Essays on Financial Coordination

Essays on Financial Coordination

Essays over de financieel coördinatie

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by the command of
rector magnificus

Prof.dr. R.C.M.E Engels

and in accordance with the decision of the Doctorate Board.

The public defense shall be held on
Thursday, 17 January 2019 at 09:30 hrs

by

Lingtian Kong
born in China

Erasmus University Rotterdam

Doctoral Committee

Doctoral dissertation supervisor:

Prof.dr. M.J.C.M. Verbeek

Other members:

Prof.dr. W.B. Wagner
Prof.dr. S. van Bekkum
Prof.dr. A.J. Menkveld

Co-supervisors:

dr. D.G.J. Bongaerts
dr. M.A. Van Achter

Erasmus Research Institute of Management – ERIM
The joint research institute of the Rotterdam School of Management (RSM)
and the Erasmus School of Economics (ESE) at the Erasmus University Rotterdam
Internet: http://www.erim.eur.nl

ERIM Electronic Series Portal: http://repub.eur.nl/

ERIM PhD Series in Research in Management, 443
ERIM reference number: EPS-2019-443-F&A
ISBN 978-90-5892-534-3

Design: Ikon Images, www.ikon-images.com

This publication (cover and interior) is printed by Tuijtel on recycled paper, BalanceSilk®.
The ink used is produced from renewable resources and alcohol free fountain solution.
Certifications for the paper and the printing production process: Recycle, EU Ecolabel, FSC®, ISO14001.
More info: www.tuijtel.com

Acknowledgments

I am deeply indebted to many people who have supported me over the years.

First of all, I would like to thank Dion Bongaerts, who as my supervisor has offered me critical support at various critical points in my PhD study. Dion has not only been offering me excellent advice on research itself, but also on how to navigate the academic career. His extremely high efficiency, the ability to maintain a rational mind under pressure, and his care for the social welfare will always be an inspiration for me.

I would also like to thank my promotor, Marno Verbeek, who has offered much guidance and support over the years. I benefitted tremendously from his expertise in econometrics and institutional investors, which has become important parts of my knowledge base.

My sincere thanks also go to Mark Van Achter, who as my other supervisor has dedicated his time and energy to my growth over the years. His passion in market microstructure has inspired me to delve into this fascinating topic.

I also greatly appreciate members of my inner dissertation committee—Sjoerd van Bekkum, Albert Menkveld and Wolf Wagner. They have been providing selfless help during my PhD career. They also spared precious time in a busy season of the year to read my manuscript so that I can defend soon. I am also immensely grateful to other members of my dissertation committee—Mathijs van Dijk, Hans Degryse, and Peter Roosenboom—for their kind help during my PhD study and informative conversations on both research and academic career. I have been fortunate to have them as mentors during my PhD study.

I have been very lucky to work with a group of supportive and considerate colleagues. I would like to express my appreciation and thanks to Mathijs Cosemans, Sarah Draus, Mintra Dwarkasing, Marc Gabarro, Ying Gan, Egemen Genc, Abe de Jong, Yigitcan Karabulut, Thomas Lambert, Melissa Lin, Mike Mao, Anjana Rajamani, Claus Schmitt, Dirk Schoenmaker, Fabrrizio Spargoli, Marta Szymanowska, Francisco Urzua Infante, Yan Wang and Ran Xing for their priceless support. I would also like to extend my appreciation and gratitude to Myra and Flora for assisting me tremendously over the years, especially during the job market and the defense trajectory. Bálint, Kim, Miho and Natalija from ERIM PhD office and Yvonne from the HR office also provided excellent support to make this thesis possible. Also, I thank my fellow PhD cohorts including Aleks, Darya, Eden, Gelly, Jose, Marina, Pooyan, Rex, Rogier, Roy, Romulo, Shuo, Teng, Teodor, Vlado, Yu Wang, Xiao Xiao and Yuhao for the pleasant time and memorable conversations that we have had together.

In addition, I would like to express my gratitude towards my pre-EUR mentors Richard McGehee and Nicolai Krylov for their guidance during my early pursuit of mathematical and economic sciences. I would also like to share this achievement with Chun Guo and Dr. Yu Wang who served as witnesses and good companions when I first embarked on this journey.

Last but not the least, I would like to thank my family: my father for being a role model for his achievements as well as his cultivation of my intellectual pursuit from an early age; my mum for being a caring mother, a good listener as well as a role model. Their unconditional support and selfless provision kept me going through all the turbulence during my career.

Lingtian Kong

Nov. 21st, 2018

Contents

Chapter 1

Introduction

The financial history for the past two decades has seen three heart throbbing episodes of ebb and flow. From 1998 to 2001, the flourishing information technology boosted the equity market of the US and the world to a historic high and then crashed in a few months. From 2002 to 2008, the real estate bubble developed in the aftermaths of the previous boost only to bust violently in the end as a banking crisis. Different in nature from the previous two episodes but similarly intriguing is the advent of the "high frequency traders", the population of whom rapidly increased from 2006 to 2009 and revolutionized how financial securities are traded. Interestingly, financial journalist Michael Lewis published three books to document each of these three episodes of the financial market, indirectly pointing to the importance to these seemingly unrelated events.

In reverse chronological order, the three chapters of this PhD thesis explore the commonality of these three periods of history. In all three episodes, to some extent, the invisible hand fails to motivate the players to act optimally for the common good of the society.

In the classical economic framework pioneered by works of Adam Smith, each agent only needs to optimize her own welfare. But in the end by virtue of the invisible hand, the resources are allocated in a way that is socially optimal. No deliberate coordination is required as any agents that deviate will be disciplined by the market. But during the three episodes mentioned earlier, coordination seemed to be crucial since market discipline did not always serve to guide the behavior of the players towards societal optimum. The lack of coordination on other hand, may lead the players to deviate from such optimum.

In Chapter 2, in the context of trading speed competition, I investigate a simple form of miscoordination. In the trading platforms, there are two groups of players that fulfil complementary functions: market makers and market takers. Within each group, on the other hand, players compete with each other in trading speed so as to be the winner in a winner-take-all trade rush. As a result, there are not only mutually beneficial strategic complementarities between a maker and a taker, but there are also zero-sum strategic substitutions between any two makers or any two takers. Due to the lack of coordination, none of the two types of externalities, the complementarities or the substitution can be fully internalized by individual market participants, hence the invisible hand can under- or over- allocate resources compared to the social optimum. In this setup, in particular, any form of coordination is not possible in the anonymous trading platforms. Thus an arms-race like speed competition is a likely outcome amongst the makers or amongst the takers. But such speed competition is not without any cost. In the modern times, trade speed can only be achieved through expensive investments in information technologies.

In this chapter, I demonstrate that whether the cost outweighs the benefit depends crucially on how the marginal contribution of trading speed to the gains from trade changes with trading speed itself. I show theoretically that when it is decreasing, the market discipline by the technological expenses borne by the traders themselves is not sufficient to curb the speed competition – the traders would over-invest in speed and exhibit "arms-race behavior". In the end of this chapter, I provide a micro-foundation for this decreasing gain from trade of trading speed, based on the classical Merton portfolio re-balancing problem.

