may lead to potential problems when interpreting the results unless all such differences or selection bias is fully accounted for. While our paper controls for many observable firm characteristics that are thought to affect innovation and perform matching on exogenous covariates such as industry and location of incorporation, we still would like to acknowledge potential omitted variable bias for unobserved heterogeneity and that causal interpretation should be taken with caveats.[40] Panel C of Table 2 reports the difference in average firm characteristics before and after IPO. Size inevitably increases after the IPO. Firms on average experience a drop in R&D intensity after IPO, potentially because of asset expansion outpacing the growth rate in innovation opportunity. The same is true for capital expenditure but to a much smaller extent. Patents and Citations are both higher in the post-IPO periods.[41] Firms also seem to be more profitable after IPO with a much lower leverage on average.

4. Identification and Methodology

In the context of this paper, we refer to the general partners of VC funds as the VCs which are directly identified from SDC.[42] It is common that such partnerships between fund investors and

[40] We have tested the parallel assumption in the difference-in-differences setup and conducted robustness checks such as testing on pseudo-event dates to ensure our results are not spurious.

[41] Although the level of firm R&D intensity is lower after IPO, the absolute amount of R&D expenditure is increased due to access to capital market and the large proceeds received after IPO.

[42] Most venture capitals around the world are limited partnership, with wealthy individuals or institutional investors being the limited partners and the fund managers being the general partners (Cumming and Johan, 2008). The limited partners assume limited liability by not interfering with the general partner's active management who assume unlimited liability and receive compensation in the form of carried interests and management fees.

managers usually last no longer than a decade with an option to extend for at most a few years had the investments not yet come to maturity. Upon IPO, VC will liquidate its position and distribute the proceeds back to their investors.[43] To gauge innovation, we primarily focus on the inputs as measured by the level of R&D investment. R&D is generally considered an investment with benefits to be reaped only after some years.[44] The costs of R&D are usually substantial and with great uncertainty in the future payoffs. Such uncertainty tends to be highest at the initial stage of the investment.

Our study employs a set of firm-level control variables related to R&D as inspired by Kochhar and David (1996), which is common in prior innovation literature and is believed to be highly related to innovation. Profitability measures such as *ROA* can influence the decision of the management in R&D investment and it may also be a direct outcome of R&D investment. Leverage adds to potential financial distress and may affect the budget planning of a firm in allocating resources. Size is a firm characteristic to commonly account for the structural or operational difference between big and small firms. Capital expenditure is the investment in fixed assets which may affect overall budget allocation within a firm due to limited resources. Operating cash flow indicates the ability of a firm to generate enough cash flow for daily operation and future investment. For detailed constructions of the variables used in different specifications, please refer to Appendix A.[45] Continuous variables with extreme observations are winsorized at 1% level of each tail to minimize noise without dropping observations.

[43] Some VCs may gradually distribute the proceeds to their investors over at most a year or two after the offer which can be affected by the lock-up provisions. The lock up period for VCs are usually between 90-270 days based on the data from Thomson Financial's Global New Issues Database in SDC.
[44] As Hall (2002) suggests, R&D should not be analyzed in a static framework since an optimal R&D strategy has options-like features.
[45] For all empirical specifications, control variables are lagged except for *CFO* and *CAPEX* as we believe that concurrent cash flow from operation has a greater impact on the decision of R&D investment than that of previous

To test how the exit of VCs will affect innovation, this paper adopts a multiperiod difference-in-differences framework by treating the IPO-exit of VC as the "treatment" and VC backed firms as the "treated group" since only VC backed firms will experience the exit of VCs after IPO.[46] Intuitively, this treatment dummy can also be viewed as a "reverse treatment" of VCs' monitoring as VCs leave. The post-treatment period is the year of IPO and thereafter and the pre-treatment period before the IPO year. We also include an IPO-year-fixed effect to alleviate potential systematic differences in the way VCs exit at a different point in time. This paper investigates how the IPO-exit of VCs will affect a firm's innovation and the dependent variable of interests will ideally be the measures that proxy for inputs to do R&D. In this context, R&D intensity measures the actual R&D expenditures normalized by firm size for comparison of innovation levels across the board. We specify our baseline model as follow:

$$R\&D_{i,t} = \alpha + \beta_1 VC_i + \beta_2 POST + \beta_3 VC_i * POST + Firm_{FE} + FYear_{FE} + IPOYear_{FE} + Industry_{FE} + YearAfterIPO_{FE} + \sum Controls_{i,t} + \varepsilon_{i,t} \qquad (1)$$

Following our approach, the *POST* dummy is set to 1 for firm-year observation after IPO and 0 otherwise. The *VC* dummy is set to 1 for VC backed firms and 0 otherwise. To estimate the exit effect of VCs, we simply interact the two categorical variables (i.e., *VC * POST*). More precisely, this key interaction dummy takes a value of 1 for all firm-year observations that are VC backed and in the post IPO period, indicating the absence of VCs' monitoring after IPO. The coefficient

quarter and that the capital expenditure is likely to have substituting or complementary effect when it comes to resources allocation during the same quarter.

[46] The 'treatment' is during IPO but only VC backed firms will experience the exit of VCs. We will not be able to fully disentangle the individual IPO effect from the exit of VCs. Perhaps a concerning assumption of this paper is that we assume no systematic difference in the way firms transform from IPO per se and all observed differences are solely because of the confirmed exit of VCs. We argue that there is no obvious reason to believe that VC backed firms will undergo a difference transformation process during IPO compared to a similar non-VC backed counterpart that systematically affects innovation.

β_3 is key to address our research questions. The interaction term gives an estimate of the exit effect, assuming other controls and fixed effects are well addressed in our specifications. If our first hypothesis is true, we should observe the β_3 to be negative and statistically significant. To test hypothesis 2, we use the total known amount of capital invested by VCs as a proxy for direct VC involvement. We then indicate a firm with high VC involvement as one (i.e., *HIGHINVOLVE* = 1) when the VC involvement of a firm is higher than its corresponding 2-digit SIC industry and zero otherwise.[47] A bigger stake in the investment indicates a greater alignment of interests within a start-up for the VCs and hence a stronger motivation for involvement and monitoring before IPO. To investigate the heterogeneous effect of firms with different levels of pre-IPO VC involvement, we apply a triple difference-in-differences approach. An alternative version of our baseline model to be tested is as follow:

$$\begin{aligned} R\&D_{i,t} = \alpha + \beta_1 VC_i + \beta_2 POST + \beta_3 VC_i * POST + \beta_4 HIGHINVOLVE_i + \beta_5 VC_i \\ * POST * HIGHINVOLVE_i + \beta_6 VC_i * HIGHINVOLVE_i + \beta_7 POST \\ * HIGHINVOLVE_i + Firm_{FE} + FYear_{FE} + IPOYear_{FE} + Industry_{FE} \\ + YearAfterIPO_{FE} + \sum Controls_{i,t} + \varepsilon_{i,t} \qquad (2) \end{aligned}$$

Intuitively, a greater stake from the VC should induce the firm to invest more in innovation due to greater monitoring and hence we expect β_4 to be positive. Should VC involvement play a significant role in shaping future R&D investment behavior, we would likely observe β_5 to be statistically significant regardless of the sign. Our second hypotheses jointly test if VC's exit will

[47] Many VCs are quite specialized in specific industries and there likely exists unobserved heterogeneity in the way VCs interact with the entrepreneurial firms across industries. Identifying firms with high involvement using industry benchmark (i.e., industry media) may indicate better the level of involvement. Nonetheless, our results are unchanged if we use a continuous proxy for VC involvement. Also, note that non-VC backed firms may also have VC participation, but the amount of capital invested by VCs in non-VC led IPOs are much less significant compared to the size of the firm before IPO.

negatively impact R&D investment and whether such an effect is stronger for firms with greater VC involvement. The sign of the β_5 will coincide correspondingly to one of the two opposite hypotheses and indicate whether the withdrawal effect will dominate any inheritance of positive traits from the VC with regards to innovation.[48] Nevertheless, our empirical setup does not eliminate the possibility of both effects canceling out each other even when we observe an insignificant β_5.

We defined *HIGHINST* as an indicator variable that equals one if a firm has a yearly average institutional ownership above its respective 2-digit SIC industry median. It is probable that a threshold level of institutional ownership has to exist for institutions to participate in monitoring and such threshold may differ across industries. Unlike smaller individual investors in the secondary market, institutional investors have to file with the SEC Form 13F to disclose their respective holdings every quarter due to significant ownership. Following the same strategy in testing *H2*, we propose the below model to examine *H3*:

$$\begin{aligned} R\&D_{i,t} = \; & \alpha + \beta_1 VC_i + \beta_2 POST + \beta_3 VC_i * POST + \beta_4 HIGHINST_i + \beta_5 VC_i * POST \\ & * HIGHINST_i + \beta_6 VC_i * HIGHINST_i + \beta_7 POST * HIGHINST_i + Firm_{FE} \\ & + FYear_{FE} + IPOYear_{FE} + Industry_{FE} + YearAfterIPO_{FE} \\ & + \sum Controls_{i,t} + \varepsilon_{i,t} \end{aligned} \tag{3}$$

Should there be any significant impact from high institutional monitoring, we should observe β_5 to be positive and significant. The intuition behind our prior is that institutional investors are

[48] Another way to find evidence on whether pre-IPO involvement has a role in R&D investment after IPO is to empirically test only the VC backed firms and check if firms associated with higher pre-IPO involvement exhibit a greater drop in R&D after IPO. But such set-up does not provide a complete insight on the effect of VC's exit when compared to the non-VC backed firms. Note that some of the non-VC led firms can also have some VC investments before IPO.

sophisticated enough and capable of replacing the monitoring role of VCs in allocating sufficient and optimal resources to firms' innovation. The presence of institutional investors can also alleviate potential short-termist behavior during IPO-exits due to their longer investment horizon.

5. Results and Discussions

5.1 VC's Exit

Table 3 summarizes the results of our model (1) and documents evidence in favor of our first hypothesis that the exit of venture capitalists will negatively impact a firm's innovation. Column (1) represents a baseline difference-in-difference setup where VC backed firms were the treatment group and non-VC backed the control without any firm-level controls or fixed effects. Column (2) and (3) add a different combination of fixed effects to (1). Column (4) and (5) include a set of firm-level controls for (2) and (3) respectively.

[Table 3 Here]

In line with prior studies, VC backed firms indeed exhibit a higher level of innovation as shown from the unambiguously positive coefficients of the VC dummy. The coefficients of the interaction term (i.e., *POST * VC*) yield consistent sign and statistical significance throughout all 5 specifications. VC backed firms seem to face a significant drop in R&D intensity after IPO as compared to the non-VC backed firms. It is interesting to note that all firms will experience a drop in the level of R&D after the IPO. This can be potentially due to asset expansion from the IPO outpacing firm's R&D investment opportunity.[49] While the positive sign of the VC dummy fits

[49] We do find that there may be a denominator effect (e.g., expansion in total assets) in which we saw a greater dip in capital expenditure after IPO for VC backed as well. However, the magnitude of such drop is relatively small (i.e., about 1.15%) and hence we believe that our results is not solely due to a systematic differential in the expansion of

consistently with the extant literature, the interaction term provides evidence for a greater decline in R&D intensity for the VC backed post-IPO. A direct interpretation of results in Table 3 is that VC firms are associated with higher R&D intensity on average but there is a systematic drop in R&D intensity for all after IPO and such drop is greater for VC backed firms. While interpretations might be different, our results are seemingly consistent with that of Darrough and Rangan (2005) who document a seemingly more extreme behavior of the insiders who manipulates earnings upwards by cutting R&D expenditures immediately after IPOs. Due to the limited sample size and periods in Darrough and Rangan (2005), they are unable to fully generalize the innovation behavior of the firms around IPO. In this paper, we use panel data in a multi-period DiD framework and explicitly compare VC and non-VC backed IPOs. We argue a more natural explanation is that firms do not fully inherit all the positive traits from the VCs who no longer actively engage in the monitoring and the promotion of innovation. Nonetheless, we do not fully eliminate the possibility that the decrease in R&D intensity is due to the deliberate short-termist behaviors (e.g., suppressing long-term projects) from the VCs when exits become imminent.[50] Another interesting observation from Table 3 is that higher cash flow from operation is negatively associated with R&D intensity. A potential explanation is that firms with higher cash flow may be an indication of relatively having invested in more short-term tangible projects which in turn translate into higher contemporary cash flows. Capital expenditures seem to be complementary with R&D investment.

---

total assets for the VC backed firms after IPO versus the non-VC backed. Not shown in the paper, we included a variable that indicates the magnitude of total asset expansion during IPO and our results are not changed.

[50] We do observe a slight bounce back of R&D intensity the year after IPO as shown in Fig 2, but it is far from having enough evidence to make the conclusion about any potential short-termist behavior of the VCs.

5.2 Parallel Assumptions

Results from Table 3 are to be interpreted with caveats if some of the key assumptions are not addressed properly. The difference-in-differences setup entails many assumptions and interpretation of causal relation is potentially invalid should any of the key assumptions fail. One of the most debated concerns is the parallel trend assumption where treatment and control groups should exhibit the same trend before the treatment and the trend will continue had treatment not occurred. A way to look at the parallel trend assumption is to look at changes in R&D intensity during the years before IPO and the results are shown in Table 4. For Column (1) and (2), we include all observations within the 5 years before IPO and empirically test if there exists a difference in time trend. Both columns include firm-level controls. Column (2) added industry fixed effects on top of (1). It appears that there is no statistically significant difference in R&D intensity trend before IPO between the VC and non-VC backed firms. Column (3) and (4) perform similar tests by interacting the individual *Year*(*t*) dummies (e.g., *Year(-1)* represent the year before IPO) with the VC dummies. We do not observe a significant difference in the change of R&D intensity before IPO between the treated and the control group. The above evidence suggests that there is no violation of the parallel trend assumptions, at least in the pre-event periods.

[Table 4 Here]

5.3 Placebo Event Tests

Table 5 exhibits the results of our placebo tests under which we hypothetically assume that VCs achieved confirmed exit at alternative dates. Column (1) assumes that the exit event occurs in the year before the actual IPO. Column (2) to (4) assume that VCs achieved exit 1 year after, 3 years before and 5 years after the actual IPO year correspondingly. The sample windows are shown in the parentheses below each column number.[51] Firm-level controls and all fixed effects are included as well. Under the DiD framework, the coefficients of the interaction term unanimously do not yield any statistical significance across all specifications. The larger decrease in R&D intensity for VC backed firms statistically disappear when hypothetical event dates were assumed. This again confirms our main findings that it is the IPO-exit that is affecting the R&D intensity.

[Table 5 Here]

5.4 Serial Correlation

Another concern regarding multiple periods difference-in-differences specification is serial correlations which can lead to serious over-estimation of *t*-statistics (Bertrand *et al*., 2004). One way to address this is to collapse the multi-period time series data into two effective periods, before and after IPO-exit. The same baseline model is tested but under collapsed regression settings where all variables were averaged within pre-specified sample periods before and after IPO into two collapsed periods. Industry and IPO year fixed effects were included for columns (3) and (6). Regardless of different specifications, we observe consistent evidence that VC backed firms still seem to experience a greater drop in R&D intensity at an economically significant level of potentially around 6% on average with the fixed effects and controls included. To further test the robustness of our results, we conducted the same empirical tests with different sampling windows. Results were shown in Table 7 with columns (1) to (3) including IPO year observations in the calculation of mean and (4) to (6) not. We observe consistent results both statistically and magnitude-wise regardless of sampling windows. The exit of VCs is associated with about 5%

[51] We use only sub samples and did not repeat our baseline regression using all sample by merely shifting one or two years around the actual IPOs because statistical significance of the interaction term may still show up if the real economic effect indeed exists and is dominating in the other time periods.

reduction in R&D intensity on average after controlling for firm-specific characteristics and fixed-effects.

[Table 6 and 7 Here]

5.5 Propensity Score Matching

Propensity score matching is another useful way to test for change in R&D intensity between the VC backed firms and the non-VC backed. We first calculate the propensity score of being selected as the treated (i.e., VC backed) using the logit model. The covariates include a combination of Fama-French 48 industry classification, State, IPO year, pre-IPO mean firm size, and pre-IPO mean R&D intensity. The pre-IPO means are calculated using available observations from the last 5 years before the IPO year. Controls (i.e., non-VC backed) are selected based on the nearest propensity score with replacement for each VC backed firm. We then compare the average change in the R&D intensity after IPO, which is the mean post-IPO R&D intensity within 6 years after IPO subtracting the mean pre-IPO R&D intensity (i.e., the time-series mean decrease in R&D intensity after IPO), between the treated and the control groups. Column (1) to (3) report the estimated average treatment effects with different sets of covariates in the first stage respectively. Column (4) to (6) report the same estimates but without including IPO year observations in calculating the means. On average, VC backed firms experienced approximately 6% more decline in R&D intensity when matched on the aforementioned dimensions. The coefficients from Table 8 largely concur with that of Table 3. One potential concern with all prior estimations is the problem of endogeneity. Propensity score matching will work ideally if we can select on firm characteristics before the initial financing stage, but such data is not publicly available. Firm characteristics that are available after the infusion of VCs already reflect the selection of the VCs and their subsequent influences. Column (1) and (4) hence represent the most exogenous matching since their location and industry classification are not influenced by VC intervention (Hochberg, 2011). Nonetheless, we also included results that match on additional observable characteristics as additional checks.

[Table 8 Here]

5.6 The Output of Innovation

We focus on the level of R&D expenditures in prior sections as a proxy for firms' willingness to innovate before and after IPO. Nonetheless, we further explore the effect on the outcome of R&D investment since there is certainly a strong correlation between the input and the output. Most studies tend to use patent counts or forward citations as a proxy for outputs of innovation. It should be noted that outputs from R&D are certainly influenced by the individual translation efficiency where more efficient firms may produce more patents of higher quality even if the investment to R&D is the same. One concern of using patent data as R&D intensity is with regards to the life cycle and the timing of patent filing. Patents are not an immediate product from R&D investment and it may take years before a firm eventually succeeds and manages to file it. It is therefore difficult to trace the exact amount of R&D expenses and time spent behind each filing. We do not expect immediate changes of large magnitude between the two groups after VCs exit in terms of R&D outputs. However, in the long-run, we may observe a greater drop in the output or quality of innovation for VC backed firms should our first hypothesis be true. Results in Table 9 confirm our conjecture. We measure innovation output as the natural log of patent count filed each year plus one and we do the same for the total yearly citation counts of each firm. We run the same baseline regression with various sample windows. Column (1) to (3) report the results for patent counts whereas (4) to (6) report the results for patent citations. In general, we see that not only does the coefficient of the interaction term become statistically more significant in longer windows, but the magnitude also increases. The result is consistent with our first hypothesis after considering the potential lag in the filing of patents that VC's exit seems to negatively affect a firm's investment in innovation which ultimately undermines its innovation output and quality in the future.