As shown in Chapter 2, market discipline can already run into difficulty in a simple setup with strategic substitutions and complementarities. But in most modern economic relationships, there are still other types of frictions that occur alongside the strategic substitution. For example, in Chapter 3, I consider the innate conflict of interest between debt holders and the equity holder. Given limited liability, equity holders have incentives to engage in "risk shifting": she would take on excessive risk so that she's the one that benefits when the realized outcome is good, but her creditors are the ones to suffer the downside losses. In the context of commercial banks, the market discipline in place to address this conflict of interest is by ways of short-term, demandable debts. When the creditors sense that the equity holder (represented by the bank manager due to fiduciary duty under the common law) is engaging in such a risk shifting strategy, they would withdraw their deposits before the bank asset maturity date, incurring liquidation costs on the equity buffer, punishing the equity holder. As a result, *ex ante*, the equity holder should be discouraged to take on excessive risks.

This market discipline did not fulfill its function in the period leading up to the 2008 financial crisis. The banks still took on too much risk. Demandable debts were present in the form of money market instruments. Government guarantees did not exist for these instruments, as the deposit insurance exists for the retail deposits, so that the market discipline was not neutralized by the government intervention. Had market discipline have worked, the runs of the money market funds should have occurred once the bad news of the sub-prime loans surfaced. But this was not the case and the runs occurred much later. This led me to question a long-held belief: why runs did not work as a disciplining device? As a consequence of the discussion above, this question is very important as if under some circumstances the answer is no, market discipline by the short-term debt holders loses its incentive compatibility premise and foundation under those circumstances.

In Chapter 3, I show theoretically that the answer is indeed no sometimes. In particular, higher risk taken by the bank may in fact discourage runs when the bank equity buffer is low. I demonstrate that this is a result of a common fact in finance: the higher the asset risk, the higher also is the asset upside payoff. The short term debt holder, known as "depositors" in this chapter, effectively becomes the residual claimant of the asset at its maturity. Thus a higher asset risk means the asset would pay more when it succeeds, which benefits the depositors if they do not run and hold their deposits until maturity. This occurs only when equity buffer is low, because in this case, it is more likely that the bank would be insolvent in the end, which is the circumstance under which the depositors become effective residual claimants. This mechanism is particularly relevant for banks issuing runnable deposits, as I show that runs are able to decrease the equity level, creating the necessary condition of the effect mentioned earlier. This theory explains well the risk management practice of counting high payoff assets as substitutes for capital amongst commercial banks.

In Chapter 3, I assume that the manager represents wholeheartedly the best interest of the equity holders. But because of agency problems, this is hardly the case in most corporations. Managers (CEOs) may strive to maximize their own welfare instead of that of the shareholders or simply do not work hard enough (a behavior known as "shirking"). The equity holder may directly tell the CEO what to do ("voice" or "monitoring") or simply sell their shares to depress the share prices so the company more likely becomes a target of acquisition ("exit"). This philosophy of market discipline underlies the academic discipline of corporate governance.

But in order to tell the CEO what to do, the shareholders first have to be able to evaluate whether the CEO has done a good job. This can be difficult, because the CEO is hired as an expert which the shareholders are not. The performance evaluation is in particular difficult when it comes to investment projects that reveal their benefit only in the long term, such as innovation projects involving significant research and development. So it is theoretically feasible that the shareholders are not able to judge the value of R&D while it is being performed. Thus if given too much power to discipline the CEO, the stock market may hamper corporate innovation.

In Chapter 4, I set out to empirically test this theoretical prediction. I show that indeed if given the opportunity to monitor the CEO too frequently, shareholders discipline may make the CEO reluctant to undertake R&D projects that are costly in the short term but valuable in the long term. I establish the causal relationship between evaluation horizon (the inverse of the frequency of "voice" mentioned above) and corporate innovation by a requirement by the SEC in 2011. All US public firms are asked to conduct shareholder voting to approve the CEO compensation proposal (called the "say on pay", or "SOP" for short) either once every year, or once every three years. By comparing the innovation outcomes of companies with different frequencies of SOP, I provide evidence of the relationship between evaluation horizon and innovation. But if done as is, an endogeneity problem may arise since if given the choice of the horizon, companies that specialize in long term innovations may also be the ones that choose infrequent evaluations. To address this problem, I take advantage of a special term and condition imposed by the SEC: the frequency of evaluation itself has to be determined for each company by shareholder votes. This allows me to restrict the comparison to firms that voted *narrowly* in favor of high frequency and those that voted *narrowly* in favor of low frequency. The firms that narrowly passed three years are the considered the treatment group and the firms that narrowly failed to pass a three year horizon are considered the control group. Because of the narrow margin, the aforementioned endogeneity problem is likely to be negligible. In summary, in this chapter I demonstrated that in the context of corporate innovation, stock market discipline may stifle innovation when the market is not able to accurately valuate innovations. This result speaks to the internet frenzy in late 1990s when tech companies are valuated much higher than their true value.

Chapter 2

Trading speed competition: Can the arms race go too far?*

*This project is conducted in collaboration with Dion Bongaerts and Mark Van Achter.

Abstract

In our model, liquidity providers and demanders endogenously adopt costly speed technology. Competition induces negative externalities on same-side traders, and leads to over-investment in speed. However, execution probabilities increase in transaction speed, which generates positive externalities on other-side traders. Contrary to popular belief, liquidity demanders are shown to be more prone to wasteful arms races when marginal gains from trade (GFTs) are constant in transaction speed. Yet, this results reverts with declining marginal GFTs, a setting which we provide micro-foundations for.

2.1 Introduction

In recent years, financial markets have been completely transformed by a newly-emerging group of market participants: high-frequency traders (HFTs), which provide liquidity using computer algorithms at a millisecond pace. As of 2010, HFTs generate at least 50% of volume, and even more so in terms of order traffic in the US equity market.[1] Facing such radical changes, policy makers in the US and the European Union have called for a welfare assessment of HFTs, in order to design appropriate regulation.[2]

The HFT emergence also induced a fierce academic debate. Proponents, like Burton Malkiel, argue that "competition among HFTs serves to tighten bid-ask-spreads, reducing transaction costs for all market participants".[3] In contrast, the opponents, including Paul Krugman, are concerned that HFTs undermine markets and use resources that could have been put to better use.[4] Meanwhile, the vastly growing empirical literature on HFTs has provided evidence consistent with both claims. For example, Brogaard et al. (2014) conclude that HFTs contribute to price discovery, Malinova et al. (2013) show they improve liqudity, and Carrion (2013) finds they provide liquidity when it is scarce. In contrast, several papers document that HFTs reduce liquidity provision significantly in stressful times, in contrast to traditional market makers (see Anand and Venkataraman (2015) and Korajczyk and Murphy (2015)). Moreover, HFT technology is arguably very expensive for society (see Biais et al. (2015b) for a discussion). For instance, in 2010, Spread Networks installed a new $300 million high-speed fiber optic cable connecting New York and Chicago, to reduce the latency of the existing route from 16 to 13 milliseconds. Meanwhile, that improvement has already become virtually obsolete by the introduction of wireless microwave technology in 2011, which managed to almost shave off an additional 5

[1]See the SEC (2010) concept release on equity market structure, and "Casualties mount in high-speed trading arms race", *Financial Times*, Jan 22, 2015.

[2]See "The Morning Risk Report: Future of High Frequency Trading Regulation is Murky", *Wall Street Journal*, January 30, 2014,. and "High-Frequency Traders Get Curbs as EU Reins In Flash Boys", *Bloomberg News*, Apr 14, 2014, respectively.

[3]See "High frequency trading is a natural part of trading evolution", *Financial Times*, Dec 14, 2010.

[4]See "Rewarding bad actors", *New York Times*, Aug 2, 2009.

milliseconds.[5] Moreover, having some of the brightest minds in the world working on the creation, detection or academic analysis of HFT algorithms implies a large opportunity cost for society.

Our model sketches the relevant trade-offs to reconcile these two opposing views. As such, it allows to analyze whether HFTs facilitate allocative efficiency in portfolios sufficiently to justify its (opportunity) costs. In particular, we present two counter-balancing effects HFTs induce on welfare. On the one hand, only the first trader to react to a trading opportunity gains from her investments. As a result, other traders which also invested in trading technology did so in vain (at least, for that trading opportunity), as they arrived (often marginally) later. This negative externality, which we label "substitution effect", materializes both at the liquidity-supply and the liquidity-demand side.[6] Biais et al. (2015a), among others, provide empirical evidence that faster traders indeed obtain larger profits. On the other hand, speedier HFT liquidity provision enlarges opportunities for liquidity demanders to successfully transact, thereby stimulating market participation and investments in trading technology from liquidity demanders. Hence, the interaction between both market sides entails a positive externality, which we label as the "complementarity effect". This effect incorporates and even goes beyond the competition argument put forward by Malkiel. We show that which of the two effects dominates crucially depends on the expected value of each additional trade, or in other words the expected marginal gains from trade (GFT). By and large, the existing literature has (implicitly) assumed the expected marginal GFT to be constant in transaction speed, mainly for tractability reasons. We relax this assumption and show that this relaxation influences which of the two effects dominates. In particular, if expected marginal GFTs increase in transaction speed the complimentarity effect is strengthened and over-investment in technology becomes less likely. If, on the other hand, expected marginal GFTs decrease in transaction speed, the complementarity effect

[5] See Budish et al. (2015) and "Networks Built on Milliseconds", *Wall Street Journal*, May 30, 2012. for a discussion. Other infrastructure-related examples include the cost of co-location services and of individual high-frequency data feeds.

[6] While the early empirical literature mainly focused on the changes in liquidity provision induced by the emergence of high-frequency traders, a similar race is ongoing at the liquidity-demanding side which increasingly applies high-speed algorithmic trading strategies.

is weakened and over-investment in technology becomes more likely. We provide micro-foundations for the latter by working through a portfolio rebalancing problem with risk-aversion and stochastic and discrete trading opportunities. The reason that expected marginal GFTs decline in the average frequency of trading opportunities is that the likelihood of large portfolio dislocations and hence rebalancing needs increases with the time elapsed since the previous trading opportunity at a rate that is slower than linear.

For our analysis we borrow and extend the model by Foucault et al. (2013), who use it to analyze the effect of make-take fees on market quality, but do not use it to analyze arms race effects. The model is a stochastic monitoring model with two types of agents: market makers which fill the book and market takers emptying the book. When a transaction takes place, trading gains are realized as further explained below. Each agent competes with agents of her own type for these gains in a winner-take-all fashion.[7] A speed improvement implies makers and takers have better chances to be the first ones to respectively arrive to an empty or filled book. Yet, lowering latency also implies a cost which is quadratically increasing in monitoring intensity.[8] When optimizing their monitoring speed, both agent types account for the associated (marginal) costs and the obtainable marginal gains from trade (labeled “marginal GFT”). The substitution and complementarity effects can be identified from the optimization problems for both agent types. Both drive resource allocations away from first best in opposite directions. The expected marginal GFT essentially function as a weight on the complementarity effect. With more takers than makers, arms race effects are more likely to be seen among takers than among makers as the group on which negative externalities are exerted is relatively large and the group on which positive externalities are exerted is relatively small. This preliminary conclusion would go against the popular perception that arms races would be more

[7] Thus, the fastest trader is the only one that profits from a standing trading opportunity. In particular, the first maker to arrive to an empty book can post a sell limit order. Subsequent makers arriving to the book need to wait till the book is empty again. The first taker to arrive to a filled book can transact at the standing sell order. Subsequent takers arriving to the book need to wait till the book is filled again.

[8] This reflects the increasingly costly investments in human capital and IT-infrastructure (which become increasingly costly the more we approach the speed of light).

prevalent among makers.

Assuming the expected marginal GFTs to be constant leads to a tractable model with closed-form solutions. This is probably one of the reasons why many papers have made this assumption (albeit often implicitly; see for example, Biais et al. (2015b), Foucault et al. (2013) and Hoffmann (2014))). We generalize the model by assuming the expected marginal GFTs to be a strictly monotonic and differentiable function of average transaction speed. This compromises tractability as solving the model now involves finding the roots of a quintic function, which is analytically generally not possible. Yet, this generalization is important as it (i) materially affects the weight on the complementarity effect, and (ii) is much more plausible, as later on shown by our micro-foundations. We manage to show the incremental effect by analyzing the first-order conditions for makers and takers and highlighting additional terms that are monotonic in the dependence of GFTs on transaction frequency. More specifically, we show that with decreasing expected marginal GFTs, the additional effects we identify increase the likelihood of over-investment (and hence arm races) among makers, and reduce the likelihood of over-investment among takers.[9] These findings could bring our results more in line with the popular belief that arms races are more prevalent among makers. We illustrate these findings with numerical examples.

The last step of our main analysis provides micro-foundations that expected marginal GFTs are declining in transaction frequency. Or in other words, that on average, each additional trade that takes place as a result from upgrades in trading speed technology is less valuable than the previous one. To this end, a portfolio rebalancing problem for an investor with log utility and access to a risky and risk-free asset is set up. Due to the log utility, the investor would like the weights on both assets in his portfolio constant. This creates a continuous rebalancing motive as prices of the risky asset continuously move. However, trading is only possible at stochastically determined discrete points in time. For every trade the gain from trade is given by the increase in investor utility due to the trade. This utility improvement is increasing in the price move of the risky asset since the previous trade. Since the

[9] Naturally, the effects work the opposite way when expected marginal GFTs increase with transaction frequency.

variance of returns increases in the square root of time, each additional trading opportunity is valuable, but its expected value is decreasing in the frequency at which such trading opportunities arise. A similar case can be made for a risk-averse investor dynamically hedging a non-linear derivative (portfolio).

From a modeling perspective, the closest paper to ours is by Foucault et al. (2013), as we largely draw on their model. The economic effects of the two opposing forces on over- or under-investment are also in their model, but are only used in the discussion on make-take fees. We link their model to the aggregate welfare question of HFTs, and most crucially show the importance of properly modeling the marginal GFTs as these will serve as weights on the two countervailing forces. Our paper provides a further contribution in outlining the micro-foundations for declining marginal GFT with risk-averse takers.