[Table 9 Here]

5.7 VC Involvement

Our second set of opposing hypotheses aims to test if VC has a lasting influence in the R&D investment decision even after exit. While it is clear from our above discussion that VCs' exit potentially withdraws positive influences on innovation that they brought into the firms at the initial stage, current literature does not shed much light on the interactive role of VC's involvement before exit. If VC involvement does not influence the future investment behavior of a firm at all, we will likely not observe any statistical significance in the coefficients of the three-way interaction term, β_5, in our alternative model (2). If VC involvement does play a role in a firm's future investment agenda, we may observe a statistically significant negative or positive sign which either corresponds to the dominance of either the withdrawal or persistency effect in the post-exit period. Table 10 reports the interaction effect of VC's pre-IPO involvement on R&D intensity. Column (1) is a triple DiD set up without firm controls and fixed-effects whereas column (2) and (3) include firm controls as well as a different set of fixed effects respectively. VC involvement is proxied by the relative level of capital VCs invested normalized by total assets the year before IPO compared to one's corresponding industry group median. We argue that

HIGHINVOLVE is a reasonably well proxy for the amount of effort VC is putting in since the incentive to achieve IPO is greater when VC owns a larger fraction of the firm.

[Table 10 Here]

It is affirmative of our first hypothesis that the interaction term between *VC* and *POST* still yields the expected sign with large statistical significance, indicating a negative impact on VC's exit on firm's investment in innovation. There is some evidence suggesting that VC involvement has a positive relation with the firm's incentive to do innovation. The key three-way interaction term exhibits a consistent negative sign with statistical significance even after including all fixed effects. This result is largely in favor of *H2a* that the negative impact on a firm's innovation is greater for VC backed firms after IPO-exit, and such negative effect is even stronger for firms with greater VC involvement. A plausible explanation is that the withdrawal effect is at least a more first-order factor in determining R&D intensity than any positive traits inherited from VC's involvement. Potentially, firms with higher VC involvement were driven to conduct more extensive innovation such that the magnitude of reversion in R&D investment is bigger when VCs are gone. From Table 10, firms with high pre-IPO VC involvement seem to at least experience a 3% greater drop than similar VC backed firms when compared to the non-VC backed.

[Table 11 Here]

5.8 Institutional Monitoring

Institutional investors are sophisticated and often take on monitoring roles due to having significant stakes in a firm. They represent a group of investors with different incentive horizons than the VCs after successful IPOs and their presence in some ways overlaps with the role of VCs.

Table 11 reports the impact of high institutional presence on firms' innovation around IPO-exit. By empirically testing model (3), we observe a heterogeneous effect from institutional monitoring on the exit of VCs which is captured by the three-way interaction term. Column (1) reports the results without firm-level controls and fixed effects. Column (2) and (3) both included controls and a set of fixed effects. Consistent with prior literature, the presence of high institutional holdings seems to be positively associated with innovation. While the DiD estimator yields robust results in favor of our *H1* across all 3 specifications, the three-way interaction terms in all three specifications yield positive signs with an average magnitude of about 5.6%. This result is consistent with our third hypothesis that the presence of institutional holdings will mitigate the negative impact on R&D investment to a significant extent when VC exits. Cadman and Sunder (2014) argue that VCs cannot anticipate subsequent institutional monitoring to take pre-emptive measures. Any differential effect from institutional monitoring between VC and non-VC backed firms is unlikely a direct result from the presence of VCs. In the presence of high institutional ownership, firms may also engage less in myopic behavior to achieve short term performance at the costs of long-term firm value, such as suppressing R&D investments that may be undervalued by the market due to information asymmetry. In essence, our empirical evidence does suggest that the institutional investors favor long-run investments such as R&D which may not yield immediate payoffs and alleviate potential short-termist behaviors.

6. Conclusion

This paper documents a negative impact on a firm's innovation when the VCs achieved an IPO-exit. Our results suggest that VCs indeed promotes innovation, however, firms will experience a drop in R&D when they leave, suggesting little to no inheritance of the positive influences from VCs. Plausibly, a firm's innovation may also be affected when there is a turnover of major stakeholders during an IPO. While extant literature documents mainly the evidence of a positive correlation between innovation and VC, we investigate the relation between VCs and innovation from a “reverse treatment” perspective as VCs exit. Our evidence concurs with the impression that VCs promote innovation and VC backed firms are on average more innovative both before and after IPO-exit. However, we do find evidence of withdrawal effect in innovation when VCs leave. This paper also fills the void in the literature by investigating whether VCs shape firms’ future innovation agenda after they exit. Our empirical evidence suggests that VCs play an important role in promoting firm innovation and this positive influence seems not entirely inherited by entrepreneurial firms. Our paper also complements the literature on institutional investors and finds that their presence may replace the monitoring role of the VCs after IPO and potentially reduces short-termist behaviors of the insiders that overweight current performance over future prospects. As many governments perceive VCs as the fueling agent for the economy and have specific policies and tax incentives in place to encourage VC financing, we hope to present some useful evidence on the impact of VCs during different phases of a firm and provide important implications for policies aimed at enhancing economic growth.

References

Aggarwal V. A., and D. Hsu. 2014. "Entrepreneurial Exits and Innovation." Management Science 60, 867-887.

Baker, M., and Paul A. Gompers, 2003. "The determinants of board structure at the initial public offering." Journal of Law and Economics 46, 569-598.

Bascha, A., and U. Walz. 2001. "Convertible Securities and Optimal Exit Decisions in Venture Capital Finance." Journal of Corporate Finance 7, 285-306.

Bergemann, D., & Hege, U. (1998). Venture capital financing, moral hazard, and learning. Journal of Banking & Finance, 22(6-8), 703-735.

Bernstein, Shai. 2015. "Does Going Public Affect Innovation?" Journal of Finance 70, 1365-1403.

Bertrand, M., E. Duflo, and S. Mullainathan. 2004. "How Much Should We Trust Differences-In-Differences Estimates?" The Quarterly Journal of Economics 119, 249-275.

Black, B. S., and R. J. Gilson. 1998. "Venture Capital and the Structure of Capital Markets: Banks versus Stock Markets." Journal of Financial Economics 47, 243-277.

Bolton, P., Scheinkman, J., & Xiong, W. (2006). Executive compensation and short-termist behaviour in speculative markets. The Review of Economic Studies, 73(3), 577-610.

Brander, J., R. Amit, and W. Antweiler. 2002. "Venture Capital Syndication: Improved venture Selection versus the Value-added Hypothesis." Journal of Economics and Management Strategy 11, 423-452.

Cadman, B., and Sunder, J. 2014. "Investor horizon and CEO horizon incentives." The Accounting Review 89(4), 1299-1328.

Caselli, S., Gatti, S., and F. Perrini. 2009. “Are Venture A Catalyst for Innovation or Do They Simply Exploit It?” European Financial Management Journal 15, 92-111.

Cornelli, F., and O. Yosha. 2003. “Stage Financing and the Role of Convertible Debt.” Review of Economic Studies 70, 1-32.

Cumming, D. J. 2006. “The Determinants of Venture Capital Portfolio Size: Empirical Evidence.” Journal of Business 79, 1083-1125.

Cumming, D. J. 2008. “Contracts and Exits in Venture Capital Finance.” The Review of Financial Studies 21, 1947-1982.

Cumming, D. J., and S. A. Johan, 2012. “Is Venture Capital in Crisis?” World Economic Review, 69-72.

Cumming, D. J., and S. A. Johan, 2008. “Information Asymmetries, Agency Costs and Venture Capital Exit Outcomes Venture Capital” An International Journal of Entrepreneurial Finance 10, 197-231.

Cumming, D. J., and D. Li. 2013. “Public Policy, Entrepreneurship, and Venture Capital in the United States.” Journal of Corporate Finance 23, 345-467.

Da Rin, Marco, Thomas Hellmann, and Manju Puri. 2013 "A survey of venture capital research." Handbook of the Economics of Finance. Vol. 2. Elsevier, 2013. 573-648.

Darrough, M., and Rangan, S. 2005. “Do insiders manipulate earnings when they sell their shares in an initial public offering?” Journal of Accounting Research 43(1), 1-33.

Dushnitsky, G., and M. J. Lenox. 2005. “When Do Firms Undertake R&D by Investing in New Ventures?” Strategic Management Journal 26, 947-965.

Engel, D. and M. Keilbach. 2007. "Firm-level implications of Early Stage Venture Capital Investment – An Empirical Investigation." Journal of Empirical Finance 14, 150-167.

Ertimur, Y., Sletten, E., and Sunder, J. 2014. "Large shareholders and disclosure strategies: Evidence from IPO lockup expirations." Journal of Accounting and Economics 1(58), 79-95.

Giot, P., and A. Schwienbacher. 2007. "IPOs, Trade Sales and Liquidations: Modeling Venture Capital Exits Using Survival Analysis." Journal of Banking and Finance 31, 679-702.

Gompers, P. A. 1995. "Optimal Investment, Monitoring, and the Staging of Venture Capital." Journal of Finance 50, 1561-1489.

Gompers, P. A. 1999. "Ownership and Control in Entrepreneurial Firms: An Examination of Convertible Securities in Venture Capital Investments." Harvard, Working Paper.

Gompers, P. A., and J. Lerner. 1999. "What Drives Venture Capital Fund Raising." Brookings papers on economic activity, 149-192.

Hall, H., Griliches, Z., and Hausman, J.A. 1986. "Patents and R&D: Is There a Lag?" International Economic Review 27, 265-283.

Hall, H. Bronwyn. 2002. "The Financing of Research and Development." Oxford Review of Economics 18, 35-51.

Hellmann, T. (2006). IPOs, acquisitions, and the use of convertible securities in venture capital. Journal of Financial Economics, 81(3), 649-679.

Hellmann, T., Puri, M. 2000. "The Interaction Between Product Market and Financing Strategy: the role of Venture Capital." Review of Financial Studies 13, 959-984.

Hirukawa, M., and M. Ueda. 2011. “Venture Capital and Innovation: Which Is First?” Pacific Economic Review 16, 421-465.

Hochberg, Y. V. 2011. “Venture capital and corporate governance in the newly public firm.” Review of Finance 16(2), 429-480.

Kaplan, S. N., and P. Stromberg. 2003 “Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts.” Review of Economics Studies 70, 281-315.

Kanniainen, V., and C. Keuschnigg. 2003. “The Optimal Portfolio of Start-Up Firms in Venture Capital Finance.” Journal of Corporate Finance 9, 521-534.

Kanniainen, V., and C. Keuschnigg. 2004. “Start-Up Investment with Scarce Venture Capital Support.” Journal of Banking and Finance 28, 1935-1959.

Keuschnigg, C. 2004. “Taxation of a Venture Capitalist with a Portfolio of Firms.” Oxford Economic Papers 56, 285-306.

Kochhar, R., & David, P. (1996). Institutional investors and firm innovation: A test of competing hypotheses. Strategic management journal, 17(1), 73-84.

Kortum, S. and J. Lerner. 2000. “Assessing the Contribution of Venture Capital to Innovation.” RAND Journal of Economics 31, 674-692.

Lerner, J. 1994. “The Syndication of Venture Capital Investments.” Financial Management 23, 16-27.

Puri, M., & Zarutskie, R. (2012). On the life cycle dynamics of venture-capital-and non-venture-capital-financed firms. The Journal of Finance, 67(6), 2247-2293.

Schwienbacher, A. 2008. “Innovation and Venture Capital Exits.” The Economic Journal 118, 1888-1916.

Fig 1. Mean R&D Intensity Pre-IPO vs Post-IPO

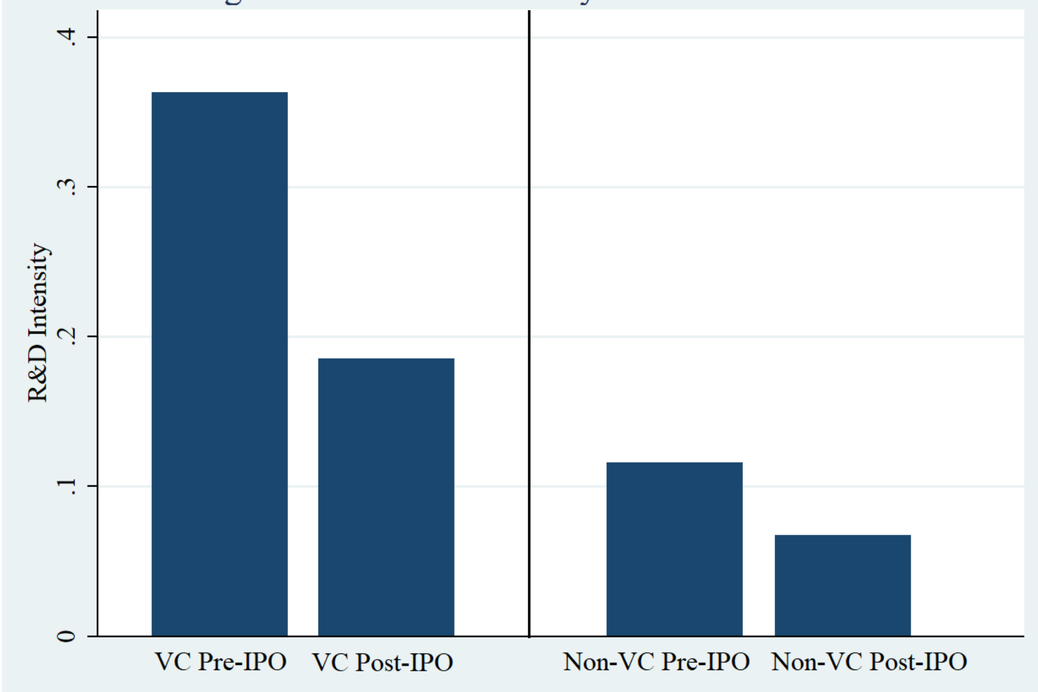


Fig 2. R&D Intensity around IPO

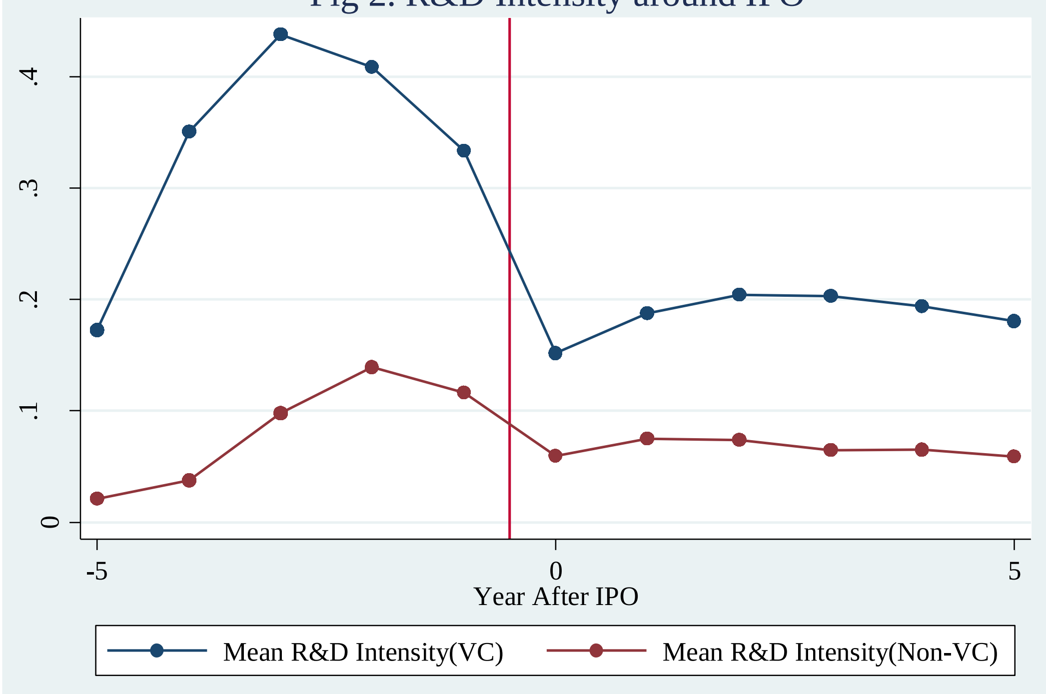


Table 1. Fama-French Industry Distribution

Panel A

Industry	**Non-VC**	**VC**	All
Food Products	36	16	52
Beer & Liquor	9	4	13
Tobacco Products	4	0	4
Recreation	76	18	94
Printing and Publishing	22	6	28
Consumer Goods	41	4	45
Apparel	40	6	46
Healthcare, Medical Equipment, Pharmaceutical Products	163	470	633
Chemicals	29	6	35
Textiles	13	2	15
Construction and Construction Materials	68	4	72
Steel Works Etc	33	2	35
Fabricated Products and Machinery	75	24	99
Electrical Equipment	23	9	32
Automobiles and Truck	25	3	28
Aircraft, Ships, and Railroad Equipment	10	2	12
Precious Metals, Non-Metallic, and Industrial Metal Mining	10	1	11
Coal	9	0	9
Petroleum and Natural Gas	84	17	101
Personal and Business Services	303	524	827
Business Equipment	194	324	518
Business Supplies and Shipping Containers	18	4	22
Transportation	86	15	101
Wholesale	96	17	113
Retail	165	73	238
Restaurants, Hotels, Motels	58	17	75
Others	55	11	66
Total	1,745	1,579	3,324

Panel B

Industry	R&D Intensity			
	VC		Non-VC	
	Mean	Sd	Mean	Sd
Consumer (Non-Durable)	0.10	0.13	0.029	0.038
Consumer (Durable)	0.12	0.11	0.026	0.019
Manufacturing	0.10	0.09	0.037	0.066
Energy	0.09	0.02	0.007	0.010
Chemicals	0.11	0.11	0.032	0.041
Business Equipment	0.13	0.10	0.096	0.079
Wholesale	0.01	0.03	0.001	0.012
Healthcare	0.23	0.17	0.153	0.212
Other	0.09	0.09	0.022	0.049

Table 2. Descriptive Statistics

The table reports the summary statistics of the main variables used in this paper. Panel A reports the number of observations, *N*, mean, stand deviations, and the value of the variable at various percentile. Panel B reports the results for *t*-test on the difference of mean between VC and non-VC backed firms and Panel C reports the difference before and after IPO. *, **, or *** indicates a significance level at 0.10, 0.05, or 0.01 respectively. For a detailed definition of variables, please refer to Appendix A.