Our paper also contributes to the rapidly expanding theoretical literature on the effect HFTs have on welfare. Many recent papers focus on the asymmetric information channel (i.e., the pick-off risk slow traders face) when evaluating the welfare consequences of speed technology (e.g., Biais et al. (2015b), Budish et al. (2015), Cespa and Vives (2015), Hoffmann (2014), Jovanovic and Menkveld (2015), Menkveld and Yueshen (2012), and Rojcek and Ziegler (2016). Other papers, such as Aït-Sahalia and Saglam (2014), explore the welfare impact of the inventory channel (i.e., HFTs are more efficient in optimizing their inventories over time). We document that physical and human capital (opportunity) costs alone suffice to induce a wasteful arms race, and we obtain welfare losses even in the absence of the aforementioned channels. Furthermore, we show that the commonly-made assumptions of constant marginal GFT and risk-neutral investors may induce an underweighting (in the model) of the negative substitution externality HFTs exert. Adverse selection is then needed to generate arms race effects on the maker side.

Taking a broader view, our model shows similarities with traditional imperfect competition models such as Cournot (1838). The intensity in our model is (to a large extent) equivalent with quantity in such models. In these models, producers typically face a downward sloping demand curve. The declining marginal gains from trade we provide micro-foundations for are consistent with such a downward sloping demand curve. Yet, there are also key differences with these traditional models. First, our model features competition on both sides of a trade, because makers and takers both compete for profitable trading opportunities. Second, the way individual monitoring intensities translate to transaction intensities generates interesting patterns. The stochastic winner-takes-all feature of trading induces more over-investment. In addition, the interaction of how buy and sell side monitoring intensities translate to transaction intensities generates the complementarity effect.

2.2 Setup

The setup of our model is based Foucault et al. (2013). We consider a market with two types of participants: market makers and market takers. Each maker $i \in \{1, 2, \ldots, M\}$ monitors the market at discrete points in time and arrives according to a Poisson process with (endogenously chosen) intensity parameter $\mu_i \geq 0$. Similarly, each taker $j \in \{1, 2, \ldots, N\}$ arrives with (endogenously chosen) intensity $\tau_j \geq 0$. The total numbers of makers and takers, M and N, are exogenous.[10] By assumption, there is no form of information asymmetry present.

We assume that there is a market mechanism to bring together liquidity demand and supply. Crucially, at this stage we write the model in its most general form and the market mechanism can take many forms. The way in which maker and taker intensities map into expected trading intensities depends on the market mechanism. More generally, we define r as the transaction intensity and further define

$$E(r) = f(\sum_i \mu_i, \sum_j \tau_j), \tag{2.1}$$

where the underline indicates a vector of intensities. We do require the expected trading intensity to be a strictly increasing function of both the maker and the taker

[10] This setup has a winner-takes-all feature from an ex-post perspective (i.e., the one conducting the trade is the only one benefitting). From an ex-ante perspective, the fastest trader is not guaranteed to always execute. This setup is chosen based on the notion of order processing uncertainty: the fast trader never knows what is going to happen after she submits the order and before it reaches the server of the exchange. A similar argument can be found in Yueshen (2014).

intensities. In other words, we require that

$$\frac{\partial f(\underline{\mu}, \underline{\tau})}{\partial \mu_i} > 0, \qquad \frac{\partial f(\underline{\mu}, \underline{\tau})}{\partial \tau_j} > 0. \tag{2.2}$$

The market operates according to the following timeline: before trading begins, each maker i chooses an intensity μ_i to maximize her expected trading profit $\Pi_m(\mu_i)$.[11] Similarly, each taker j chooses τ_j to maximize her expected trading profit $\Pi_t(\tau_j)$. Once the trading begins, each player monitors the market following a Poisson process with the intensity previously chosen. Every time the trader arrives to the market he can submit an order to trade one unit of the security. For the moment, we assume that the trading phase of the model repeats itself an infinite number of times. One should notice that while trading happens indefinitely, the model is in essence a one-shot model as arrival intensities are chosen once at the start of the game and kept constant throughout.

In our model, monitoring speed does not come for free. In particular, we assume the per period monitoring costs (in clock time, not transaction time!) for both trader types to be quadratically increasing in the monitoring frequency chosen. These costs reflect the required investments in IT-infrastructure and human capital. The marginal cost of technology is increasing, reflecting the observation that as latency approaches zero, the cost for such advancement becomes higher and higher.[12] Monitoring costs can differ between the makers and the takers. This difference allows to assess the impact of heterogeneity in know-how (i.e., a knowledge endowment) among the market participants. We denote the cost per unit of time for maker i by $C_m(\mu_i) = \beta\mu_i^2/2$, while for taker j it equals $C_t(\tau_j) = \gamma\tau_j^2/2$. One should note that since trading continues indefinitely, we are interested in the per period costs compared to the per period revenues.

Finally, when a transaction takes place, the gains from trade (GFT) are split between between maker and taker according to the fractions π_m and $\pi_t = 1 - \pi_m$,

[11] In this sense the agents in our models are the prop traders in Biais et al. (2015a)

[12] For example, to improve latency from a second to a tenth of a second, one would "only" need to automate the trading using algorithms. To get to the millisecond level, however, co-location and super-fast communication lines are required, which are significantly more costly.

respectively, where $\pi_m \in (0,1)$.[13] The *expected GFT for an additional trade* or the expected marginal GFT originated by liquidity taker j are given by a (weakly) monotonically increasing and continuously differentiable function $G\left(\frac{\tau_j}{\bar{\tau}}r\right)$ of his expected trading frequency. We assume that the GFT arise from the portfolio selection and consumption need of the *takers*. This is supported by the reality that in financial markets, the market takers are usually the parties with the intention to hold the security over an horizon exceeding a day, while the makers mainly serve as short term intermediaries. In Section A.1, we will provide further micro-foundations for this assumption. Relatedly, the *expected GFT per unit of time* is given by $W\left(\frac{\tau_j}{\bar{\tau}}r\right) = \left(\frac{\tau_j}{\bar{\tau}}r\right)G\left(\frac{\tau_j}{\bar{\tau}}r\right)$. $W(\cdot)$ is our measure of social welfare. One can prove that concavity of $W(\cdot)$ is equivalent to the $G(\cdot)$ function being uniformly decreasing in expected trading speed on its domain.

2.3 Equilibrium Analysis and Welfare

In this section, we first set out to define the first best solution and equilibrium outcome of the model. Moreover, we define what we mean with over- and under-investment in trading technology. Next, we solve the model for the case of constant expected marginal GFT, i.e., $G(\cdot) = G_0$. Thereafter, we continue by solving the model in its more general form, namely for any monotonic and differentiable function $G(\cdot)$. Because tractability in the general case is low, the argument can only be made by analyzing the difference between first-order conditions a social planner faces and those that market participants face. Having established a general result that depends on the expected marginal GFT function, we provide micro-foundations for the shape of this function. We finish the section by illustrating our model outcomes with graphical representations of numerical examples.

2.3.1 First best and equilibrium definitions

In this subsection, we define the first-best outcome, the equilibrium outcome, and under- and over-monitoring. Let us start by deriving the first best outcome as the solution of a social planner's problem, where the social planner only cares about

[13] Participation incentives dictate that $\pi_m \in (0,1)$ must hold on average. Because we abstract from informed trading, this must even hold for each individual trade.

aggregate welfare.

Definition 1 ***Social Planner's Problem***

A social planner chooses $\{\mu_i\}_{i=1,2,...,M}$ *and* $\{\tau_j\}_{j=1,2,...,N}$ *to maximize aggregate social welfare:*

$$\sum_{j=1}^{N} W\left(\frac{\tau_j}{\bar{\tau}}r\right) - \sum_{i=1}^{M} \frac{\beta\mu_i^2}{2} - \sum_{j=1}^{N} \frac{\gamma\tau_j^2}{2}, \tag{2.3}$$

such that $\mu_i \geq 0, \forall i \leq M; \tau_j \geq 0, \forall j \leq M$.

As can be gauged from this objective function, the social planner only focuses on the aggregate gains from trade which are realized by the interaction between makers and takers. It does not account for the distribution of these gains between makers and takers (i.e., π_m does not show up in this equation).

Next, we define an equilibrium as the outcome of the setting in which makers and takers individually optimize their intensities given the optimal strategies of all other players (i.e., we look for a Nash equilibrium).

Definition 2 ***Equilibrium***

The equilibrium that we consider is a Nash equilibrium defined by intensity choices $\{\mu_i\}_{i=1,2,...,M}$ *and* $\{\tau_j\}_{j=1,2,...,N}$, *such that:*

1. *Given all taker intensities,* $\{\tau_j\}_{j=1,2,...,N}$, *as well as all other maker intensities:* $\{\mu_n\}_{n=1,2,...,i-1,i+1,...,M}$ *each maker* i *maximizes her profit after cost:*

$$\pi_m \frac{\mu_i}{\bar{\mu}} \sum_{j=1}^{N} W\left(\frac{\tau_j}{\bar{\tau}r}\right) - \frac{\beta\mu_i^2}{2}$$

2. *Given all maker intensities,* $\{\mu_i\}_{i=1,2,...,M}$, *as well as all other taker's choice:* $\{\tau_n\}_{n=1,2,...,j-1,j+1,...,N}$ *each taker* j *maximizes her profit after cost:*

$$(1-\pi_m)W\left(\frac{\tau_j}{\bar{\tau}r}\right) - \frac{\gamma\tau_j^2}{2}$$

where $\mu_i \geq 0$, *and* $\tau_j \geq 0$.

We define over- and under-investment by makers and takers as the equilibrium $\hat{\tau}$ and $\hat{\mu}$ respectively exceeding and falling short of their first-best counterparts.

We will now compare equilibrium outcomes to first best outcomes, and will explicitly consider the functional form of $G(\cdot)$. Doing so will prove to be crucial to determine whether wasteful arms races occur. The current literature mostly assumes that $G(\cdot)$ is constant, irrespective of the trading speed (for example, Biais et al. (2015b), Foucault et al. (2013) and Hoffmann (2014)). We therefore first solve the model for constant $G(\cdot)$ as a benchmark.

2.3.2 Benchmark Case: constant expected marginal GFT

In this case, each transaction that takes place generates the same amount of social welfare, and we have that $G(\cdot) = G_0$. Hence, the aggregate GFT is linear in the expected trading frequency of the taker involved.

In first best, the FOCs of the social planner's optimization problem should hold, whereas in equilibrium those of the individual market participants should hold. The first order conditions are given by the following lemma.

Lemma 3 ***FOCs of constant GFT setting***

The first order conditions of the SPP and the individual maker optimization problems are given by

$$G_0 \frac{\partial r}{\partial \mu_i} = \beta \mu_i, \qquad \textit{(SPP FOC } \mu\textit{)} \tag{2.4}$$

$$G_0 \pi_m \left(\frac{\mu_i}{\bar{\mu}} \frac{\partial r}{\partial \mu_i} + \frac{\mu_i}{\bar{\mu}^2} r \right) = \beta \mu_i. \qquad \textit{(maker FOC)} \tag{2.5}$$

The equivalent expressions for takers are fully symmetric.

Proof. See Appendix. □

Using the concavity properties of r, one can show that if π_m were large, over monitoring would occur, whereas if π_m is sufficiently small under-monitoring occurs. The intuition for over-monitoring is that individual makers do not endogenize the negative effect their speed increase has on the transaction likelihood of other makers (substitution effect). The intuition for under-monitoring is that due to the fact that only a part of the surplus can be captured, there is insufficient incentive to invest in monitoring capacity (complementarity effect).

To make our setting more concrete we now assume a trading mechanism in the form of a limit order book that has capacity for only one limit order at a time. In order for a transaction to take place, the limit order book needs to be filled and a market order needs to arrive. By the properties of the Poisson distribution, the aggregate monitoring process of all makers jointly also follows a Poisson distribution with the following intensity: $\bar{\mu} = \sum_{i=1}^{M} \mu_i$. Similarly, the aggregate monitoring intensity of all takers jointly equals $\bar{\tau} = \sum_{j=1}^{N} \tau_j$. Consequently, the expected interval between a transaction and replenishment of the book equals $D_m = 1/\bar{\mu}$. Analogously, the expected interval between the posting of a limit order and transaction is given by $D_t = 1/\bar{\tau}$. Thus on average, the duration between two trades is $D = D_m + D_t$, and the average trading frequency equals $r = (D_m + D_t)^{-1} = \frac{\bar{\mu}\bar{\tau}}{\bar{\mu}+\bar{\tau}}$. Similarly, one can show that for a given liquidity taker j with monitoring intensity τ_j, the expected trading frequency is given by $\frac{\tau_j \bar{\mu}}{\bar{\mu}+\bar{\tau}} = \frac{\tau_j}{\bar{\tau}} r$.

Proposition 1 ***First Best with Constant Marginal GFT***

The first best monitoring intensities are symmetric and given by:

$$\hat{\mu} = \frac{N^2}{(M\hat{r}+N)^2}\frac{G_0}{\beta}, \hat{\tau} = \frac{M^2\hat{r}^2}{(M\hat{r}+N)^2}\frac{G_0}{\gamma}; \qquad \textit{where } \hat{r} = \left(\frac{N^2}{M^2}\frac{\gamma}{\beta}\right)^{\frac{1}{3}}.$$

The resulting aggregate welfare equals:

$$\hat{\Pi} = \frac{M\hat{\mu}N\hat{\tau}}{M\hat{\mu}+N\hat{\tau}}G_0 - \frac{\beta M\hat{\mu}^2}{2} - \frac{\gamma N\hat{\tau}^2}{2}.$$

Proof. See Appendix. □

The economic intuition behind this solution is as follows. First, the optimal maker intensity $\hat{\mu}$ increases in the expected marginal GFT, G_0, because higher G_0 justifies higher investments in monitoring technology to be made. Second, $\hat{\mu}$ decreases in the marginal monitoring cost for makers, β, due to the increasing marginal cost of monitoring intensity. Third, $\hat{\mu}$ decreases in the number of makers, M, since it is the aggregate intensity that the social planner cares about and individual costs are quadratic in monitoring intensity. The more makers there are, the lower the required frequently for each individual maker; this we call the "substitution effect." In addition to the within-type effects outlined above, cross-type effects can also be observed. First, the optimal maker intensity $\hat{\mu}$ decreases in the marginal cost of the takers, γ. This happens because maker intensity and taker intensity are complementary. After all, high monitoring by takers is only useful if the book is likely to be F and high monitoring intensity by makers is only worthwhile if the book is likely to be E. Hence, makers and takers impose positive externalities on each other; this we call the "complementarity effect." Second, $\hat{\mu}$ increases in the numbers of takers N due to the same type of complementarity. Third, this complementarity effect can also be seen from the first term of the formula for the aggregate welfare (i.e., $\frac{M\hat{\mu}N\hat{\tau}}{M\hat{\mu}+N\hat{\tau}}G_0$). Unilateral increases in maker intensity $\hat{\mu}$ will increase the total welfare not only by a factor of M, but also by N, the number of takers. All six interpretations above apply to the taker intensities $\hat{\tau}$ too (for the constant expected marginal GFT, the problem is symmetric).