Panel a. Full sample

Variables	N	Mean	Sd	Min	.25	Mdn	.75	Max
R&D INTENSITY	22890	0.1412	0.2055	0.0000	0.0083	0.0765	0.1776	1.2593
ROA	33422	-0.0763	0.3438	-2.0610	-0.0821	0.0287	0.0796	0.3188
LEVERAGE	33469	1.2134	3.8806	-15.0565	0.2696	0.6706	1.4856	24.5030
CAPEX	33094	0.0644	0.0726	0.0002	0.0194	0.0405	0.0793	0.4001
SIZE	33568	5.1631	1.8215	0.7381	3.8931	5.1000	6.4154	9.5254
CFO	31862	-0.0025	0.2716	-1.4407	-0.0309	0.0660	0.1302	0.4059
HIGHINVOLVE	37844	0.6949	0.4605	0.0000	0.0000	1.0000	1.0000	1.0000
PATENTS	25062	0.3898	0.8174	0.0000	0.0000	0.0000	0.0000	3.8712
CITATIONS	23936	0.7300	1.5361	0.0000	0.0000	0.0000	0.0000	6.1506
HIGHINST	37844	0.3974	0.4894	0.0000	0.0000	0.0000	1.0000	1.0000

Panel b. VC vs. non-VC backed

	VC			Non-VC			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(5) - (2)**
	N	*Mean*	*Sd*	*N*	*Mean*	*Sd*	*T-test*
R&D INTENSITY	12298	0.2066	0.2286	10592	0.0652	0.1408	-0.141***
ROA	13953	-0.1729	0.4110	19469	-0.0071	0.2653	0.166***
LEVERAGE	13930	0.7879	3.1786	19539	1.5167	4.2862	0.729***
CAPEX	13842	0.0557	0.0649	19252	0.0706	0.0771	0.015***
SIZE	13992	4.7690	1.7066	19576	5.4448	1.8487	0.676***
CFO	13696	-0.0776	0.3286	18166	0.0541	0.2013	0.132***
HIGHINVOLVE	16085	0.5867	0.4924	21759	0.7749	0.4177	0.188***
PATENTS	10167	0.5648	0.9167	14895	0.2703	0.7179	-0.295***
CITATIONS	9674	1.0922	1.7933	14262	0.4844	1.2764	-0.610***
HIGHINST	16085	0.4357	0.4959	21759	0.3692	0.4826	-0.067***

Panel c. Pre-IPO vs. Post-IPO

	Pre-IPO			Post-IPO			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(5) - (2)**
	N	*Mean*	*Sd*	*N*	*Mean*	*Sd*	*T-test*
R&D INTENSITY	4158	0.2399	0.3108	18732	0.1192	0.1659	-0.121***
ROA	6314	-0.1760	0.4998	27108	-0.0531	0.2910	0.123***
LEVERAGE	6388	1.5291	5.7036	27081	1.1389	3.3030	-0.390***
CAPEX	6228	0.0803	0.0850	26866	0.0607	0.0689	-0.020***
SIZE	6404	4.0500	1.8948	27164	5.4255	1.7008	1.376***
CFO	5210	-0.1137	0.4215	26652	0.0192	0.2249	0.133***
HIGHINVOLVE	10637	0.6500	0.4770	27207	0.7125	0.4526	0.063***
PATENTS	8440	0.2658	0.6706	16622	0.4527	0.8759	0.187***
CITATIONS	8257	0.5973	1.4556	15679	0.7999	1.5723	0.202***
HIGHINST	10637	0.5677	0.4954	27207	0.3309	0.4705	-0.237***

Table 3. IPO-Exit

This table reports the impact of VC's IPO-exit on R&D intensity. Column (1) to (3) did not include firm-level controls whereas (4) and (5) did. All continuous variables are winsorized at 1% at both tails. All standard errors were clustered by firms and reported in the parentheses. Various important fixed effects are included in different specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For a detailed definition of variables, please refer to Appendix A.

R&D INTENSITY	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
VC	0.2687***	0.1644***		0.0759***	
	(0.012)	(0.011)		(0.012)	
POST	-0.0343***	-0.0384***		-0.0425***	
	(0.007)	(0.007)		(0.007)	
*POST * VC*	-0.1563***	-0.1274***	-0.1130***	-0.0456***	-0.0593***
	(0.011)	(0.010)	(0.009)	(0.011)	(0.010)
ROA				-0.0067	0.0056
				(0.006)	(0.006)
LEVERAGE				-0.0006**	-0.0000
				(0.000)	(0.000)
CAPEX				0.0869***	0.1883***
				(0.029)	(0.026)
SIZE				-0.0067***	-0.0136***
				(0.002)	(0.002)
CFO				-0.3668***	-0.3548***
				(0.013)	(0.014)
CONS	0.0934***	0.1401***	0.1909***	0.1726***	0.2121***
	(0.007)	(0.008)	(0.004)	(0.013)	(0.012)
Firm FEs	No	No	Yes	No	Yes
Industry FEs	No	Yes	Yes	Yes	Yes
Year FEs	No	Yes	Yes	Yes	Yes
Year After IPO FEs	No	No	Yes	No	Yes
IPO Year FEs	No	Yes	Yes	Yes	Yes
Adjusted R^2	0.19	0.38	0.64	0.61	0.78
Observations	22890	22889	22828	19752	19602

Table 4. Parallel Trend

This table shows the result of tests on the parallel trend assumption before the IPO-exit. Column (1) and (2) include all samples within 5 years prior to IPO. Column (3) and (4) include all samples 4 years prior to IPO. Firm-level controls are included for all specifications. All continuous variables are winsorized at 1% at both tails. All standard errors are clustered by firms and reported in the parentheses. Various fixed effects are included in different specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, or 0.01 respectively. For a detailed definition of variables, please refer to Appendix A.

Δ R&D Intensity	**(1)**	**(2)**	**(3)**	**(4)**
*TREND * VC*	0.0015	0.0176		
	(0.034)	(0.035)		
TREND	0.2075	0.3618		
	(0.606)	(0.622)		
VC	-0.0225	-0.0529	1.0511	1.1741
	(0.124)	(0.127)	(1.141)	(1.137)
*VC * YEAR(-3)*			-1.2040	-1.0887
			(1.189)	(1.122)
*VC * YEAR(-2)*			-0.9586	-1.0481
			(1.108)	(1.093)
*VC * YEAR(-1)*			-0.9633	-1.0846
			(1.148)	(1.141)
YEAR(-3)			-0.1119	-0.0701
			(0.113)	(0.111)
YEAR(-2)			-0.1338	-0.0681
			(0.089)	(0.064)
YEAR(-1)			-0.0626	0.0390
			(0.112)	(0.122)
ROA	0.2945***	0.3264***	0.2951***	0.3236***
	(0.069)	(0.071)	(0.075)	(0.071)
LEVERAGE	-0.0085**	-0.0064	-0.0088**	-0.0066
	(0.004)	(0.004)	(0.003)	(0.004)
CAPEX	1.4733***	1.5675***	1.5153***	1.6167***
	(0.530)	(0.525)	(0.378)	(0.427)
SIZE	-0.1748***	-0.1757***	-0.1782***	-0.1773***
	(0.024)	(0.025)	(0.025)	(0.028)
CFO	0.0282	0.0174	0.0329	0.0227
	(0.082)	(0.085)	(0.098)	(0.102)
CONS	1.2174***	1.1469***	1.3113***	1.2133***
	(0.227)	(0.218)	(0.154)	(0.170)
Industry FEs	No	Yes	No	Yes
IPO Year FEs	Yes	Yes	Yes	Yes
Adjusted R^2	0.13	0.16	0.13	0.16
Observations	1207	1200	1194	1188

Table 5. Placebo Test

This table reports the results of placebo tests with pseudo-event dates. Column (1) assumes that VCs achieve confirmed exit a year before the actual IPO. Column (2) to (4) assume exit occurs 1 year after, 3 years before and 5 years after the actual IPO year. Sample windows are shown right below each column number. Firm-level controls are included for all specifications. All continuous variables are winsorized at 1% at both tails. All standard errors are clustered by firms and reported in the parentheses. Several fixed effects are also included in the regressions. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, or 0.01 respectively. For a detailed definition of variables, please refer to Appendix A.

R&D INTENSITY	**(1)** *(-2,-1)*	**(2)** *(0,+1)*	**(3)** *(-5,-1)*	**(4)** *(0,+10)*
*POST * VC*	0.0083	0.0083	-0.0474	-0.0021
	(0.033)	(0.006)	(0.055)	(0.005)
ROA	0.0614	0.0409***	0.0468	0.0097
	(0.088)	(0.010)	(0.045)	(0.007)
LEVERAGE	0.0010	0.0001	0.0002	0.0002
	(0.001)	(0.000)	(0.001)	(0.000)
CAPEX	0.2370	0.1833***	0.5882	0.1483***
	(0.322)	(0.052)	(0.366)	(0.024)
SIZE	-0.0039	0.0067*	-0.0091	-0.0135***
	(0.058)	(0.004)	(0.020)	(0.003)
CFO	-0.5479***	-0.3347***	-0.5516***	-0.3313***
	(0.059)	(0.030)	(0.056)	(0.018)
CONS	0.1816	0.0780***	0.1809*	0.1772***
	(0.284)	(0.016)	(0.101)	(0.014)
Firm FEs	Yes	Yes	Yes	Yes
Industry FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Year After IPO FEs	Yes	Yes	Yes	Yes
IPO Year FEs	Yes	Yes	Yes	Yes
Adjusted R^2	0.80	0.81	0.88	0.78
Observations	278	3600	380	13407

Table 6. Collapsed Regressions

This table reports the impact of VC's IPO-exit on R&D intensity under collapsed regression settings where all variables were averaged within 5 years before and 6 years after IPO. Column (1) to (3) included IPO year observations in the calculation of means whereas (4) to (6) did not. All continuous variables are winsorized at 1% at both tails. All standard errors were clustered by firms and reported in the parentheses. Industry and IPO year fixed effects were included in columns (3) and (6). Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, or 0.01 respectively. For a detailed definition of variables, please refer to Appendix A.

R&D INTENSITY	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
VC	0.2502***	0.1070***	0.0724***	0.2534***	0.1060***	0.0690***
	(0.014)	(0.011)	(0.010)	(0.014)	(0.011)	(0.011)
POST	-0.0551***	-0.0305***	-0.0358***	-0.0509***	-0.0304***	-0.0363***
	(0.008)	(0.006)	(0.005)	(0.008)	(0.006)	(0.006)
*POST * VC*	-0.1344***	-0.0652***	-0.0672***	-0.1286***	-0.0598***	-0.0626***
	(0.012)	(0.009)	(0.009)	(0.012)	(0.010)	(0.009)
ROA		-0.3590***	-0.3444***		-0.3663***	-0.3521***
		(0.023)	(0.021)		(0.024)	(0.022)
LEVERAGE		-0.0008	-0.0008*		-0.0007	-0.0007
		(0.001)	(0.000)		(0.001)	(0.000)
CAPEX		-0.2625***	0.0173		-0.2697***	0.0187
		(0.053)	(0.055)		(0.056)	(0.057)
SIZE		-0.0190***	-0.0062***		-0.0203***	-0.0072***
		(0.002)	(0.002)		(0.002)	(0.002)
CFO		-0.0007	-0.0032		-0.0016	-0.0041
		(0.020)	(0.016)		(0.021)	(0.018)
CONS	0.1278***	0.1886***	0.1483***	0.1302***	0.1948***	0.1542***
	(0.009)	(0.012)	(0.012)	(0.010)	(0.013)	(0.013)
Industry FEs	No	No	Yes	No	No	Yes
IPO Year FEs	No	No	Yes	No	No	Yes
Adjusted R^2	0.17	0.64	0.70	0.15	0.63	0.69
Observations	4323	4108	4085	4310	4087	4065

Table 7. Robustness Tests

This table reports the impact of VC's IPO-exit on R&D intensity for alternative sample periods. Column (1) to (3) included IPO year observations whereas (4) to (6) did not. Firm-level controls include *ROA*, *LEVERAGE*, *CAPEX*, *SIZE*, and *CFO*. All continuous variables are winsorized at 1% at both tails. All standard errors were clustered by firms and reported in the parentheses. All fixed effects were included in all specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, or 0.01 respectively. For a detailed definition of variables, please refer to Appendix A.

R&D INTENSITY	**(1)** (-1,+1)	**(2)** (-5,+5)	**(3)** (-5,+10)	**(4)** (-1,+1)	**(5)** (-5,+5)	**(5)** (-5,+10)
*POST * VC*	-0.0599***	-0.0577***	-0.0597***	-0.0360***	-0.0454***	-0.0498***
	(0.010)	(0.010)	(0.010)	(0.014)	(0.011)	(0.011)
Firm-Level Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year After IPO FEs	Yes	Yes	Yes	Yes	Yes	Yes
IPO Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.78	0.80	0.78	0.77	0.81	0.80
Observations	4855	11003	14882	1926	8828	12705

Table 8. Propensity Score Matching

This table reports the average treatment effects on the time-series mean change in R&D intensity between VC backed and non-VC backed firms after IPO using one to one propensity score matching. Logit model is used to calculate propensity score and the covariates for different specifications include Fama-French 48 industry classification, IPO year, pre-IPO mean firm size, and pre-IPO mean R&D intensity. Each non-VC control is matched with replacement. Means were calculated within 5 years before IPO and 6 years after IPO. Column (1) to (3) includes IPO year observation and column (4) to (6) do not. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, or 0.01 respectively.

Δ R&D Intensity	(1)	(2)	(3)	(4)	(5)	(6)
Average Treatment Effect						
VC vs. Non-VC	-0.0644***	-0.0850***	-0.0208*	-0.0659***	-0.0855***	-0.0251*
	(0.018)	(0.016)	(0.011)	(0.020)	(0.018)	(0.014)
Matched on:						
Industry	Yes	Yes	Yes	Yes	Yes	Yes
State	Yes	Yes	Yes	Yes	Yes	Yes
IPO year	No	Yes	Yes	No	Yes	Yes
Pre-IPO Size	No	No	Yes	No	No	Yes
Pre-IPO R&D Intensity	No	No	Yes	No	No	Yes
Observations	2115	2115	2115	1905	1905	1905

Table 9. Patents and Citations

This table shows the impact of VC's IPO-exit on patent counts and patent citations. Column (1) to (3) report the results for patent counts whereas (4) to (6) report the results for patent citations. All continuous variables are winsorized at 1% at both tails. All standard errors are clustered by firm and reported in the parentheses. All fixed effects are included in all specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, or 0.01 respectively. For a detailed definition of variables, please refer to Appendix A.

	PATENTS			**CITATIONS**		
	(1) *(-5,+5)*	**(2)** *(-5,+10)*	**(3)** *All Sample*	**(4)** *(-5,+5)*	**(5)** *(-5,+10)*	**(6)** *All sample*
*POST * VC*	-0.0686*	-0.1090***	-0.1086***	-0.1670**	-0.2440***	-0.2349***
	(0.038)	(0.039)	(0.039)	(0.084)	(0.087)	(0.088)
ROA	-0.0353	-0.0469**	-0.0321	-0.0966**	-0.0953**	-0.0718
	(0.023)	(0.021)	(0.022)	(0.049)	(0.046)	(0.047)
LEVERAGE	0.0005	0.0001	-0.0003	-0.0010	-0.0010	-0.0023
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
CAPEX	0.2010**	0.1948**	0.2409***	0.1999	0.0126	0.1661
	(0.088)	(0.088)	(0.091)	(0.196)	(0.186)	(0.194)
SIZE	0.0557***	0.0641***	0.0550***	0.0534**	0.0316	0.0115
	(0.012)	(0.012)	(0.012)	(0.025)	(0.024)	(0.024)
CFO	0.0280	0.0405	0.0582*	0.0797	0.1429**	0.1725**
	(0.033)	(0.031)	(0.033)	(0.074)	(0.069)	(0.075)
CONS	0.1920***	0.1828***	0.2200***	0.6574***	0.7664***	0.8221***
	(0.056)	(0.056)	(0.059)	(0.110)	(0.110)	(0.115)
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year After IPO FEs	Yes	Yes	Yes	Yes	Yes	Yes
IPO Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.76	0.74	0.72	0.71	0.69	0.65
Observations	11074	14932	17010	10663	14266	15984

Table 10. VC Involvement

This table reports the interaction effect of VC's pre-IPO involvement on R&D intensity upon VC's exit. Column (1) does not include firm-level controls and fixed effects whereas (2) and (3) do. All continuous variables are winsorized at 1% at both tails. All standard errors are clustered by firm and reported in the parentheses. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, or 0.01 respectively. For a detailed definition of variables, please refer to Appendix A.

R&D INTENSITY	**(1)**	**(2)**	**(3)**
*POST * VC * HIGHINVOLVE*	-0.0724***	-0.0385**	-0.0296*
	(0.019)	(0.019)	(0.016)
HIGHINVOLVE	0.0876***	0.0184	0.0124
	(0.011)	(0.013)	(0.009)
VC	0.2389***	0.0642***	
	(0.012)	(0.013)	
POST	-0.0062	-0.0171*	
	(0.006)	(0.009)	
*POST * VC*	-0.1094***	-0.0358***	-0.0445***
	(0.012)	(0.013)	(0.011)
*POST * HIGHINVOLVE*	-0.0512***	-0.0165	-0.0246***
	(0.010)	(0.013)	(0.008)
*VC * HIGHINVOLVE*	0.0602***	0.0415**	0.0460**
	(0.020)	(0.019)	(0.020)
ROA		-0.0093	0.0058
		(0.006)	(0.006)
LEVERAGE		-0.0007***	-0.0000
		(0.000)	(0.000)
CAPEX		0.0531*	0.1856***
		(0.028)	(0.026)
SIZE		-0.0033***	-0.0137***
		(0.001)	(0.002)
CFO		-0.3621***	-0.3510***
		(0.013)	(0.014)
CONS	0.0339***	0.1314***	0.2069***
	(0.006)	(0.012)	(0.013)
Firm FEs	No	No	Yes
Industry FEs	No	Yes	Yes
Year FEs	No	No	Yes
Year After IPO FEs	No	No	Yes
IPO Year FEs	No	Yes	Yes
Adjusted R^2	0.21	0.61	0.78
Observations	22890	19752	19602

Table 11. Institutional Monitoring

This table reports the interaction effect of institutional monitoring on R&D intensity upon VC's exit. Column (1) does not include firm-level controls and fixed effects whereas (2) and (3) do. All continuous variables are winsorized at 1% at both tails. All standard errors are clustered by firms and reported in the parentheses. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, or 0.01 respectively. For a detailed definition of variables, please refer to Appendix A.