In reality, however, the first best outcome typically does not materialize. Therefore, we now proceed by solving for the equilibrium of this economy and compare it with the first best outcome outlined above. While solving in closed form is not possible, we follow Foucault et al. (2013) and obtain an implicit solution for the equilibrium with constant expected marginal GFT.

Proposition 2 ***Equilibrium with Constant Marginal GFT***

The equilibrium monitoring intensity for makers and takers is given by:

$$\mu^* = \frac{M + (M-1)\tau^*}{M(1+r^*)^2}\frac{G_0\pi_m}{\beta}, \tag{2.6}$$

$$\tau^* = \frac{r^*((1+r*)N-1)}{N(1+r^*)^2}\frac{G_0\pi_t}{\gamma}, \tag{2.7}$$

respectively, where r^* *is the positive real solution to:*

$$Nr^3 + (N-1)r^2 - (M-1)zr - Mz = 0$$

with $z \equiv \frac{\gamma\pi_m}{\beta\pi_t}$.

The resulting aggregate welfare is given by:

$$\Pi^* = \frac{M\mu^* N\tau^*}{(M\mu^* + N\tau^*)} G_0 - \frac{\beta M {\mu^*}^2}{2} - \frac{\gamma N {\tau^*}^2}{2}. \tag{2.8}$$

Proof. See Appendix. □

An easy comparison between the equilibrium and first best monitoring intensities is not possible because the equilibrium expressions involve r^* which is implicitly defined. However, we are able to obtain intuition of the forces at work by comparing the first order conditions behind the first-best and the equilibrium outcomes.

Substituting for r and $\frac{\partial r}{\partial \mu_i}$ in (2.4) and (2.5), we get the following first order conditions of the SPP and the individual maker optimization problems

$$G_0 \frac{\bar{\tau}^2}{(\bar{\mu} + \bar{\tau})^2} = \beta\mu_i, \qquad \textit{(SPP FOC } \mu\textit{)} \tag{2.9}$$

$$G_0\pi_m \frac{\bar{\tau}^2 + \mu_{-i}\bar{\tau}}{(\bar{\mu} + \bar{\tau})^2} = \beta\mu_i, \qquad \textit{(maker FOC)} \tag{2.10}$$

where $\mu_{-i} \equiv \sum_{j \neq i}^{j \leq M} \mu_j$. The equivalent expressions for takers are fully symmetric.

In these FOCs, we can see that the complementarity effect and the substitution effect give rise to positive and negative externalities. First, observe that the LHS of the *SPP-FOC* has a multiplier G_0, while in the *maker-FOC* this multiplier is given by $G_0\pi_m$. Since $\pi_m \in (0,1)$, this multiplier is smaller, and hence an individual taker does not fully endogenize the complementarity effect her activity induces on her counter-parties (this makes it a positive externality). As a result, makers are in equilibrium inclined to under-monitor as compared to the first best. Secondly, the numerators of the LHSs of the two FOCs differ by a term $\mu_{-i}\bar{\tau}$, which increases the equilibrium intensities relative to the first-best intensities. The intuition behind this is as follows. Increasing one's own intensity increases the transaction probability for the individual maker or taker, but also reduces the effectiveness of orders sent by all competitors as limit orders are more likely to hit a full book and market orders more likely to hit an empty book. Because this effect on competitors is not endogenized by individual makers or takers, this substitution effect gives rise to a negative externality on same-side market participants.

For takers the problem is completely symmetric here and all intuition carries over. The two effects described above are (partially) offsetting. Which of the two dominates depends on parameters. Also here the FOCs provide guidance. A larger π_m reduces the complementarity effect for makers and strengthens it for takers (as $\pi_t = 1 - \pi_m$). If M is small, all liquidity needs to be provided by a small number of makers. Because costs are quadratic in monitoring intensity, and because the number of competitors is small in this case, we have that $\frac{\mu_{-i}\bar{\tau}}{(\bar{\mu}+\bar{\tau})^2}$ is small as well, limiting the substitution effect and making an arms race less likely.

2.3.3 Relaxing the constant expected marginal GFT assumption

In this subsection we conduct a similar analysis as in the previous section, but then with a more general functional form for the expected marginal GFT. In particular, if $G(\cdot)$ is a monotonic and differentiable function of transaction frequency $\frac{\tau_j\bar{\mu}}{\bar{\mu}+\bar{\tau}}$, then we can obtain the respective FOCs for the SPP and individual market participants by subsequently applying the product rule and chain rule for differentiation. Imposing symmetry among players of the same type, we get the following expressions for the FOCs w.r.t. maker and taker intensities:

Lemma 4 *If $G(\cdot)$ is a monotonic and differentiable function of transaction frequency $\frac{\tau_j\bar{\mu}}{\bar{\mu}+\bar{\tau}}$, the first order conditions of the SPP and the individual maker and taker opti-*

mization problems are given by:

$$\beta\mu_i = G\left(\frac{\bar{\mu}\bar{\tau}/N}{\bar{\mu}+\bar{\tau}}\right)\frac{\bar{\tau}^2}{(\bar{\mu}+\bar{\tau})^2} + G'\left(\frac{\bar{\mu}\bar{\tau}/N}{\bar{\mu}+\bar{\tau}}\right)\frac{\bar{\mu}\bar{\tau}^3/N}{(\bar{\mu}+\bar{\tau})^3}, \qquad \textit{(SPP FOC } \mu\textit{)} \quad (2.11)$$

$$\beta\mu_i = \left(G\left(\frac{\bar{\mu}\bar{\tau}/N}{\bar{\mu}+\bar{\tau}}\right)\frac{\bar{\tau}(\mu_{-i}+\bar{\tau})}{(\bar{\mu}+\bar{\tau})^2} + G'\left(\frac{\bar{\mu}\bar{\tau}/N}{\bar{\mu}+\bar{\tau}}\right)\frac{\bar{\mu}\bar{\tau}^3/N}{(\bar{\mu}+\bar{\tau})^3}\frac{\mu_i}{\bar{\mu}}\right)\pi_m, \qquad \textit{(maker FOC)} \quad (2.12)$$

$$\gamma\tau_j = G\left(\frac{\bar{\mu}\tau_j}{\bar{\mu}+\bar{\tau}}\right)\frac{\bar{\mu}^2}{(\bar{\mu}+\bar{\tau})^2} + G'\left(\frac{\bar{\mu}\tau_j}{\bar{\mu}+\bar{\tau}}\right)\frac{\bar{\mu}^3\tau_j}{(\bar{\mu}+\bar{\tau})^3}, \qquad \textit{(SPP FOC } \tau\textit{)} \quad (2.13)$$

$$\gamma\tau_j = \left(G\left(\frac{\bar{\mu}\tau_j}{\bar{\mu}+\bar{\tau}}\right)\frac{\bar{\mu}(\bar{\mu}+\tau_{-j})}{(\bar{\mu}+\bar{\tau})^2} + G'\left(\frac{\bar{\mu}\tau_j}{\bar{\mu}+\bar{\tau}}\right)\frac{\bar{\mu}^2\tau_j(\bar{\mu}+\tau_{-j})}{(\bar{\tau}+\bar{\mu})^3}\right)(1-\pi_m). \quad \textit{(taker FOC)} \quad (2.14)$$

What is clear from both sets of equations is that compared to the constant expected marginal GFT, there is now a term involving $G'(\cdot)$ that enters with a strictly positive coefficient since intensities are always (strictly) positive. The effect of these terms differs (i) between makers and takers and (ii) between welfare maximization and individual profit maximization of market participants. In particular, this term in the FOC on the μs is less prominent for individual makers than for a social planner. Hence, if $G' > 0$, makers tend to under-monitor more and if $G' < 0$ they tend to over-monitor more. The reason is that if $G' < 0$, an increase in speed lowers at the margin the welfare that is created by each additional trade. However, since each maker can only expect to be present in a fraction $\frac{\mu_i}{\bar{\mu}} < 1$ of all trades, this is not fully endogenized and enters as a negative externality. For takers, on the other hand, the reverse result holds. if $G' > 0$, takers tend to over-monitor more and if $G' < 0$ they tend to under-monitor more. If $G' < 0$, an increase in speed reduces the marginal value of each trade for a specific taker, just as in the social planner's problem, because numerator of the argument of $G(\cdot)$ only involves the transaction speed of an individual taker. However, for the social planner, a higher taker intensity lowers the rate at which marginal GFT deteriorate for other takers because of the substitution effect. An individual taker does not endogenize this and hence, the term involving $G'(\cdot)$ is relatively more important. Therefore, individual takers are more prone to under- rather than over-invest in monitoring speed when expected marginal gains from trade are declining in transaction speed. We can also analyze the moderating effect of relative bargaining power by looking at these FOCs. The term involving G' in (2.12) is relatively more important when π_m is large. One should note that this is already the situation in which even with constant expected marginal gains from trade over-investment by makers is more likely. In this case, under-investment on the taker side is likely, but the additional term involving G' gets little weight, such that under-investment problems are hardly amplified. By contrast, if π_m is low, under-investment on the maker side is likely, and the term involving G' attenuates the under-investment problem, but only a little. On the taker side, a low π_m leads to likely over-investment on the taker side, which is heavily attenuated by the term involving G'. Our findings are summarized in the following proposition.

Proposition 3 *Compared to the constant marginal gains from trade case, if expected marginal gains from trade are monotonically decreasing in expected transaction speed, then in equilibrium*

1. *Makers are more likely to over-invest and less likely to under-invest*
2. *Makers over-invest more or under-invest less*
3. *Takers are more likely to under-invest and less likely to over-invest*
4. *Takers under-invest more or over-invest less*
5. *These effects positively interact with π_m for makers and negatively for takers, such that arms races are particularly likely and severe among makers and primarily when makers have relatively much bargaining power.*

The reverse holds if expected marginal gains from trade are uniformly increasing in expected transaction speed.

Proof. See Appendix. □

These findings have important ramifications. We typically observe in markets the number of takers to be much larger than the number of makers. As a result, with constant expected marginal GFT, investments in monitoring speed by takers lead to a relatively larger substitution effect. After all, the group of peers that are negatively affected by investments in speed technology of a given taker is larger. For makers, the substitution effect should be relatively larger, leading to under-investment as the group on which they extend positive externalities is larger. For takers, the complementarity effect should be relatively larger, leading to over-investment as the group on which they extend negative externalities is larger. If expected marginal GFT are declining in transaction speed, as we will argue in the next section, the additional terms in Equations (2.11) to (2.14) counter these effects, both for takers as well as for makers, and in particular when bargaining power for makers is high. Hence, we are much more likely to see over-investment by makers, a concern often expressed by regulators and policy makers.

2.3.4 Are expected marginal GFT increasing or decreasing with speed?

In the previous section we saw the crucial importance of the shape of the expected marginal GFT function $G(\cdot)$ when analyzing welfare effects of HFTs. One would naturally like to know which shape of $G(\cdot)$ would be most plausible. We argue that $G(\cdot)$ is most likely to be downward sloping in transaction speed. We present two very much related settings, micro-foundations, in which such shape materializes naturally.

In the first setting, we consider an economy with two assets; one risky and one risk-free. The value of the risky asset follows a geometric Brownian motion. There is an investor with initial wealth and log utility with risk-aversion coefficient δ. This investor needs to continuously optimize a consumption and portfolio allocation problem. Due to his log utility, the investor would like to maintain fixed portfolio weights on the risky and risk-free assets. Because the price of the risky asset fluctuates continuously, this investor has a continuous rebalancing need. However, rebalancing is possible only at stochastic but discrete points in time. Whenever a rebalancing trade takes place a fraction π_m of the welfare gain resulting from the trade accrues to the liquidity provider of the trade and hence is a welfare loss to the investor. In the end, we are interested in how average aggregate welfare created by trading depends on the arrival frequency of trading opportunities. This tells us how $G(\cdot)$ depends on $\frac{\tau_j \bar{\mu}}{\bar{\mu}+\bar{\tau}}$. We solve this portfolio rebalancing problem in Appendix A.1. A portfolio rebalancing problem with transaction costs is traditionally very hard to solve, because for fixed transaction costs there are thresholds for trading to take place. As a result, trading will not happen at each opportunity. Because we impose transaction costs that are proportional to welfare created by trading, trading takes place when the opportunity arises with probability 1. As a result, we can solve an equivalent problem that does not feature transaction costs. For this equivalent problem, we then show that the expected marginal gains from trade are decreasing in the arrival frequency of trading opportunities as the need to trade closely after the previous trade is not very high. It is simply very unlikely that since the previous trade the price of the risky asset has moved a lot and hence, the portfolio is still close to its optimum.

A very similar result materializes for a risk-averse hedger or arbitrage trader that wants to dynamically hedge a position in an (exotic) non-linear derivative by trading in the spot and money market. Since non-linear derivatives typically have a Gamma exposure, there is a need to constantly rebalance the position. However, trading is only possible at stochastic but discrete points in time. Because of risk-aversion, each trade creates welfare. As before, transaction costs proportional to welfare created are irrelevant for trading decisions and therefore welfare patterns (they only lead to a level shift). For similar reasons as above, the marginal value of an additional trade declines with the arrival frequency of trading opportunities.

2.3.5 Numerical illustration

To illustrate the effect of the non-constant expected marginal GFT, we present numerical examples to illustrate our point.

In this exercise, we visualize the maker and taker intensities relative to first best. To enable 3D-plotting, we reduce the number of free parameters by restricting our model parameters as follows: we set $G_0 = 1$ and $\beta = \gamma$. Then, we make plots of the under- and over-monitoring of either type as a function of the profit split ratio π_m and the cost coefficient β, for several combinations of (M, N). In particular, we plot the following combinations of M and N: $(2, 2); (2, 20); (50, 250)$. We set the number of takers larger or equal to the number of makers as, in reality, there are usually more liquidity demanders than suppliers in any particular market.