R&D INTENSITY	**(1)**	**(2)**	**(3)**
*POST * VC * HIGHINST*	0.0625**	0.0554**	0.0489**
	(0.025)	(0.026)	(0.024)
HIGHINST	0.0649***	0.0362***	0.0132*
	(0.012)	(0.010)	(0.008)
VC	0.3184***	0.1405***	
	(0.020)	(0.024)	
POST	0.0067	-0.0057	
	(0.008)	(0.007)	
*POST * VC*	-0.1977***	-0.1071***	-0.1057***
	(0.019)	(0.023)	(0.023)
*POST * HIGHINST*	-0.0903***	-0.0305***	-0.0162**
	(0.014)	(0.011)	(0.008)
*VC * HIGHINST*	-0.0867***	-0.0671***	-0.0552**
	(0.023)	(0.025)	(0.024)
ROA		-0.0098	0.0051
		(0.006)	(0.006)
LEVERAGE		-0.0007***	-0.0000
		(0.000)	(0.000)
CAPEX		0.0556**	0.1912***
		(0.028)	(0.025)
SIZE		-0.0033***	-0.0127***
		(0.001)	(0.002)
CFO		-0.3638***	-0.3515***
		(0.013)	(0.014)
CONS	0.0606***	0.1196***	0.2344***
	(0.008)	(0.010)	(0.016)
Firm FEs	No	No	Yes
Industry FEs	No	Yes	Yes
Year FEs	No	No	Yes
Year After IPO FEs	No	No	Yes
IPO Year FEs	No	Yes	Yes
Adjusted R^2	0.20	0.61	0.78
Observations	22894	19761	19609

Appendix A. Variable Definitions

Variable	Descriptions
VC	A dummy that equals to 1 if a firm is VC backed
HIGHINVOLVE	A dummy equal to one if the total known amount of capital invested in the firm by the VCs after normalized by the total asset of the year before IPO is higher than its corresponding the 2-digit SIC industry median.
HIGHINST	A dummy equal to one if the average quarterly mean institutional ownership as disclosed by Form 13F is higher than its corresponding 2-digit SIC industry median.
POST	A dummy that equals to 1 if the fiscal year is after IPO.
ROA	Net income divided by total assets.
LEVERAGE	Total debt over total common equity.
CAPEX	Total capital expenditures normalized by total assets.
SIZE	Natural log of total assets.
TREND	A time trend variable for the 5 years before IPO.
YEAR(t)	A dummy equal to 1 if it is *t* year after IPO.
CFO	Operating cash flow normalized by total assets.
PATENT	Natural log of 1 + total number of patents filed in a year.
CITATION	Natural log of 1 + total number of patent citations in a year.

Chapter 4. Is Import Competition Good for Domestic Innovation?

1. Introduction

This paper examines the long-debated relationship between product market competition and firm innovation. On the one hand, the view that competition hinders innovation can trace back to an early view proposed by Schumpeter (1942) that the presence of monopoly rents generates incentives for innovation whereas perfect competition is not optimal for innovations (i.e., Schumpeter effect). On the other hand, Arrow (1962) holds the alternative view that competition promotes innovation because monopolist has a weaker incentive to innovate since by innovating the monopolist merely replaces itself (i.e., Arrow effect). More recent theories have taken a middle ground to try reconciling the mixed arguments and arrived at a non-linear relationship.[52] It is no surprise that the prediction of theoretical models relies heavily on their assumptions. On the empirical side, it is equally challenging to derive a causal relation between competition and innovation due to the absence of a perfect proxy for competition and the innate endogeneity problem of such measures since an increase in innovation per se affects the competitive landscape of the market. While Gilbert (2006) concludes that current literature remains far from a general theory of innovation competition, our paper revisits this much-debated relation empirically by not only directly documenting evidence of causal inference between product market competition and innovation but also by providing insights on how this relationship might change under different business and market settings. Interestingly, some of our results seem to reconcile the two opposing theories.

[52] See (Aghion *et al.*, 2005; Schmidt, 1997; Aghion and Griffith, 2008; Hashmi, 2013).

Traditional market competition measures such as the Herfindahl index (*HHI*) and Concentration Ratio (*CR*) are calculated at the industry level. It is likely that firms, though in the same industry, face different levels of threats depending on their unique competitive position. Applying a generalized measure across all firms in the same market may eliminate important variations. As documented in Gilbert (2006), prior empirical studies do not arrive at a robust conclusion using such industry-level measures. This paper revisits the controversial relation between market competition and firm innovation using two newly developed firm-specific measures of competition threat which are constructed using computational linguistics over a large sample of Form 10-Ks. We first examine the association between competitive threat and various innovation measures using pooled Ordinary Least Square (OLS) regression. We find a consistent positive relation between competitive threat and the firm's innovation activities. Our results bolster the argument that competition may force firms to minimize costs and to develop new products through innovation in order to escape the competition. A one-standard-deviation increase in firm-level competition threat faced by an average firm increases the number of patent applications (R&D intensity) by 9.3% (2.7%). We address the endogeneity concerns related to potential omitted variable bias or reverse causality using Instrumental Variables (IV) estimation as well as by identifying sudden reductions in import tariffs as exogenous shocks in competition. The results from IV regressions and propensity score matching are consistent with our OLS results.

We then explore the interactive effects of competition under different business settings. We find that the positive effect of competition on innovation is more prominent in firms with higher past profitability. Interestingly, we observe a negative relation between past profitability and innovation suggesting that incumbent firms that were already capturing a good portion of the profits in the past on average are less innovative. The above pair of results seemingly reconcile somehow the

Schumpeter effect and the Arrow effect – Competition that reduces pre-innovation profit spurs innovation because firms have incentives to retain their current profit level, but on the other hand, firms with significant sales in past are relatively less innovative, ceteris paribus.

Our study also documents the interactive effect of industry outlook on competition. Using the average past five-years industry growth as a proxy for future industry growth, we find support that a better industry-level prospect seems to encourage more innovation. However, such a positive effect is less prominent in the presence of greater competition. While competition in general spurs innovation, the positive effect from good industry prospects is moderated under competition, which indicates that any potential sharing of future profits due to greater competition may discourage innovation. Our paper also provides evidence that the positive effect on innovation from heightened competition is more pronounced for firms with higher financial constraints. Intuitively, firms that lose competitive edges in the innovation race may eventually go out of business. As argued by Aghion *et al.* (1999), managers innovate more in competitive markets because the bankruptcy risk is lower should they manage to escape competition and establish monopolistic power. We also provide insights on how the current inventory of technology affects the competitive landscape of an industry. Firms owning different levels of technology will face different levels of competitive threats looking forward which affects their incentives to innovate. In the spirit of Aghion *et al.* (2005), we argue that firms having a large inventory of patents have a higher probability of escaping the competition through innovation whereas the laggard firms have less incentives to invest in innovation without foreseeing a promising outcome. We also observe some evidence that product market competition reduces innovation efficiency which potentially points to greater resources wasting or inefficient resource allocation upon heightened competition. Nonetheless, our main results show that competition increases both the quantity and quality of innovation in general.

2. Review of Literature

The extant literature has not arrived at a full consensus on the effect of competition on innovation. The debate traces back to the theory advanced by Schumpeter (1942) that the presence of monopoly rents generates incentives for innovation whereas perfect competition is not the optimal market structure for innovation. The discussion between competition and innovation intensified when Arrow (1962) proposes that competition rather than monopoly encourages innovation. The main argument of the Arrow (or replacement effect) is that the monopolist does not gain much from additional innovation as it already captures most of the market, whereas a competitor has no pre-existing profit to replace. After the initial spark on the debate, many theoretical studies either draw a similar conclusion as Schumpeter (e.g., Gilbert and Newbery 1982; Greenstein and Ramey, 1998; Chen and Schwartz, 2013) or Arrow (e.g., Reinganum, 1983; Weinberg, 1992). Following the prior theoretical foundations, a new strand of literature emerged in an attempt to reconcile the two opposing conclusions and arrive at non-linear (e.g., inverted-U shape) relationships (e.g., Schmidt, 1997; Boone, 2001; Aghion *et al.*, 2005; Aghion and Griffith, 2008; Hashmi, 2013). However, no unambiguous conclusion can be drawn as the theoretical relation between competition and innovation is yet confounded by the complexity in the market structures, characteristics of innovation, and the dynamics of discovery (Kamien and Schwartz, 1975 and Gilbert, 2006).[53]

[53] See Kamien and Schwartz (1975) and Gilbert (2006) for comprehensive surveys about the theories and empirical evidence on the relation between market competition and innovation.

While no robust relationship exists between competition and innovation from the theory side, empiricists have been striving equally hard to find real-world evidence but provide no clearer answer. For instance, Scherer (1965) finds no correlation between concentration ratio and R&D intensity, whereas Mansfield *et al*. (1977) document only some evidence of a positive correlation between market concentration and R&D expenditure but none when concentration is above moderate levels. Angelmar (1985), however, concludes that such a positive correlation only stands in industries with low barriers to imitations and becomes negative in industries with high barriers to imitation. A survey study conducted by Gilbert (2006) has provided a broad list of earlier empirical studies that find positive, negative, no, or mixed relation between competition and innovation. Inevitably, prior empirical research usually takes the form of examining cross-section partial correlations between market structure and innovation, and such studies are deemed uninformative since they do not control adequately for technological opportunities that vary across industries and are correlated with traditional industry-level competition measures (Nickell, 1996).

The most concerning caveat in the interpretation of prior empirical studies is the endogeneity between innovation and competition. The confounding relation is complex and reverse causality is highly likely because conducting innovation per se changes the competitive landscape of a market. The simultaneous effects of competition and innovation make causal inferences on the effect of competition difficult. Xu (2012) introduced the use of tariff rates and foreign exchange rates as IVs for import competition and find that competition significantly reduces expected profits and firm leverage. We follow the approach in Xu (2012) by addressing the endogeneity problem using both import tariffs and foreign exchange rates as IVs for product market competition.[54]

[54] Li and Zhan (2018) follow a similar strategy to study the impact of competition on stock crash risk.

Another challenge confronting the empiricists is the construction of a satisfactory measurement for market competition. The traditional competition measures such as *HHI* or the *CR* of the largest firms are embedded with several limitations. These measures are calculated using historical data and at the industry level. The competitive threat faced by a firm also may not fully be accounted for since only sales data of public firms are available most of the time. Scherer (1984) reiterates that conclusions drawn from using traditional industry-level measures of competition on innovation may be merely an artifact of inadequate controls for differences among firms and industries in opportunities for R&D.

Borrowing from computational linguistics, Hoberg *et al.* (2014) construct a new measure of product market threats at firm-level through the textual analysis a large sample of 10-K filing (i.e., *Fluidity*). They find that increasing product market competition reduces the likelihood of dividend payouts or share repurchases and that such firms hold more cash. Li *et al*. (2013) also develop a simple but novel measure of competition using 10-K filings (i.e., *Pctcomp*) and find that this new firm-level measure is only weakly related to traditional competition measures but can itself reconstructs an industry-level measure. They highlight that the new measure is generally useful for financial statement analysis. An important improvement over the traditional measures is that *Fluidity* and *Pctcomp* are forward-looking and capture the competitive threats from the perspectives of the managers without being bound by the definition of industries.[55] Both measures will be used in our study with a greater emphasis on the former.[56]

[55] Traditional industry classification (e.g., SIC or NAICS) may not be a perfect indication of direct competition among firms. Firms today can compete across numerous different industries and even overseas.
[56] While both measures capture competition threats in a forward-looking manner through firms' qualitative disclosure, we believe that the construction of *Fluidity* from product descriptions resembles more closely the definition of product market competition. Nonetheless, we repeat most regressions with *Pctcomp* and achieved very similar results.

We contribute to the literature in several important ways. Our paper first attempts to shed light on the complex relation between competition and innovation. Second, by using firm-level measures, we can better capture the true competitive pressure confronting individual firms from the perspective of the managers than traditional measures do. Third, we tackle the potential endogeneity issue and provide causal inference for the impact of competition on innovation. Fourth, we provide some empirical insights that may reconcile some of the opposing theories in the literature. Last but not least, we hope to provide some useful implications for policymakers when designing foreign trade policy or anti-trust regulation in an attempt to promote innovation.

The rest of the paper is organized as follows: Section 3 elaborates on the construction of variables and our empirical strategy; Section 4 discusses the main empirical results; Section 5 concludes.

3. Sample Construction and Methodology

3.1 Measuring competition

Our main measure of competition is the *Fluidity* variable developed by Hoberg *et al.* (2014) using firms' 10-K filings. Intuitively, *Fluidity* is defined as a cosine similarity between a firm's own word usage vector with the word usage vector that reflects rival actions. If there is a higher overlap between a firm's products and the changes in the competitors', then the firm is facing greater competition. This measure is obtainable from the Hoberg-Philips Data Library. We also use *Pctcomp* as an alternative measure of competition as in Li *et al.* (2013) mainly for robustness checks.[57] *Pctcomp* measures the number of occurrences of competition-related words which is an indication of competitive pressure faced by the firm from the perspective of the managers. Both

[57] The data for *Pctcomp* is available from Feng Li's website on http://webuser.bus.umich.edu/feng/

measures are firm-specific and based on the textual analysis of management's disclosures in 10-K filings, unlike the traditional measures which are calculated at the industry-level. The above measures are also forward-looking rather than a market representation of the past. Importantly, the above measures also consider competitive threats from non-public firms which constitute a significant portion of the product market.

3.2 Measure Innovation

We use the newest patent database from Kogan *et al.* (2017) which contains information of all patents granted by the U.S. Patent and Trademark Office (USPTO) between 1926 and 2010. The average lag between filing and the issuing of a patent is approximately 2 years. This database potentially does not cover completely the patents applied between 2009 and 2010. Therefore, we drop the samples after 2008 as suggested by Hall *et al.* (2001). We mainly focus on the number of patents filed as a proxy of innovation activity due to the lag between the application and the granting of patents. Nonetheless, we use the number of patents issued as a check. We also use Research and Development (R&D) expenditures from the Compustat Database as a proxy for the inputs to innovation.

3.3 Addressing Endogeneity and Measurement Error

Empirical evidence from the reduced-form regressions is usually plagued with the issue of endogeneity. An omitted factor could exist that affects both competition and innovation or there can be a feedback loop between innovation and competition since innovation itself can affect the competitive landscape of a product market. As emphasized by Aghion *et al.* (2018), finding exogenous variation in competition measure is difficult which can be coupled with the additional problem of measurement error. To address potential omitted variables bias or reverse causality, we followed Xu (2012) who uses import tariff and foreign exchange rates as instrumental variables for the product market competition.[58]

Due to the availability of the import data, our IV regressions are restricted to the firms in the manufacturing industry only. Li and Zhan (2018) have reported that the firm characteristics of this sub-sample are not significantly different from other industries. We show that these two IVs satisfy the relevance condition because they are highly correlated with our abovementioned competition measures. Import tariff is an important policy instrument that regulates competition from foreign firms. Lower import tariff inevitably leads to heightened competition from abroad. Foreign exchange rates also affect the competitiveness of imported goods since cheaper foreign currency encourages imports. We use the updated version of import data as in Schott (2008) and define the tariff rate as the total amount of general import charges divided by the total general import values of the U.S. manufacturing sector.[59] As for calculating the foreign exchange rate, we use the average yearly nominal exchange rate of all US trade partners weighted by the percentage value of their yearly import for each 3-digit SIC industry every year.[60] All exchange rate data were obtained directly from the International Financial Statistics (IFS) online database under the International Monetary Fund (IMF). Both IVs satisfy the exclusion condition because changes in such macro variables are arguably not directly related to firm-level decisions (e.g., innovation) through channels other than changes in the domestic competition landscape.[61]

---

[58] A recent working paper by Pancost and Schaller (2019) argues that instrumental approach not only can be applied to address omitted variables or simultaneity, it also alleviates attenuation bias from classical errors in variables.

[59] The trade data can be obtained from Schott's International Economics Resource Page Trade Data and Concordances at http://faculty.som.yale.edu/peterschott/sub_international.htm

[60] Our results are consistent if we were to convert to real exchange using CPI data from IMF statistics benchmarked to year 2010.

[61] While most of the recent trade agreements are negotiated at a higher political and economic level between countries and under international institutions (e.g., World Trade Organization), we argue that the decision to change tariff is not directly correlated with the domestic firm innovation other than through the channel of product market competition.

Following prior literature, we also make use of large reductions in import tariffs as a natural experiment that represents exogenous shock in product market competition (e.g., Fresard, 2010; Valta, 2012, Li and Zhan, 2018). Intuitively, a reduction in the import tariff reduces trade barriers and increases competition from abroad. A large reduction in import tariffs will significantly intensify the competition of a product market. Following Fresard (2010), Valta (2012), and Li and Zhan (2018), we create a dummy that indicates a negative shock in import tariffs to be one if the tariff reduction is larger than three times the median drop in the same 3-digit SIC industry over the full sample period.[62] As suggested by Li and Zhan (2018), we also drop the events when the tariff rate is less than 1% since the impact from a further decrease in import tariff for an already low rate industry is likely minimal. We then match the firms identified in each industry that experience a negative shock with control firms using propensity matching using firm size, year, ROA, Tobin's Q, and leverage.[63]

3.4 Sample Construction and Descriptive Statistics

We obtain the information of all U.S. public firms from Compustat from 1979 to 2019 and merge the data with the patent database of Kogan *et al.* (2017). Following prior literature, we removed utility and financial firms (i.e., SIC code that begins with 49 or 6) which are heavily regulated and with very different innovation activities. We further merged the resulting sample with *Fluidity* data (1996 to 2017) from the updated Hoberg-Philips Data Library as in Hoberg *et al.* (2014) and with *Pctcomp* data (1994 to 2009) from Feng Li's online database as in Li *et al.* (2013). We end up with a raw sample of 71,847 firm-year observations without missing control variables.

[62] As in Fresard (2010), we also check whether such the negative drop is transitory. Such transitory is few and excluding transitory tariff reduction event does not affect our results.

[63] Please note that the negative shock is industry-specific, therefore, all control firms are from other untreated industries with similar matching characteristics

[Table 1 Here]

Table 1 shows the summary statistics of our variables used in this study. For a detailed definition of the variables, please see Appendix Table A1. Interestingly, a median firm over our entire sample period does not have any explicit record of tangible innovation inputs or outputs. An average firm is having a negative return on assets (ROA). The mean total assets of firms are about $276 million and the average R&D and capital expenditures both account for about 7% of the total assets. The mean leverage ratio of a firm is moderately high about 62%. A firm on average holds about 23% of total assets as cash.

4. Results

A first glimpse of the pairwise correlations in column (1) of Appendix Table A2 suggests a positive correlation between competition and patents filed. While *HHI* and CR_8 are highly correlated (84%), *Fluidity* and *Pctcomp* are only weakly related to the former pair of traditional measures.[64] Surprisingly, the two text-based competition measures are themselves also weakly correlated (15%) to each other. While the methodological difference may be the reason for the weak correlation between our main measures, it is also likely that each measure captures different aspects of the actual competitive threats perceived by the managers. We provide complete correlations between variables in Appendix Table A2.

4.1 Industry Level Competition Measures

Herfindahl-Hirschman Index is arguably the most popular measure for market competition and it has become a widely used control for industry-level competition in both accounting and finance

[64] Which are consistent with Li *et al.* (2013) and Hoberg *et al.* (2014).

research. Following prior literature, we use firms' sales data as a proxy for market shares within each 2-digit SIC industry.[65] Specifically, *HHI* is calculated as follows for each industry:

$$HHI = S_1^2 + S_2^2 + S_3^2 + \cdots + S_N^2$$

where N denotes the number of firms in the industry and S denotes market share for each firm. A higher *HHI* implies a higher market concentration and hence lower competition. Industry concentration, on the other hand, considers only the market shares of the few largest firms. For example, an eight-firm concentration ratio (CR_8) is calculated as the combined market shares of the eight largest firms in a specific industry divided by industry total sales. CR_8 takes a theoretical range between 0 to 1 and indicates the degree of market concentration (competition). An extremely low concentration ratio may be an indication that the industry is close to perfect competition. *HHI* and CR_8 both attempt to measure competition based on the concentration of sales and hence the high correlation.

Table 2 presents the regression results of innovation on *HHI* and CR_8. Column (1) of Table 2 shows a statistically significant negative correlation between *HHI* and the number of patents filed, implying that a lower *HHI* (i.e., more competitive market) is associated with more patents filed. The same can be said for CR_8 in column (3). Column (2) and (4) introduces a square term of the two respective measures to account for potential non-linearity between competition and innovation as documented in prior literature.[66] Interestingly, we observe a moderating effect from the square term. This observation seemingly lends some support to the inverted-U relation as documented by some prior research if we were to plot the number of patents filed against the level of competition.

[65] We also run regressions using *HHI* constructed at 3-digit SIC industry, but the results are substantially weaker in explaining innovation activities under similar empirical setups.
[66] We include the square terms to check for potential non-linear relation (e.g., inverted-U).

However, statistical significance is only observed for *HHI*. For column (4) to (8), we do not observe any statistical significance for both measures and their square terms by adding industry fixed effect on top of firm controls and year fixed effect. By adding industry fixed effect, we account for the unobserved time-invariant heterogeneity within each industry which can be a quick control of the general differences in opportunities for R&D across industries.[67] However, the time variation alone in both *HHI* and CR_8 seems unable to explain significantly the variation in patent counts at firm-level after accounting for industry fixed effects.[68]

There are at least two plausible reasons for the above insignificance: 1) industry-level competition measures do not fully capture or explain the variance in firm-level innovation activities. 2) The way *HHI* or CR_8 was constructed involves random measurement error of actual competition level which leads to attenuation bias that reduces statistical power to obtain significant results. The first caveat is an innate shortcoming of traditional industry-level measures that not only are static and backward-looking, but also ignore the within-industry variation in competition at firm-level (Li *et al.*, 2013 and Hoberg *et al.*, 2014). Another plausible shortfall of using sales data is that competition from private competitor firms is not explicitly addressed, leading to poor proxies for actual industry concentration (Ali *et al.,* 2009 and Bens *et al.*, 2011). To address this concern, we repeat identical tests using the two recently developed forward-looking competition measures which are constructed at firm-level.[69] To overcome endogeneity as well as attenuation bias from

[67] Note that adding *Industry* x *Year* fixed effect will drop both measures from the regression entirely due to collinearity.
[68] Results from Table 2 is based on the final sample we used to run regression on our main competition measure, *Fluidity*, for better cross comparison. When we consider the entire Compustat sample, we do observe some statistical evidence between the correlation of *HHI* or CR_8 with the number of patents filed but it is far from being as robust as *Fluidity* or *Pctcomp* under identical setups.
[69] Li *et al.* (2013) has shown that *Pctcomp* is correlated with traditional measures but at a relatively low level, implying that *Pctcomp* has substantial unique variation. They also provide evidence that this measure is not merely a noisy version of prior competition measures.

measurement error, we adopt the instrumental approach as suggested by Pancost and Schaller (2019).

[Table 2 Here]

4.2 Baseline results on Competition and Innovation

We empirically test the relation between the two new forward-looking competition measures and innovation activities. We first run a baseline pooled OLS regression of innovation variables on *Fluidity* and *Pctcomp*. We included a vector of common variables that are thought to affect a firm's innovation activities as inspired by Aghion *et al.* (2005). We control for lagged firm size, asset tangibility, book leverage, cash holding, Tobin's Q, and profitability. We also include 2-digit SIC industry and year fixed effects for all specifications to generally account for differences in innovation opportunities across industries and over the years. All continuous variables are winsorized at 1% level at each tail of the distribution. The standard errors are all clustered at the firm level.

[Table 3 Here]

Table 3 reports the relation between competition and various measures of firm innovation. All independent and control variables are lagged by one period. Column (1) to (4) show the results for our main competition variable, *Fluidity*, and column (5) to (8) for our alternative competition measure, *Pctcomp*. All specifications yield a positive correlation with the number of patents filed or issued. The results are also consistent across the two competition variables when we use alternative measures of innovation such as patent filed per employee and R&D intensity. The results from Table 3 suggest that not only competition is positively associated with the incentives to invest, but also is correlated with the quantity of innovation output. A one-standard-deviation change in *Fluidity* and *Pctcomp* will lead to about 9.3% and 4.9% increase in the number patent application or 2.7% and 3.5% in the R&D intensity respectively in an average firm over our entire sample period.[70] The economic magnitude is significant considering the patent filed and the R&D intensity for an average firm. It is also reassuring to see that our results being even more robust when we included an additional *Industry* x *Year* fixed effect as shown in Appendix Table A3.[71] This implies that the within-firm variations in competitive threats are explaining the variation in innovation activities in a way substantially better than traditional measures such as *HHI* or *CR*. Bigger firms seem to have a higher number of patent applications and capital expenditure seems to be mostly complementary towards innovation.[72] A higher patent count is also negatively related to profitability but there is some evidence that the output of innovation is positively correlated with market valuation (i.e., *TobinsQ*). Interestingly, a positive increase in cash holdings is associated with higher innovation activity. This result coincides with the findings by He and Wintoki (2016) that the increasing cash holdings were almost entirely driven by R&D intensive firms arguably because cash is a strategic asset to ensure smooth investment when sources of finance become volatile. Our baseline results using the two new firm-specific measures support the theory that competition encourages innovation.

[Table 4 Here]

One might raise a concern about the above results that a subset of our sample is driving what we observed. It is plausible that such a positive relationship between competition and innovation only exists in the extremes of the distribution and that the relation may not be as linear as what we

[70] Note that $(e^{3.51*0.0184} - 1) / (e^{0.54} - 1) = 9.3\%$ and $(e^{0.47*0.0737} - 1) / (e^{0.54} - 1) = 4.9\%$.
[71] This is also equivalent to controlling for HHI or CR in Table 3.
[72] A negative correlation between size and R&D intensity can be due to the denominator effect in the construction of the R&D intensity variable for column (4) and (8).

thought.[73] To shed more light on this front, we constructed dummies that indicate whether a firm is in the top 30% percentile (i.e., *HighFluidity*), middle 40% percentile (i.e., *MidFluidity*), or bottom 30% percentile (i.e., *LowFluidity*) in terms of *Fluidity* from the previous period. We then repeat the same regressions as in Table 3 using our new indicator dummies and the results are reported in Table 4. The indicator dummies in column (1), (3), and (5) are constructed by sorting the full sample and the dummies in column (2), (4), and (6) by sorting within each 2-digit SIC industry. The observed results are overall consistent and do not seem to differ much by how we sort *Fluidity*. Being a firm facing high competitive threats seems to file more patents and the opposite is true for firms having low competitive threats. Firms having a moderate level of competition do not seem to exhibit any statistically significant relation with patent counts. This linearity is further confirmed when we repeat the setups from Table 3 using a new rank variable, *FluidityRank*, that takes the number from 1 to 10 with 1 being in the decile of the lowest competitive threats and 10 the highest. The results are reported in Appendix Table A4 which indicates a very similar pattern. We, therefore, focus on reduced form linear regressions with IVs for competition measures as in Xu (2012) and Li and Zhan (2018) in the next section.

4.3 Evidence from Import Tariff

4.3.1 Instrumental Variables

To address the potential concerns on endogeneity between competition and innovation, we implement instrumental variables (IV) regression. Following Xu (2012), we use import tariff and foreign exchange rate as IVs for our main competition variable, *Fluidity*.[74] As abovementioned, it

[73] Considering that several prior studies have documented a non-linear relationship between competition and innovation, we have also run regressions with quadratic term. However, the square term is not statistically significant for our main competition measure.

[74] We also obtained very similar results using *Pctcomp*.

satisfies the relevance condition because it is highly correlated with our competition measures and is arguably not directly related to firm-level innovation through channels other than competition.

[Table 5 Here]

Table 5 reports the results for our first stage and second-stage IV estimates. As shown in column (8), we regress *Fluidity* on import tariffs, foreign exchange rate, firm-level controls, and the year and industry fixed effects in the first-stage. The negative sign on *Import Tariff* shows that a higher import tariff reduces competition. Intuitively, an increase in the tariff reduces the price competitiveness of foreign firms. A negative coefficient in foreign exchange rates (i.e., U.S. dollar/foreign currency) implies that a higher valuation of the U.S. dollar encourages imports because it reduces the prices of foreign goods.[75] We then use the predicted value of *Fluidity* from our first-stage IV regression in the second-stage. To examine the validity of our instruments formally, we conduct the Hansen *J*-statistics test for the overidentifying restrictions and the null that the instruments are valid. The *J*-statistics are not statistically significant with a *p*-value above 10%, indicating that the instruments are uncorrelated with the error terms of the model, therefore, satisfying the exclusion condition. We also formally perform Kleibergen-Paap rk LM statistics test for under-identification with the null that there is no correlation between the IVs and the endogenous variable. The *p*-values of LM statistics clearly reject the null that the IVs are weak instruments, hence satisfying the relevance condition. Column (1) to (4) reports the second stage IV regression estimates for various innovation measures. We observed unambiguous positive

[75] A drop in foreign exchange rate (US$/Foreign$) will increase *Fluidity* (i.e., competition) because imports are cheaper and vice-versa.

coefficients for *Fluidity* in our second-stage estimations across various measures of innovation and our results support a causal inference that competition increases firm innovation.

4.3.2 Exogenous Shock in Tariff

We further confirm the causal relation between product market competition and innovation using exogenous decreases in the tariffs as a robustness check. From the U.S. import data, we observe numerous industry-specific large decreases in tariffs between 1990 to 2008.[76] Following Fresard (2010), Valta (2012), Li and Zhan (2018), we define an exogenous negative shock to be one in each industry-year when the import charge reduction is more than three times the median drop in the same industry over our sample period. We found a total of 193 exogenous tariff shocks identified during our sample period. We identified 4,035 firm-year observations that experience the exogenous reduction in import charges. Using propensity score matching, we match firms that experience negative tariff decrease in a specific year with control firms of very similar firm characteristics.[77] Table 6 reports the average treatment effects (i.e., changes in patent filing activity) on the treated firms versus similar control firms. We matched on year, firm size, ROA, Tobin's Q, and leverage. A caveat in the interpretation of Table 6 is that matching could not account for unobserved heterogeneity between the treated and the untreated and that even the closest controls are from different industries. A benefit of matching is that it makes no explicit assumption on functional forms as compared to conventional OLS regression. Nonetheless, the results are consistent with our previous results that competition stimulates innovation in terms of patent applications.

[76] Sample period is the overlap of both the patent and the import data.
[77] By construction, the covariates of the control firms follow a statistically similar distribution as the treated.

[Table 6 Here]

4.4 Heterogeneous Effect from Competition

While the previous section presents evidence on the causal relation between competition on innovation, we also attempt to examine how firm-level characteristics, the current state of technology, and industry outlook will affect the role of competition in promoting innovation. Their potential interactive effects with competition may provide some insights into the diverse theoretical and empirical predictions of extant literature.

4.4.1 Past Profitability

Internal funding is a key source of financing for innovation which is related to the past profitability of the firm (Tang, 2006). We investigate how the past profitability of a firm will affect the firm's innovation under greater competition. We measure the average ROA of a firm in the past five years (i.e., *ROA_l5*) and interact it with *Fluidity*. Table 7 reports the second stage IV estimates using import tariff and foreign exchange rate and their respective interactions with *ROA_l5* as the IVs for *Fluidity* and its interaction with *ROA_l5* respectively.[78] Our evidence first suggests that the positive effect of competition on innovation is greater when firms were very profitable in the past five years. Competition that potentially reduces a firm's pre-innovation profit seems to stimulate innovation. Firms that are making profits before competition have greater incentives to retain their profits by innovating as argued by Schumpeter (1942). Interestingly, the negative association between past profitability and innovation suggests that incumbent firms that already captured a good portion of the total profits in the past are on average less innovative. In Arrow (1962), a monopolist has a reduced incentive to spend in innovation than that in a competitive industry

[78] Using IVs alleviate potential endogeneity concerns and measurement errors. Nonetheless, not shown in the paper, the pooled OLS versions of the regressions yields similar results and conclusions.

because the firm is already earning positive profit and innovation only replaces the profitable monopoly with another. While the first finding favors the argument in Schumpeter (1942), the latter seems consistent with the replacement effect in Arrow (1962). Our empirical evidence seems to reconcile somehow the arguments of the two studies. However, we did not find a statistically significant impact on innovation productivity per employee.

[Table 7 Here]

4.4.2 Future Growth

The goal of a firm is to create values for its stakeholders which is often evaluated by its profitability. Industry outlooks that determine the potential profits from the future inevitably play a crucial role in firms' incentives to innovate. We use the past 5 years industry mean sales growth as a proxy for future outlooks. Following the methodology from Table 7, we test the interactive effect of the proxy for industry outlook on our instrumented competition measure. Table 8 reports the second-stage results from IV regressions. Unsurprisingly, we observe evidence that a better industry-level prospect promotes innovation. The positive causal relation between *Fluidity* and our innovation measures remain consistent and robust. However, the negative coefficients of the interaction terms in all three columns indicate that such a positive effect is less prominent in the presence of greater competition. A plausible explanation is that firms have a greater incentive for innovation when the industry is booming with growing future profits, but such incentive is diminishing the presence of greater competitive threats that might potentially erode post-innovation profits. In other words, competition lowers the future level of expected returns from innovation, hence the incentive to invest in R&D.

[Table 8 Here]

4.4.3 Financial Constraints

Innovation is costly with huge uncertainty, especially at its initial stage. The financial status of a firm undoubtedly plays an influential role in any R&D investment decisions. Following an identical strategy as in Table 7, we explore the competition effect for firms with high financial constraints. Without financial flexibility, firms are more susceptible to aggressive market strategies from rival firms. We measure financial constraints using the WW index as in Whited and Wu (2006). We assign 1 to firms that have above WW index higher than that of the median firm in their respective 2-digit SIC industry in the same year.[79] Firms have to allocate resources consistently for continuous innovation and those experiencing financial constraints are less flexible in securing positive NPV projects (Campello 2006). This argument is in line with the negative sign of *Fin_Constraint* that firms in the high financially constrained group on average experience a lower level of innovation activity. The full results are shown in Table 9. Interestingly, we observe a positive coefficient for the interaction, *Fluidity* x *Fin_Constraint*. The positive effect on innovation from heightened competition is more pronounced for firms with higher financial constraints. Plausible explanations include that firms having financial constraint is more incentivized to escape competition through innovation. Losing competitive edges in the innovation race may kick them out of the competition entirely. Aghion *et al.* (1999) argue that managers innovate more in competitive markets because the bankruptcy risk is lower should they manage to escape competition and establish monopolistic power. Again, we did not find statistically significant results on the innovation productivity per employee for firms under different financial constraints.

[79] Different industries might have different threshold for financial constraints. Our results are the same if we instead use our sample median or 3-digit SIC median as a benchmark.

[Table 9 Here]

4.4.4 Distance to Technology Frontier

Harris and Vickers (1987) and Lippman and McCardle (1987) use exponential discovery to model the dynamics of R&D competition. While in these models, even though a lagging firm can catch up or jump ahead of the competing firms, a leader still has a greater probability to win the race since completing a phase in R&D typically increases the incentive to invest more. In the modern world, not all innovation is built from scratch and many breakthroughs are built onto a firm's existing knowledge base or patents. The current technology state of a firm certainly plays a pivotal role in deciding its innovation policy. Firms facing a similar level of competitive threats may adopt an entirely different innovation strategy because of differences in their existing state of technology. It thus may also affect the expected costs and returns from innovation in an attempt to escape competition. Therefore, we examine how the current inventory of technology affect the relation between competition and innovation. Inspired by the concept of technological frontier, we measure how far the firm's existing technology (patent) inventory is to an average firm in the same industry. We define both *Dist_Tech* and *Dist_Tech_l5* as the proxies for the relative distance of a firm to the industry mean patent inventory level in the same 3-digit SIC industry. More specifically, *Dist_Tech* measures the difference between the total number of patents filed by a firm and its industry mean and normalized by the industry mean whereas *Dist_Tech_l5* measures the difference between the total number of patents filed by a firm in the past 5 years and its industry mean over the same period and then normalized by the corresponding 5-year industry mean. Following Table 7, we explore the interactive effect of competition on innovation for firms having different current states of technology. The results are shown in Table 10. We find that firms running ahead of the innovation race in the past years are more innovative and such incentive to innovate is even greater under competition. The regressions yield positive signs for the interaction terms indicating a stronger positive effect on innovation for firms with a higher state of technology. Intuitively, firms that are closer to the technology frontier (i.e., technologically leading firms) have a higher probability of escaping the competition through innovation whereas the laggard firms have less incentives to merely catch up when anticipating a greater competition ahead. The results are consistent with the argument in Aghion *et al.* (2005), competition discourages laggard firms (i.e., firms that are further away from the technology frontier) to innovate but encourage leading neck-and-neck firms to innovate more and escape competition.

[Table 10 Here]

4.5 Quality of Innovation

This paper so far discussed the quantitative side of innovation outputs under heightened competition. It is also important for us to shed some light on the qualitative aspect of innovation output. We use patent citations and patent value as proxies for innovation quality as in Kogan *et al.* (2017). We report the combined results of *Fluidity* and its interactions with past profitability, industry outlooks, technology state, and financial constraints on both measures of innovation quality in Table 11. It is reassuring to observe a consistent and statistically significant relation between *Fluidity* and both measures of innovation quality. Our results unambiguously suggest that firms under competitions produce better quality patents in terms of patent citations and patent value. Having greater competition not only leads to a greater number of patents but also their quality. Past profitability seems to reinforce the positive impact on patent value from heightened competition as shown in Column (3) of Panel B. Better financial conditions potentially create a less pressing work environment that encourages innovation of better value. In column (4) and (6) of both panels, we find that the interaction effect from competition is greater for firms with higher financial constraints but less so for firms in a growing industry. We argue that firms having financial constraint is more incentivized to escape competition not only through more innovation, but also innovation of better quality. Both dimensions of innovation help firms differentiate their products and hence the ability to profit. Similarly, a firm's incentive for quality innovation is reduced upon the anticipation of lower expected profit shares due to greater competition. Overall, the above relations do not seem to deviate from our previous arguments about firms' incentive to increase patent counts. From column (5) of both panels, the positive coefficients of *Dist_Tech_l5* suggest a momentum effect that a firm on average produces better quality patents when it has already a large accumulation of patent pool relative to its industry peers. It is plausible that the synergy from the existing knowledge pool helps to produce innovation of better quality. However, competition per se does not seem to reinforce or diminish such impact.

[Table 11 Here]

4.6 Competition and Innovation Efficiency

Boone (2000) shows theoretically that the effect of competition on a firm's incentive to innovate is also related to the firm's efficiency level relative to that of the industry peers. Inspired by his model, we also examine empirically if the competition landscape per se affects a firm's innovation efficiency. There certainly exists huge heterogeneity in innovation efficiency among individual firms that differ vastly as first shown in Table 1. We explore how product market competition is going to affect a firm's efficiency in innovation. Intuitively, a firm can become more efficient in the innovation process under greater competitive pressure because managers have the incentives to keep the firm alive and stay competitive by reducing the waste of unnecessary resources. On the other hand, firms may also, under heightened competition, over-invest, or misallocate resources in innovation projects in the hope of escaping competition through innovation breakthroughs.

[Table 12 Here]

Table 12 reports the combined results of *Fluidity* and its interactions with past profitability, industry outlooks, technology state, and financial constraints on innovation efficiency. We measure innovation efficiency as the natural log of one plus the average number of patents filed from *t* to *t*+2 divide by the mean R&D expenditures in the same period. This measure allows us to gauge the ratio of outputs to inputs quantitatively. Results from OLS regression suggests a strong negative correlation between competition and future innovation efficiency, suggesting the possibility of greater resources wasting under competition during the innovation process. However, the evidence weakens statistically after we account for endogeneity using IVs. From column (5), firms that run ahead of others in terms of technology accumulation seem to waste less resources and exhibit greater R&D efficiency on average, but competition seems to reduce such efficiency though weakly. Other interactive effects, however, do not yield statistical significance at all.

5. Conclusion

Gilbert (2006) reiterates the inconclusiveness of the theory between the effect of competition and innovation incentives. While more recent papers try to reconcile the two opposing theories by documenting non-monotonic effects of competition on innovation, it is still far from a consensus since the heterogeneous effects of competition are often dependent on both micro- and macro-level variables and assumptions. Empirically, the evidence documented in the extant literature is even more contrasting, from positive, negative, non-monotonic, to none. This paper attempts to shed light on the long-debated relation between competition and innovation using empirical evidence. By using textual-based competition measures at firm-level, we are able to capture within-industry variation which is often absent in the traditional measures used in the prior literature. The managers who drafted 10-K filings are arguably the most appropriate persons to gauge the true competitive pressure facing their firms than any industry concentration ratios calculated from historical sales numbers. Most importantly, we provide causal inference by addressing the endogeneity issues between competition and innovation using instrument variables and exogenous shocks in import tariffs. We primarily find that product market threats encourage firms to innovate and enhance both the quantity and quality of innovation outputs. However, we also find some evidence that competition may reduce innovation efficiency. By examining how firm-level characteristics, state of current technology, and industry outlook might affect the role of competition in promoting innovation, we provide some empirical insights that seem to reconcile some of the opposing theories in the literature. Lastly, we hope to provide some useful implications for policymakers when designing trade policies or anti-trust regulations to promote domestic innovation.

References

Aghion, P., Bechtold, S., Cassar, L., & Herz, H. (2018). The causal effects of competition on innovation: Experimental evidence. The Journal of Law, Economics, and Organization, 34(2), 162-195.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., & Howitt, P. (2005). Competition and innovation: An inverted-U relationship. *The Quarterly Journal of Economics*, *120*(2), 701-728.

Aghion, P., Dewatripont, M., & Rey, P. (1999). Competition, financial discipline and growth. *The Review of Economic Studies*, *66*(4), 825-852.

Aghion, P., & Griffith, R. (2008). *Competition and growth: reconciling theory and evidence*. MIT press.

Ali, A., Klasa, S., & Yeung, E. (2008). The limitations of industry concentration measures constructed with Compustat data: Implications for finance research. The Review of Financial Studies, 22(10), 3839-3871.

Angelmar, R. (1985). Market structure and research intensity in high-technological-opportunity industries. The journal of industrial economics, 69-79.

Arrow, K. (1962). Economic welfare and the allocation of resources for invention. The rate and direction of inventive activity: economic and social factors. *N. Bureau*.

Bens, D. A., Berger, P. G., & Monahan, S. J. (2011). Discretionary disclosure in financial reporting: An examination comparing internal firm data to externally reported segment data. The Accounting Review, 86(2), 417-449.

Blundell, R., Griffith, R., & Reenen, J. V. (1995). Dynamic count data models of technological innovation. *The Economic Journal*, *105*(429), 333-344.

Boone, J. (2000). Competitive pressure: the effects on investments in product and process innovation. *The RAND Journal of Economics*, 549-569.

Boone, J. (2001). Intensity of competition and the incentive to innovate. *International Journal of Industrial Organization*, *19*(5), 705-726.

Campello, M. (2006). Debt financing: Does it boost or hurt firm performance in product markets? *Journal of Financial Economics*, *82*(1), 135-172.

Chen, Y., & Schwartz, M. (2013). Product innovation incentives: Monopoly vs. competition. *Journal of Economics & Management Strategy*, *22*(3), 513-528.

Fresard, L. (2010). Financial strength and product market behavior: The real effects of corporate cash holdings. *The Journal of finance*, *65*(3), 1097-1122.

Gilbert, R. (2006). Looking for Mr. Schumpeter: Where Are We in the Competition-Innovation Debate? *Innovation policy and the economy*, *6*, 159-215.

Gilbert, R., & David M. G. Newbery. (1982). Preemptive Patenting and the Persistence of Monopoly. *The American Economic Review*, *72*(3), 514-526.

Greenstein, S., & Ramey, G. (1998). Market structure, innovation and vertical product differentiation. International Journal of Industrial Organization, 16(3), 285-311.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). *The NBER patent citation data file: Lessons, insights and methodological tools* (No. w8498). National Bureau of Economic Research.

Harris, C., & Vickers, J. (1987). Racing with uncertainty. The Review of Economic Studies, 54(1), 1-21.

Hashmi, A. R. (2013). Competition and innovation: The inverted-U relationship revisited. *Review of Economics and Statistics*, *95*(5), 1653-1668.

He, Z., & Wintoki, M. B. (2016). The cost of innovation: R&D and high cash holdings in US firms. *Journal of Corporate Finance*, *41*, 280-303.

Hoberg, G., Phillips, G., & Prabhala, N. (2014). Product market threats, payouts, and financial flexibility. *The Journal of Finance*, *69*(1), 293-324.

Kamien, M. I., & Schwartz, N. L. (1975). Market structure and innovation: A survey. *Journal of economic literature*, *13*(1), 1-37.

Kogan, L., Papanikolaou, D., Seru, A., & Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics*, *132*(2), 665-712.

Li, F., Lundholm, R., & Minnis, M. (2013). A measure of competition based on 10-K filings. *Journal of Accounting Research*, *51*(2), 399-436.

Li, S., & Zhan, X. (2018). Product market threats and stock crash risk. *Management Science*.

Lippman, S. A., & McCardle, K. F. (1987). Dropout behavior in R&D races with learning. The RAND Journal of Economics, 287-295.

Mansfield, E. (1977). The production and application of new industrial technology. Norton.

Nickell, S. J. (1996). Competition and corporate performance. *Journal of political economy*, *104*(4), 724-746.

Pancost, N. A. & Schaller, G. (2019). Measuring Measurement Error. Working paper.

Reinganum, M. R. (1983). The anomalous stock market behavior of small firms in January: Empirical tests for tax-loss selling effects. Journal of financial economics, 12(1), 89-104.

Scherer, F. M. (1965). Firm size, market structure, opportunity, and the output of patented inventions. The American economic review, 55(5), 1097-1125.

Scherer, F. M. (1984). Innovation and Growth: Schumpeterian Perspectives Cambridge. *Mass.: MITPress*.

Schmidt, K. M. (1997). Managerial incentives and product market competition. The Review of Economic Studies, 64(2), 191-213.

Schott, P. K. (2008). The relative sophistication of Chinese exports. *Economic policy*, *23*(53), 6-49.

Schumpeter, J. A. (1942). *Socialism, capitalism and democracy*. Harper and Brothers.

Valta, P. (2012). Competition and the cost of debt. *Journal of Financial Economics*, *105*(3), 661-682.

Weinberg, J. A. (1992). Exclusionary practices and technological competition. The Journal of Industrial Economics, 135-146.

Whited, T. M., & Wu, G. (2006). Financial constraints risk. *The Review of Financial Studies*, *19*(2), 531-559.

Xu, J. (2012). Profitability and capital structure: Evidence from import penetration. *Journal of Financial Economics*, *106*(2), 427-446.

Table 1. Summary Statistics

This table reports the summary statistics of the main variables used in this study. The patent related measures were obtained from Kogan *et al.* (2017) which contains information of all patents granted by the U.S. Patent and Trademark Office (USPTO) between 1926 and 2010. *Fluidity* and *Pctcomp* are developed by Hoberg *et al.* (2014) and Li *et al*. (2013) respectively. Other accounting variables are constructed from the Compustat database. For a detailed definition of all variables, please refer to Appendix Table A1.

Variables	**N**	**Mean**	**S.D.**	**.25**	**Mdn**	**.75**
Main LHS Variables						
Patent_F	43977	0.54	1.08	0.00	0.00	0.69
Patent_I	43977	0.57	1.09	0.00	0.00	0.69
Patent_Per_Emp	42917	0.53	1.07	0.00	0.00	0.34
R&D_Efficiency	21575	0.16	0.28	0.00	0.03	0.21
Patent_Cites	43977	0.74	1.61	0.00	0.00	0.00
Patent_Value	43977	0.75	1.67	0.00	0.00	0.12
R&D_Intensity	71845	0.07	0.15	0.00	0.00	0.09
Main RHS Variable						
Fluidity	71845	6.79	3.51	4.16	6.09	8.75
Pctcomp	26234	0.57	0.47	0.23	0.42	0.76
HHI	71845	0.15	0.14	0.06	0.11	0.19
CR8	71845	0.53	0.16	0.41	0.49	0.98
Roa_l5	71845	-0.24	1.44	-0.08	0.03	0.08
Fin_Constraint	71845	0.49	0.50	0.00	0.00	1.00
Dist_Tech	41441	0.03	3.35	-1.00	-1.00	-0.78
Dist_Tech_l5	46909	0.05	3.23	-1.00	-1.00	-0.69
Growth	71845	1.69	2.52	0.17	0.48	1.99
Controls						
Size	71845	5.62	2.13	4.07	5.56	7.11
Cash	71845	0.23	0.46	0.03	0.10	0.24
Ppe	71845	0.28	0.30	0.08	0.18	0.39
Capex	71845	0.07	0.11	0.02	0.04	0.08
Leverage	71845	0.62	0.81	0.31	0.51	0.73
Roa	71845	-0.12	0.71	-0.10	0.03	0.08
TobinsQ	71845	2.35	3.17	1.14	1.58	2.51

Table 2. Pooled OLS Regression – *HHI and CR8*

This table reports the OLS regression results for the traditional industry-level competition measures, *HHI and CR8*. Standard errors were clustered by industry and reported in the parentheses. Column (1) to (4) include only year fixed effects and column (5) to (8) include both industry and year fixed effects. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For detailed definitions of variables, please refer to Appendix A1.

	(1) *Patent_F*	(2) *Patent_F*	(3) *Patent_F*	(4) *Patent_F*	(5) *Patent_F*	(6) *Patent_F*	(7) *Patent_F*	(8) *Patent_F*
HHI	-2.8194***	-7.0983***			0.1971	-0.5381		
	(0.885)	(2.547)			(0.230)	(0.605)		
HHI^2		14.4062**				2.2100		
		(6.088)				(1.405)		
CR_8			-1.0575**	-1.2825			-0.2056	-1.2332
			(0.409)	(2.793)			(0.249)	(0.792)
CR_8^2				0.1889				0.8808
				(2.182)				(0.622)
Size	0.2439***	0.2439***	0.2442***	0.2441***	0.2680***	0.2681***	0.2681***	0.2680***
	(0.041)	(0.041)	(0.041)	(0.041)	(0.040)	(0.040)	(0.040)	(0.040)
Ppe	-0.6230***	-0.5945***	-0.5950***	-0.5951***	-0.3978***	-0.3980***	-0.3978***	-0.3975***
	(0.188)	(0.185)	(0.192)	(0.191)	(0.102)	(0.102)	(0.102)	(0.102)
Capex	0.4950*	0.5038*	0.4662	0.4707*	0.5381***	0.5390***	0.5399***	0.5408***
	(0.291)	(0.296)	(0.289)	(0.275)	(0.173)	(0.173)	(0.173)	(0.172)
Leverage	-0.0933***	-0.0916***	-0.0936***	-0.0936***	-0.0624***	-0.0622***	-0.0624***	-0.0624***
	(0.027)	(0.026)	(0.027)	(0.028)	(0.020)	(0.020)	(0.020)	(0.020)
Roa	-0.0546***	-0.0485***	-0.0502***	-0.0497***	-0.0392***	-0.0391***	-0.0393***	-0.0393***
	(0.013)	(0.012)	(0.012)	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)
TobinsQ	0.0370***	0.0358***	0.0360***	0.0359***	0.0324***	0.0324***	0.0324***	0.0325***
	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
Cash	0.1039**	0.0933*	0.0986*	0.0979*	0.0554*	0.0554*	0.0553*	0.0553*
	(0.052)	(0.050)	(0.052)	(0.055)	(0.029)	(0.029)	(0.029)	(0.029)
Constant	-0.4745***	-0.3105***	-0.1019	-0.0397	-0.8531***	-0.8227***	-0.7342***	-0.4584
	(0.096)	(0.109)	(0.173)	(0.792)	(0.213)	(0.210)	(0.225)	(0.319)
Industry FE	No	No	No	No	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.22	0.22	0.22	0.22	0.36	0.36	0.36	0.36
Observations	43977	43977	43977	43977	43977	43977	43977	43977

Table 3. Pooled OLS Regression - *Fluidity* and *Pctcomp*

This table reports the OLS regression results for the two text-based firm-level competition measures, *Fluidity* and *Pctcomp*. Standard errors were clustered by firms and reported in the parentheses. Year and 2-digit SIC industry fixed effects are included in all specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For detailed definitions of variables, please refer to Appendix A1.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Patent_F	*Patent_I*	*Patent_Per_Emp*	*R&D_Intensity*	*Patent_F*	*Patent_I*	*Patent_Per_Emp*	*R&D_Intensity*
Fluidity	**0.0184***	**0.0181***	**0.0634***	**0.0128***				
	(0.004)	(0.004)	(0.004)	(0.000)				
Pctcomp					**0.0737***	**0.0402***	**0.1357***	**0.0171***
					(0.023)	(0.023)	(0.023)	(0.002)
Size	0.2664***	0.2616***	0.0583***	-0.0131***	0.3403***	0.3349***	0.0706***	-0.0021***
	(0.009)	(0.009)	(0.005)	(0.001)	(0.013)	(0.013)	(0.007)	(0.000)
Ppe	-0.3753***	-0.4101***	-0.3329***	-0.0422***	-0.3776***	-0.3837***	-0.2438***	-0.0268***
	(0.052)	(0.052)	(0.040)	(0.004)	(0.081)	(0.080)	(0.053)	(0.003)
Capex	0.4895***	0.3140***	0.4578***	-0.0391***	0.5758***	0.4425***	0.5193***	0.0186***
	(0.070)	(0.067)	(0.070)	(0.008)	(0.106)	(0.105)	(0.085)	(0.007)
Leverage	-0.0578***	-0.0514***	-0.0968***	-0.0137***	-0.1826***	-0.1747***	-0.2132***	-0.0130***
	(0.007)	(0.007)	(0.009)	(0.001)	(0.027)	(0.026)	(0.022)	(0.002)
Roa	-0.0313***	-0.0535***	-0.0080	-0.0346***	-0.2232***	-0.2634***	-0.3311***	-0.0340***
	(0.009)	(0.008)	(0.013)	(0.002)	(0.045)	(0.050)	(0.055)	(0.006)
TobinsQ	0.0312***	0.0202***	0.0279***	0.0096***	0.0840***	0.0575***	0.0743***	0.0097***
	(0.003)	(0.002)	(0.003)	(0.001)	(0.011)	(0.011)	(0.011)	(0.001)
Cash	0.0407***	-0.0069	0.1737***	0.0035	0.0748**	0.0393	0.2941***	0.0241***
	(0.012)	(0.012)	(0.018)	(0.002)	(0.036)	(0.033)	(0.043)	(0.003)
Constant	-0.9569***	-0.8499***	-0.1948***	0.0568***	-1.3387***	-1.2184***	-0.0382	0.0333***
	(0.045)	(0.045)	(0.035)	(0.004)	(0.068)	(0.068)	(0.046)	(0.004)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.36	0.36	0.27	0.49	0.41	0.40	0.25	0.33
Observations	43977	43977	42917	71845	27580	27580	27083	30560

Table 4. Pooled OLS Regression – High, Mid, and Low Competitive Threats

This table reports the OLS regression results for dummies that indicate whether a firm is in the top 30% percentile, middle 40% percentile, or bottom 30% percentile group in terms of competitive threats (i.e., *Fluidity)*. The three dummies in column (1), (3), and (5) are sorted based on full sample and column (2), (4), and (6) are sorted within each 2-digit SIC industry. Standard errors were clustered by firms and reported in the parentheses. Year and 2-digit SIC industry fixed effects are included in all specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For detailed definitions of variables, please refer to Appendix A1.

	(1) *Patent_F*	(2) *Patent_F*	(3) *Patent_F*	(4) *Patent_F*	(5) *Patent_F*	(6) *Patent_F*
HighFluidity	0.0854*** (0.023)					
HighFluidity(SIC2)		0.0897*** (0.020)				
MidFluidity			0.0044 (0.016)			
MidFluidity(SIC2)				-0.0074 (0.016)		
LowFluidity					-0.0904*** (0.022)	
LowFluidity(SIC2)						-0.0776*** (0.019)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.36	0.36	0.36	0.36	0.36	0.36
Observations	43977	43977	43977	43977	43977	43977

Table 5. IV regression

This table presents the two-stage least squares (2SLS) regression results for our main competition measure, *Fluidity*. Column (1) to (4) reports the second-stage estimation and column (5) shows the result for the first-stage regression. The instrumental variables for competition are import tariff, which is available only for firms in the manufacturing industry (SIC 2000-3999), and foreign exchange rate. Standard errors were clustered by firms and reported in the parentheses. Year and 2-digit SIC industry fixed effects are included in all specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For detailed definitions of variables, please refer to Appendix A1.

	Second-Stage Estimation				First-Stage Estimation
	(1) *Patent_F*	(2) *Patent_I*	(3) *Patent_Per_Emp*	(4) *R&D_Intensity*	(5) *Fluidity*
Fluidity	0.1515***	0.1458***	0.1783***	0.0321***	
	(0.029)	(0.029)	(0.021)	(0.002)	
Import Tariff					-0.4142***
					(0.050)
Exchange Rate					-0.8926***
					(0.266)
Size	0.3772***	0.3810***	0.0909***	-0.0151***	0.0627***
	(0.012)	(0.012)	(0.008)	(0.001)	(0.023)
Ppe	-0.1691*	-0.2248**	-0.2565***	-0.0209**	-1.8988***
	(0.100)	(0.100)	(0.083)	(0.008)	(0.225)
Capex	0.3722**	0.1318	0.5897***	-0.0607***	2.7763***
	(0.163)	(0.162)	(0.150)	(0.014)	(0.381)
Leverage	-0.0847***	-0.0806***	-0.1416***	-0.0139***	-0.3140***
	(0.021)	(0.022)	(0.021)	(0.002)	(0.050)
Roa	0.0087	-0.0381*	0.0548**	-0.0422***	-0.5550***
	(0.024)	(0.023)	(0.027)	(0.003)	(0.055)
TobinsQ	0.0342***	0.0242***	0.0280***	0.0111***	0.0832***
	(0.004)	(0.004)	(0.005)	(0.001)	(0.011)
Cash	-0.0738**	-0.1622***	0.1564***	-0.0224***	0.9115***
	(0.033)	(0.033)	(0.035)	(0.004)	(0.057)
Constant					8.0137***
					(0.204)
Industry FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
LM statistic (*p*-value)	139(0.00)	139(0.00)	126(0.00)	157(0.00)	
Hansen's *J*-statistic (*p*-value)	0.15(0.70)	0.03(0.86)	0.17(0.68)	1.99(0.16)	
Adj. R^2	0.27	0.28	0.07	0.28	0.46
Observations	23103	23103	22657	38831	38831

Table 6. Exogenous Negative Shock on Import Tariff

We define an exogenous negative shock to be one in each industry-year when the import charge reduction is more than three times the median drop in the same industry over the sample period between 1990 to 2008. We identify a total of 193 exogenous tariffs reduction and 4,035 firm-year observations that experience such a shock. We then use propensity score matching with replacement to match each treated firm with a control firm of similar characteristics. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For detailed definitions of variables, please refer to Appendix A1.

Log (1 + Δ Patent Filed)	(1)	(2)	(3)
Average Treatment Effect	0.0662***	0.0614***	0.0992***
	(0.021)	(0.021)	(0.028)
Matched on:			
Year	Yes	Yes	Yes
Size	Yes	Yes	Yes
ROA	No	Yes	Yes
TobinsQ	No	No	Yes
Leverage	No	No	Yes

Table 7. Competition and Past Profitability

This table presents only the results for the second-stage regression estimates. We use import tariff and foreign exchange rate and their interactions with *ROA*_l5 as the IVs for *Fluidity* and *Fluidity* x *ROA_l5* respectively. Tariff rates are available only for firms in the manufacturing industry (SIC 2000-3999) and hence the smaller sample size. In untabulated results, we observed similar outcomes using Pooled OLS regression without IVs. Standard errors were clustered by firms and reported in the parentheses. Year and 2-digit SIC industry fixed effects are included in all specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For detailed definitions of variables, please refer to Appendix A1.

	(1)	(2)	(3)
	Patent_F	*Patent_I*	*Patent_Per_Emp*
Fluidity x *ROA_l5*	0.0103**	0.0127***	0.0061
	(0.004)	(0.003)	(0.006)
Fluidity	0.1519***	0.1482***	0.1787***
	(0.029)	(0.030)	(0.021)
ROA_l5	-0.0788**	-0.1177***	-0.0023
	(0.036)	(0.026)	(0.049)
Size	0.3771***	0.3795***	0.0898***
	(0.012)	(0.012)	(0.008)
Ppe	-0.1801*	-0.2348**	-0.2661***
	(0.099)	(0.099)	(0.083)
Capex	0.3597**	0.1129	0.5840***
	(0.166)	(0.166)	(0.151)
Leverage	-0.0869***	-0.0736***	-0.1458***
	(0.019)	(0.019)	(0.020)
TobinsQ	0.0341***	0.0237***	0.0295***
	(0.004)	(0.004)	(0.005)
Cash	-0.0626*	-0.1351***	0.1571***
	(0.035)	(0.036)	(0.035)
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R^2	0.28	0.28	0.07
Observations	23103	23103	22657

Table 8. Competition and Industry Growth

This table presents only the results for the second-stage regression estimates. We use import tariff and foreign exchange rate and their interactions with *Growth* as the IVs for *Fluidity* and *Fluidity* x *Growth* respectively. Tariff rates are available only for firms in the manufacturing industry (SIC 2000-3999) and hence the smaller sample size In untabulated results, we observed similar outcomes using Pooled OLS regression without IVs. Note that the R^2 from IV estimation can be negative because *SSR* for IV can be larger than *SST* (Wooldridge, 2006). Standard errors were clustered by firms and reported in the parentheses. Year and 2-digit SIC industry fixed effects are included in all specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For detailed definitions of variables, please refer to Appendix A1.

	(1)	(2)	(3)
	Patent_F	*Patent_I*	*Patent_Per_Emp*
Fluidity x *Growth*	-0.0235***	-0.0213***	-0.0229***
	(0.006)	(0.006)	(0.004)
Fluidity	0.2757***	0.2605***	0.3025***
	(0.057)	(0.056)	(0.044)
Growth	0.1340***	0.1188***	0.1280***
	(0.040)	(0.040)	(0.030)
Size	0.3680***	0.3724***	0.0819***
	(0.013)	(0.013)	(0.009)
Ppe	-0.0758	-0.1363	-0.1607
	(0.118)	(0.118)	(0.101)
Capex	0.0864	-0.1338	0.2995
	(0.217)	(0.216)	(0.193)
Leverage	-0.0553**	-0.0531**	-0.1113***
	(0.025)	(0.026)	(0.024)
Roa	0.0221	-0.0248	0.0680**
	(0.027)	(0.026)	(0.028)
TobinsQ	0.0258***	0.0164***	0.0192***
	(0.005)	(0.006)	(0.006)
Cash	-0.1281***	-0.2131***	0.0983**
	(0.044)	(0.044)	(0.042)
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R^2	0.15	0.17	-0.04
Observations	23103	23103	22657

Table 9. Competition and Financial Constraint

This table presents only the results for the second-stage regression estimates. We use import tariff and foreign exchange rate and their interactions with *Fin_Constraint* as the IVs for *Fluidity* and *Fluidity* x *Fin_Constraint* respectively. Tariff rates re available only for firms in the manufacturing industry (SIC 2000-3999) and hence the smaller sample size. In untabulated results, we observed similar outcomes using Pooled OLS regression without IVs. Standard errors were clustered by firms and reported in the parentheses. Year and 2-digit SIC industry fixed effects are included in all specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For detailed definitions of variables, please refer to Appendix A1.

	(1)	(2)	(3)
	Patent_F	*Patent_I*	*Patent_Per_Emp*
Fluidity x *Fin_Constraint*	0.0241***	0.0295***	0.0062
	(0.006)	(0.006)	(0.005)
Fluidity	0.1346***	0.1227***	0.1813***
	(0.034)	(0.035)	(0.024)
Fin_Constraint	-0.0730**	-0.0526	-0.1119***
	(0.036)	(0.036)	(0.032)
Size	0.3900***	0.4028***	0.0789***
	(0.019)	(0.018)	(0.012)
Ppe	-0.1575	-0.2148**	-0.2400***
	(0.102)	(0.101)	(0.085)
Capex	0.4097**	0.1830	0.5808***
	(0.170)	(0.169)	(0.155)
Leverage	-0.0794***	-0.0760***	-0.1352***
	(0.022)	(0.022)	(0.021)
Roa	0.0241	-0.0206	0.0631**
	(0.023)	(0.022)	(0.027)
TobinsQ	0.0351***	0.0255***	0.0273***
	(0.005)	(0.005)	(0.005)
Cash	-0.0569*	-0.1425***	0.1617***
	(0.034)	(0.034)	(0.036)
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R^2	0.27	0.28	0.06
Observations	23103	23103	22657

Table 10. Competition and the Current State of Technology

This table presents only the results for the second-stage regression estimates. We use import tariff and foreign exchange rate and their interactions with *Dist_Tech (Dist_Tech_l5)* as the IVs for *Fluidity* and *Fluidity* x *Dist_Tech (Dist_Tech_l5)* respectively. Tariff rates are available only for firms in the manufacturing industry (SIC 2000-3999) and hence the smaller sample size. In untabulated results, we observed similar outcomes using Pooled OLS regression without IVs. Standard errors were clustered by firms and reported in the parentheses. Year and 2-digit SIC industry fixed effects are included in all specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For detailed definitions of variables, please refer to Appendix A1.

	(1)	(2)	(3)	(4)
	Patent_F	*Patent_I*	*Patent_F*	*Patent_I*
Fluidity x *Dist_Tech*	0.0120***	0.0109***		
	(0.004)	(0.004)		
Dist_Tech	0.0931***	0.0869***		
	(0.017)	(0.018)		
Fluidity x *Dist_Tech_l5*			0.0136**	0.0125**
			(0.006)	(0.006)
Dist_Tech_l5			0.1106***	0.1027***
			(0.022)	(0.022)
Fluidity	0.0941***	0.1036***	0.0855**	0.1073***
	(0.025)	(0.024)	(0.036)	(0.035)
Size	0.2494***	0.2511***	0.2404***	0.2318***
	(0.010)	(0.010)	(0.010)	(0.010)
Ppe	-0.1718**	-0.2298***	-0.1126	-0.1614**
	(0.072)	(0.072)	(0.079)	(0.078)
Capex	0.5084***	0.2654**	0.4716***	0.2334*
	(0.115)	(0.111)	(0.127)	(0.123)
Leverage	-0.0755***	-0.0704***	-0.0748***	-0.0688***
	(0.016)	(0.016)	(0.016)	(0.016)
Roa	0.0176	-0.0282*	0.0302*	-0.0121
	(0.016)	(0.016)	(0.018)	(0.017)
TobinsQ	0.0265***	0.0163***	0.0254***	0.0144***
	(0.003)	(0.003)	(0.003)	(0.003)
Cash	0.0238	-0.0661***	0.0216	-0.0623**
	(0.022)	(0.022)	(0.025)	(0.025)
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Adj. R^2	0.49	0.49	0.45	0.49
Observations	22246	22246	22751	22751

Table 11. Competition and Innovation Quality

This table presents the pooled OLS regression and the second-stage results from 2SLS regressions for *Fluidity* and its interactions with *ROA_l5, Growth, Dist_Teeceh_l5,* and *Fin_Constraint*. The instrumental variables used are import tariffs and foreign exchange rates and their respective interactions with the corresponding firm and industry characteristic variables above. Tariff data is available only for firms in the manufacturing industry (SIC 2000-3999). Panel A reports the results for the dependent variable of *Patent_Cites* and Panel B reports that for *Patent_Value.* Standard errors were clustered by firms and reported in the parentheses. Year and 2-digit SIC industry fixed effects are included in all specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For detailed definitions of variables, please refer to Appendix A1.

Panel A	Pooled OLS	Second Stage IV Estimation				
Patent_Cites	(1)	(2)	(3)	(4)	(5)	(6)
Fluidity	0.0266***	0.2081***	0.2070***	0.3668***	0.1536***	0.1852***
	(0.005)	(0.038)	(0.038)	(0.075)	(0.029)	(0.049)
Fluidity x *ROA_l5*			0.0106			
			(0.008)			
ROA_l5			-0.0639			
			(0.067)			
Fluidity x *Growth*				-0.0281***		
				(0.008)		
Growth				0.1463***		
				(0.054)		
Fluidity x *Dist_Tech_l5*					0.0117	
					(0.008)	
Dist_Tech_l5					0.1369***	
					(0.049)	
Fluidity x *Fin_Constraint*						0.0297***
						(0.009)
Fin_Constraint						-0.1623***
						(0.055)
Firm-Level Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.30	0.18	0.17	0.08	0.33	0.17
Observations	43979	23105	23105	23105	22753	23105

Panel B	Pooled OLS	Second Stage IV Estimation				
Patent_Value	(1)	(2)	(3)	(4)	(5)	(6)
Fluidity	0.0207***	0.2077***	0.2082***	0.3692***	0.1446***	0.1461***
	(0.006)	(0.043)	(0.043)	(0.082)	(0.032)	(0.054)
Fluidity x *ROA_l5*			0.0158***			
			(0.006)			
ROA_l5			-0.1198**			
			(0.052)			

Table 11 Continued

Fluidity x *Growth*				-0.0286***		
				(0.008)		
Growth				0.1516***		
				(0.058)		
Fluidity x *Dist_Tech_l5*					0.0102	
					(0.008)	
Dist_Tech_l5					0.1768***	
					(0.050)	
Fluidity x *Fin_Constraint*						0.0518***
						(0.010)
Fin_Constraint						-0.0640
						(0.057)
Firm-Level Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.37	0.35	0.35	0.26	0.51	0.36
Observations	43977	23103	23103	23103	22751	23103

Table 12. Fluidity & R&D Efficiency

This table presents the pooled OLS regression and the second-stage results from 2SLS regressions for *Fluidity* and its interactions with *ROA_l5*, *Growth*, *Dist_Tech_l5*, and *Fin_Constraint*. The instrumental variables used are import tariffs and foreign exchange rates and their respective interactions with the corresponding firm and industry characteristic variables above. Tariff data is available only for firms in the manufacturing industry (SIC 2000-3999). Standard errors were clustered by firms and reported in the parentheses. Year and 2-digit SIC industry fixed effects are included in all specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For detailed definitions of variables, please refer to Appendix A1.

	Pooled OLS	Second Stage IV Estimation				
R&D_Efficiency	(1)	(2)	(3)	(4)	(5)	(6)
Fluidity	-0.0032***	-0.0117*	-0.0120*	-0.0177*	-0.0119*	-0.0152*
	(0.001)	(0.007)	(0.007)	(0.010)	(0.006)	(0.009)
Fluidity x *ROA_l5*			-0.0013			
			(0.001)			
ROA_l5			0.0173			
			(0.014)			
Fluidity x *Growth*				0.0009		
				(0.001)		
Growth				0.0009		
				(0.011)		
Fluidity x *Dist_Tech_l5*					-0.0022*	
					(0.001)	
Dist_Tech_l5					0.0309***	
					(0.008)	
Fluidity x *Fin_Constraint*						0.0035
						(0.002)
Fin_Constraint						-0.0097
						(0.012)
Size	0.0048**	0.0033	0.0027	0.0044	-0.0130***	0.0051
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Ppe	0.0892***	0.0595	0.0579	0.0497	0.0627	0.0563
	(0.031)	(0.046)	(0.046)	(0.050)	(0.044)	(0.047)
Capex	-0.0182	0.1009	0.1145*	0.1185*	0.1210**	0.1162*
	(0.044)	(0.062)	(0.063)	(0.067)	(0.060)	(0.065)
Leverage	-0.0193***	-0.0260***	-0.0243***	-0.0284***	-0.0213***	-0.0255***
	(0.004)	(0.008)	(0.007)	(0.008)	(0.007)	(0.008)
Roa	0.0053	0.0017		0.0009	0.0101	0.0032
	(0.004)	(0.007)		(0.007)	(0.007)	(0.007)
TobinsQ	0.0026***	0.0033***	0.0037***	0.0036***	0.0020**	0.0035***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Cash	0.0070	0.0183**	0.0184*	0.0204**	0.0221**	0.0212**
	(0.005)	(0.009)	(0.010)	(0.010)	(0.009)	(0.010)
Constant	0.1482***					
	(0.016)					
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.11	0.00	0.00	-0.00	0.02	-0.00
Observations	21573	15423	15423	15423	15340	15423

Appendix Table A1

Main LHS Variables	
Patent_F	Natural log of one plus the number of patents filed in year *t*
Patent_I	Natural log of one plus the number of patents granted filed in year *t*
Patent_Per_Emp	Natural log of one plus the number of patents filed in year *t* divided by the total number of employees in year *t*
R&D_Efficiency	Natural log of one plus the average number of patents filed from *t* to *t*+2 divide by the mean R&D expenditures in the same period.
Patent_Cites	Natural log of one plus total citation received from all patents filed in year *t*
Patent_Value	Natural log of one plus total patent values received from all patents filed in year *t*
R&D_Intensity	R&D expenditures divided by the total assets at the beginning of the fiscal year
Main RHS Variable	
Fluidity	A measure for competitive threat as in Hoberg *et al*. (2014)
Pctcomp	A measure for competitive threat as in Li *et al.* (2013)
HHI	Herfindahl-Hirschman Index calculated based on firms' annual sales for each 2-digit SIC industry from the Compustat database
CR_8	The eight-firm concentration ratio is calculated as the total sales of the 8 largest firms in terms of annual sales from the Compustat database scaled by industry annual sales for each 2-digit SIC industry.
Roa_l5	The average return over assets over the past 5 years.
Fin_Constraint	A dummy equal to one if the WW index is higher than that of the median firm in the same 2-digit SIC and year and zeroes otherwise. WW index is calculated as -0.91 x (*ib* + *dp*)/*ta* - 0.062 x *div* + 0.021 x *dltt*/*ta* - 0.044 x ln(*ta*) + 0.102 x *Industry Sales Growth* - 0.035 x *Sales Growth* where *ib* is income before extraordinary items; *dp* is depreciation and amortization; ta is total assets; *div* is a dummy for dividend payout; *dltt* is the total long-term debt; *Industry Sales Growth* is the average sales growth in the 3-digit SIC industry; and *Sales growth* is the firm's growth in sales.
Dist_Tech	It is the difference between the total number of patents filed by a firm and the 3-digit SIC industry mean and normalized by the industry mean.
Dist_Tech_l5	It is the difference between the total number of patents filed by a firm in the previous 5-year and the 3-digit SIC industry mean over the same period and normalized by the corresponding 5-year industry mean.
Growth	The average sales growth over the past 5 years within a 3-digit SIC industry
Controls	
Size	Natural log of total assets
Cash	Cash holdings divided by total assets at the beginning of the fiscal year
Ppe	Total property, plant, and equipment divided by total assets at the beginning of the fiscal year
Capex	Capital expenditures divided by the total assets at the beginning of the fiscal year
Leverage	Total Debt divided by the total assets at the beginning of the fiscal year
Roa	Net income divided by the total assets at the beginning of the fiscal year
TobinsQ	Book value of assets plus the market value of equity minus the book value of equity and normalized by total asset

Appendix Table A2. Pairwise Correlation Table (* indicates statistical significance at 5% level)

	Variables Names	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	*Patent_F*	1.0000						
(2)	*Patent_I*	0.8717*	1.0000					
(3)	*Patent_Per_Emp*	0.6989*	0.5549*	1.0000				
(4)	*R&D_Efficiency*	0.4355*	0.3151*	0.5114*	1.0000			
(5)	*Patent_Cites*	0.8893*	0.7382*	0.6752*	0.4643*	1.0000		
(6)	*Patent_Value*	0.9159*	0.8123*	0.5336*	0.2926*	0.8598*	1.0000	
(7)	*R&D_Intensity*	0.1496*	0.1581*	0.3714*	-0.1078*	0.1515*	0.1057*	1.0000
(8)	*HHI*	-0.1490*	-0.1538*	-0.1883*	0.0541*	-0.1351*	-0.1227*	-0.2351*
(9)	*CR_8*	-0.1549*	-0.1614*	-0.2221*	0.0937*	-0.1345*	-0.1326*	-0.3412*
(10)	*Fluidity*	0.0834*	0.0830*	0.2549*	-0.1157*	0.0716*	0.0676*	0.4207*
(11)	*Pctcomp*	0.0203*	-0.0251*	0.1484*	0.0620*	0.1048*	-0.0123	0.2190*
(12)	*Roa_l5*	0.0411*	0.0468*	-0.0447*	0.0588*	0.0322*	0.0532*	-0.2651*
(13)	*Fin_Constraint*	-0.1487*	-0.1430*	0.0884*	-0.1118*	-0.1144*	-0.1782*	0.3311*
(14)	*Dist_Tech*	0.5501*	0.5469*	0.2240*	0.2178*	0.4647*	0.5421*	-0.0153*
(15)	*Dist_Tech_l5*	0.5206*	0.5468*	0.1965*	0.1895*	0.4387*	0.5205*	-0.0223*
(16)	*Growth*	-0.0094*	0.0047	0.0886*	-0.1198*	-0.0361*	-0.0132*	0.3289*
(17)	*Size*	0.3525*	0.3581*	-0.0328*	0.0371*	0.2655*	0.4183*	-0.3383*
(18)	*Cash*	0.0198*	-0.0046	0.1506*	-0.0363*	0.0418*	-0.0003	0.2841*
(19)	*Ppe*	-0.0872*	-0.1132*	-0.1555*	0.0927*	-0.0705*	-0.0579*	-0.2477*
(20)	*Capex*	-0.0473*	-0.0832*	-0.0435*	0.0332*	-0.0159*	-0.0366*	-0.0914*
(21)	*Leverage*	-0.0449*	-0.0556*	-0.0856*	-0.0355*	-0.0405*	-0.0277*	-0.0241*
(22)	*Roa*	0.0366*	0.0446*	-0.0758*	0.0493*	0.0258*	0.0572*	-0.3313*
(23)	*TobinsQ*	0.0801*	0.0461*	0.1659*	-0.0041	0.1049*	0.0883*	0.3704*
		(8)	(9)	(10)	(11)	(12)	(13)	(14)
(8)	*HHI*	1.0000						
(9)	*CR8*	0.8377*	1.0000					
(10)	*Fluidity*	-0.2010*	-0.2944*	1.0000				
(11)	*Pctcomp*	-0.1130*	-0.1479*	0.1457*	1.0000			
(12)	*Roa_l5*	0.0761*	0.1129*	-0.1887*	-0.0116	1.0000		
(13)	*Fin_Constraint*	-0.1939*	-0.2675*	0.2523*	0.1584*	-0.1705*	1.0000	
(14)	*Dist_Tech*	0.0122*	0.0138*	-0.0050	-0.0487*	0.0444*	-0.1666*	1.0000
(15)	*Dist_Tech_l5*	0.0155*	0.0209*	-0.0227*	-0.0614*	0.0488*	-0.1707*	0.9015*
(16)	*Growth*	-0.1820*	-0.2998*	0.3892*	0.0316*	-0.1481*	0.3459*	-0.0237*
(17)	*Size*	0.1274*	0.1531*	-0.0774*	-0.2656*	0.2027*	-0.5973*	0.3515*
(18)	*Cash*	-0.1253*	-0.1761*	0.2423*	0.1416*	-0.3465*	0.2426*	-0.0465*
(19)	*Ppe*	0.1593*	0.1987*	0.0104*	-0.0794*	-0.0068	-0.1513*	0.0028
(20)	*Capex*	0.0392*	0.0409*	0.1047*	0.0280*	-0.1271*	0.0162*	-0.0237*
(21)	*Leverage*	0.0259*	0.0201*	0.0440*	-0.0770*	-0.2517*	-0.0022	-0.0039
(22)	*Roa*	0.0896*	0.1354*	-0.2183*	-0.0430*	0.5765*	-0.2129*	0.0524*
(23)	*TobinsQ*	-0.0965*	-0.1394*	0.1727*	0.1188*	-0.1825*	0.1557*	0.0119*
		(15)	(16)	(17)	(18)	(19)	(20)	(21)
(15)	*Dist_Tech_l5*	1.0000						
(16)	*Growth*	-0.0282*	1.0000					
(17)	*Size*	0.3537*	-0.1602*	1.0000				
(18)	*Cash*	-0.0585*	0.1708*	-0.1971*	1.0000			
(19)	*Ppe*	0.0017	-0.0974*	0.1875*	-0.0192*	1.0000		
(20)	*Capex*	-0.0330*	0.0031	0.2198*	0.0201*	0.7003*	1.0000	
(21)	*Leverage*	-0.0023	0.0190*	0.0667*	0.2955*	0.2723*	0.2984*	1.0000
(22)	*Roa*	0.0560*	-0.1528*	0.2533*	-0.5119*	-0.0469*	-0.2039*	-0.4336*
(23)	*TobinsQ*	0.0031	0.1380*	-0.1824*	0.2604*	-0.0684*	0.0738*	0.1080*
		(22)	(23)					
(22)	*Roa*	1.0000						
(23)	*TobinsQ*	-0.2604*	1.0000					

Appendix Table A3. Pooled OLS Regression - Fluidity and Pctcomp with *Industry* x *Year* Fixed Effects

This table reports the OLS regression results for the two text-based firm-level competition measures, *Fluidity* and *Pctcomp*. Standard errors were clustered by firms and reported in the parentheses. Year, 2-digit SIC industry, and year x industry fixed effects are included in all specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For detailed definitions of variables, please refer to Appendix A1.

	(1) *Patent_F*	(2) *Patent_I*	(3) *Patent_Per_Emp*	(4) *R&D_Intensity*	(5) *Patent_F*	(6) *Patent_I*	(7) *Patent_Per_Emp*	(8) *R&D_Intensity*
Fluidity	**0.0202*** **	**0.0196*** **	**0.0686*** **	**0.0134*** **				
	(0.004)	**(0.004)**	**(0.004)**	**(0.000)**				
Pctcomp					**0.0817*** **	**0.0586** **	**0.1355*** **	**0.0170*** **
					(0.024)	**(0.023)**	**(0.023)**	**(0.002)**
Size	0.2677***	0.2632***	0.0579***	-0.0129***	0.3429***	0.3382***	0.0709***	-0.0020***
	(0.009)	(0.009)	(0.005)	(0.001)	(0.013)	(0.013)	(0.007)	(0.000)
Ppe	-0.3767***	-0.4089***	-0.3345***	-0.0427***	-0.3894***	-0.3988***	-0.2520***	-0.0279***
	(0.053)	(0.053)	(0.041)	(0.004)	(0.083)	(0.083)	(0.054)	(0.003)
Capex	0.5076***	0.3498***	0.4668***	-0.0404***	0.6130***	0.4969***	0.5427***	0.0213***
	(0.072)	(0.069)	(0.072)	(0.008)	(0.111)	(0.110)	(0.090)	(0.007)
Leverage	-0.0600***	-0.0541***	-0.0960***	-0.0141***	-0.1826***	-0.1725***	-0.2150***	-0.0133***
	(0.008)	(0.007)	(0.009)	(0.001)	(0.027)	(0.027)	(0.022)	(0.002)
Roa	-0.0344***	-0.0564***	-0.0102	-0.0354***	-0.2190***	-0.2560***	-0.3332***	-0.0352***
	(0.009)	(0.008)	(0.013)	(0.002)	(0.045)	(0.050)	(0.056)	(0.006)
TobinsQ	0.0319***	0.0213***	0.0278***	0.0098***	0.0853***	0.0598***	0.0758***	0.0097***
	(0.003)	(0.002)	(0.003)	(0.001)	(0.012)	(0.011)	(0.012)	(0.001)
Cash	0.0427***	-0.0057	0.1756***	0.0042*	0.0718*	0.0330	0.2981***	0.0248***
	(0.012)	(0.012)	(0.018)	(0.002)	(0.037)	(0.034)	(0.044)	(0.004)
Constant	-0.9777***	-0.8732***	-0.2290***	0.0515***	-1.3591***	-1.2522***	-0.0403	0.0332***
	(0.046)	(0.046)	(0.036)	(0.004)	(0.070)	(0.070)	(0.049)	(0.004)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry x Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.36	0.36	0.28	0.49	0.40	0.40	0.24	0.32
Observations	43939	43939	42871	71771	27540	27540	27038	30519

Appendix Table A4. Rank Regression

This table reports the OLS regression results for the rank variable, *FluidityRank*, that takes the number from 1 to 10 with 1 being in the decile with the lowest competitive threats and 10 the highest. The rank variable from column (1) to (4) are sorted into deciles based on the *fluidity* of the full sample and (5) to (8) are sorted in the same way but within each 2-digit SIC industry. Standard errors were clustered by firms and reported in the parentheses. Year and 2-digit SIC industry fixed effects are included in all specifications. Coefficients with *, **, or *** indicate a significance level at 0.10, 0.05, and 0.01 respectively. For detailed definitions of variables, please refer to Appendix A1.

	(1) *Patent_F*	(2) *Patent_I*	(3) *Patent_Per_Emp*	(4) *R&D_Intensity*	(5) *Patent_F*	(6) *Patent_I*	(7) *Patent_Per_Emp*	(8) *R&D_Intensity*
FluidityRank	**0.0221*** **	**0.0212*** **	**0.0700*** **	**0.0132*** **				
	(0.004)	**(0.005)**	**(0.004)**	**(0.000)**				
FluidityRank(SIC2)					**0.0184*** **	**0.0178*** **	**0.0553*** **	**0.0100*** **
					(0.003)	**(0.004)**	**(0.003)**	**(0.000)**
Size	0.2667***	0.2619***	0.0595***	-0.0129***	0.2663***	0.2614***	0.0584***	-0.0131***
	(0.009)	(0.009)	(0.005)	(0.001)	(0.009)	(0.009)	(0.005)	(0.001)
Ppe	-0.3762***	-0.4114***	-0.3425***	-0.0456***	-0.3785***	-0.4135***	-0.3545***	-0.0480***
	(0.052)	(0.052)	(0.040)	(0.004)	(0.052)	(0.052)	(0.040)	(0.004)
Capex	0.4894***	0.3149***	0.4711***	-0.0345***	0.4901***	0.3154***	0.4847***	-0.0334***
	(0.070)	(0.067)	(0.070)	(0.008)	(0.070)	(0.067)	(0.070)	(0.008)
Leverage	-0.0579***	-0.0516***	-0.0985***	-0.0141***	-0.0584***	-0.0521***	-0.1008***	-0.0146***
	(0.007)	(0.007)	(0.009)	(0.001)	(0.007)	(0.007)	(0.009)	(0.001)
Roa	-0.0324***	-0.0546***	-0.0138	-0.0362***	-0.0330***	-0.0553***	-0.0170	-0.0373***
	(0.009)	(0.008)	(0.013)	(0.002)	(0.009)	(0.008)	(0.013)	(0.002)
TobinsQ	0.0311***	0.0201***	0.0279***	0.0097***	0.0313***	0.0203***	0.0287***	0.0099***
	(0.003)	(0.002)	(0.003)	(0.001)	(0.003)	(0.002)	(0.003)	(0.001)
Cash	0.0406***	-0.0068	0.1771***	0.0053**	0.0420***	-0.0055	0.1834***	0.0072***
	(0.012)	(0.012)	(0.018)	(0.002)	(0.012)	(0.012)	(0.018)	(0.002)
Constant	-0.9550***	-0.8456***	-0.1562***	0.0708***	-0.9293***	-0.8212***	-0.0608*	0.0899***
	(0.045)	(0.045)	(0.034)	(0.004)	(0.043)	(0.043)	(0.031)	(0.004)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.36	0.36	0.27	0.47	0.36	0.36	0.26	0.47
Observations	43977	43977	42917	71845	43977	43977	42917	71845

Chapter 5. Conclusion

Empirical studies in corporate finance often serve to validate theoretical predictions for a better understanding of real-world phenomena. Our first essay empirically explores the potential impact on disclosure behavior when firms' top executives become legally accountable under SOX for misreporting information. Related literature on the relation between litigation risk and voluntary disclosure yield mixed evidence at best and the impact on firms' mandatory disclosure from SOX is a priori even more ambiguous. In this essay, we first quantify in large scale the tones of firms' mandatory disclosure and examine whether holding key personnel might impact the way firms disclose information. We observe a structural change in the distribution of filing tones around SOX and find that firms become more conservative in reporting good or bad news. Interestingly, the investors show a stronger reaction to per unit tone change after SOX which are not entirely caused by the systematic changes in tone distribution over time. Our evidence supports the view that imposing legal accountability on top executives might reduce the firm's incentive to disclose information, however, any disclosure made after SOX might contain more information.

Our second essay examines the relation between VCs and firm innovation. While prior literature mostly demonstrates a positive correlation between the two, we provide empirical evidence from a different perspective – what could happen to a firm's innovation when VCs exit? We first document that VC-backed firms experience a greater dip in R&D intensity after IPO exit and such impact is more pronounced for firms with higher pre-IPO VC involvement and less for firms with higher institutional holdings. Potentially, VCs exit with all the positive influences that they brought into the firms at initial stages and firms do not seem to retain high R&D intensity when VCs leave. The presence of high institutional holders, however, might take over the monitoring role of VCs and continue promoting innovation after VCs leave. We also present evidence that VCs pre-select into more innovative firms at the initial stage, which is potentially the main source of endogeneity. Nonetheless, we hope to present some insights on the impact of VCs during different phases of a firm's life cycle and provide some important policy implications aimed at promoting innovation.

Our last essay revisits the debatable relation between market competition and firm innovation. Reaching a unanimous consensus among scholars has been challenging for a variety of reasons. The complexity in market structures and characteristics of innovation often leads to drastically different theoretical assumptions and predictions. Empirically, this intricacy is further confounded by endogeneity issues and the quality of competition measures. This essay attempts to address the empirical challenges by utilizing two new forward-looking text-based measures on competitive threats which are arguably better proxies for market competition at firm-level. We then use import tariffs and exchange rates as IVs to address potential endogeneity. We first document that a positive causal relation exists between competition and innovation and this result is further confirmed by the natural experiment of exogenous shock in tariff rates. We then find that a better industry prospect could encourage more innovation, but such effect is less prominent in the presence of greater competition potentially because of lower expected profits. We also provide supports that the positive impact of heightened competition is more pronounced for firms with higher financial constraints and larger patent pool. Lastly, we observe some evidence that competition may reduce firms' innovation efficiency.

The findings from this thesis not only shed light on some of the most contentious relations between innovation and VCs, innovation and competition, and legal accountability and information disclosure but also provides insights for policymakers aimed at improving domestic innovation and information transparency. Apart from complementing existing literature, the findings from this dissertation could also serve as the steppingstones for related prospective research.

References

Gilbert, R. (2006). Looking for Mr. Schumpeter: Where Are We in the Competition-Innovation Debate? *Innovation policy and the economy*, *6*, 159-215.

Hoberg, G., Phillips, G., & Prabhala, N. (2014). Product market threats, payouts, and financial flexibility. *The Journal of Finance*, *69*(1), 293-324.

Kamien, M. I., & Schwartz, N. L. (1975). Market structure and innovation: A survey. *Journal of economic literature*, *13*(1), 1-37.

Li, F., Lundholm, R., & Minnis, M. (2013). A measure of competition based on 10-K filings. *Journal of Accounting Research*, *51*(2), 399-436.