Figure A.1 presents the graphs for the constant expected marginal GFT. Let us highlight some features of the plots consistent with our intuition. First, as argued above, the smaller the cost coefficient (β, γ), the more monitoring diverges from first best (i.e., over- and under-monitoring are generated). Second, over-monitoring is more likely to happen when there are more makers and/or takers competing on the same side of the market. In this case there are more competitors, and hence more sources of negative substitution externalities.

Our main observation from this figure is that for relatively small values of M and N, neither side over-monitors severely. The first column in Figure A.1, which corresponds to the case when $M = 2$ and $N = 2$, shows no over-monitoring at all. The arms race only begins gradually from $N = 20$. Even with $(M, N) = (2, 20)$ over-monitoring is still very limited as shown in the second column of Figure A.1. When we look at the third column of the figure, we notice that as the market size grows, the substitution effect starts to dominate the complementarity effect more and more. Yet, especially for the makers, we see large parameter ranges where there is under- rather than over-monitoring. Only for relatively high values of π_m, do we see over-monitoring among makers. This also intuitively makes sense as a higher value for π_m gives more value to being the first one to execute a trade and lowers the benefit of the takers to invest in trading technology in response to an upgrade in maker technology and speed.

These results tell us that when the expected marginal GFT are constant, the complementary effect can easily start to dominate the substitution effect. Moreover, if an arms race is going on, it is on the taker rather than on the maker side, due to the relatively higher presence of the takers in the market. This is in contrast with the mainstream view on the arms race that it is more likely to occur among makers. In addition, it stands out from the figure that the over-monitoring of the takers only occurs when π_m is close to 0.5, that is, the two sides have similar bargaining power. An intuitive explanation is that when π_m is too small, there are not enough makers to generate enough positive externality to motivate sufficient monitoring of the takers, let alone the over-monitoring. On the other hand, when π_m is large, takers benefit little from trading and will invest too little. This observation suggests that the relative market power of the liquidity suppliers and demanders, in addition to their speed and their sheer numbers, can also be relevant factors to take into account when designing regulatory measures to ensure efficient monitoring.

We next present the same graphs in Figures A.2, A.3, and A.4, but for a setting in which the expected marginal GFT is linearly declining in transaction frequency with slope coefficient G_1. In particular, we set G_1 to 1, 0.2 and 5, respectively. Moreover, in order to maximize comparability, we keep the same value for G_0 as we used in the case with constant expected marginal GFT. When $G_1 = 1$ and holding constant the number of makers and takers, we observe that the tendency for makers to over-monitor is higher compared to Figure A.1. Moreover, even for small values of M, makers only over-monitor in this setup, and hardly ever in the constant expected marginal GFT setup. Figures A.3 and A.4 show that as the slope steepens, the over-monitoring becomes more severe, demonstrating that the nature of the marginal gains from trade plays an important role in determining the occurrence of the arms race.

2.4 Robustness and extensions

2.4.1 Other externalities

Competition in speed may have positive externalities in the sense that it boosts technological progress and knowledge. Other industries may benefit from this progress in unanticipated ways. As an example, the internet was developed for internal and largely military purposes, but in an unanticipated way massively improved productivity and living standards across the globe. The model is able to capture such features easily in reduced form. To this end, we can define $\tilde{\beta} = (1 - \zeta)\beta$ (and analogously $\tilde{\gamma}$) as the social costs of speed technology development and adoption after controlling for cross-product externalities captured by the term ζ. Similarly, one could incorporate negative externalities too by setting $\tilde{\beta} = (1 - \zeta + \xi)\beta$, where ξ is the negative externality, for example resulting from increased information asymmetry. The effect of such generalizations is rather straightforward. As these are externalities, social cost parameters increase or decrease while cost parameters for individual optimization problems remain unaffected. Hence, larger ζ leads to under-investment in technology, while larger ξ leads to over-investment.

2.5 Conclusion

In this paper, we have explored whether competition on speed among stock market participants is likely to trigger arms races, leading to socially wasteful investments. We highlight two opposing economic channels that influence such effect in opposing and partially offsetting ways. Competition among makers as a group and among takers as a group may indeed trigger arms races in the classical sense. However, a complementarity between the two sides, the increased success rate of trading, may offset this competition effect if the gains from trade are large enough. Therefore, the likelihood of arms races depends on how gains from trade depend on transaction frequency. The expected marginal GFT essentially acts as a weight on the complementarity effect. We show that if the expected marginal GFT is declining in transaction speed, the weight on the complementarity effect declines and arms races are more likely to occur. This we also illustrate graphically and numerically. Using a portfolio rebalancing model, we show that expected marginal GFT are indeed likely to be declining in transaction frequency. Intuitively, the gains realized in a trade shrink the smaller the time interval in between subsequent trades. Under this new more realistic specification, arms races are more likely to occur than under the standard paradigm in the literature (featuring constant expected marginal GFT). We provide several extensions to the model. For example we show that the model can incorporate in reduced form other externalities, like on unrelated technological progress.

While providing important insights, our model does make some concessions to reality. A potential concern is that it does not allow for the dual role of participants in modern limit order markets. Yet, in fact this concern is not as grave as one would think. After all, there is a group of market participants that are likely to show a net demand for liquidity. Moreover there is a group that on net will be providing liquidity. This is what in the end generates the welfare gains. Trades among makers, which currently are very common, are zero sum within the group of makers (one maker could have been the only intermediary rather than a whole chain). The single role assumption massively simplifies our problem, leading to better tractability.

Chapter 3

Asset risk and bank runs*

*This project is conducted in collaboration with Dion Bongaerts.

Abstract

We introduce a bank-run model similar to Goldstein and Pauzner (2005) with 1.) balance sheet equity and 2.) a menu of risk choices available to the bank. Elevated risk increases the chances of insolvency and hence may trigger bank runs. However, elevated risk is also associated with higher returns in good state of the world. These create additional capital buffers for debt holders and hence lower the probability of insolvency and therefore also bank runs. Since expected returns and bank capital are essentially substitutes, the latter effect dominates when the marginal benefit of additional capital is high (i.e., capital is low).

3.1 Introduction

Banks are unique financial institutions due to the composition of their balance sheets. Having an asset side primarily comprised of illiquid assets and a liability side primarily comprise of liquid overnight deposits, a banks fulfills a useful liquidity creation function in the economy. Yet, at the same time this special structure also leads to inherent fragility in the sense that banks are highly exposed to liquidity risk (and much more so than corporates). In particular, banks are exposed to the risk of bank runs. A bank run occurs when so many depositors withdraw their deposits that the long term assets of the bank have to be liquidated in a fire sale to satisfy their withdrawal. In such a run, many depositors without a liquidity need withdraw their money out of a precautionary motive. Due to fire-sale discounts, such runs usually result in losses for the bank and the depositors. Interestingly, such runs can happen even when the fundamentals of the bank assets are sound Diamond and Dybvig (1983), reflecting coordination failures among depositors. The financial crisis of 2007-2009 highlighted the potential devastating effects of such liquidity risk.

There are two standard solutions to make banks safer: 1.) reducing risk on the asset side, and 2.) increase capital ratios on the liability side (see for example, Admati and Hellwig (2014) and Cochrane (2014)). The intuition as to why these measures work can be directly derived from a simple Merton (1973) model. Reducing risk on the asset side reduces the volatility of returns to the firm value, thereby increases the distance to default and decreases default risk. Similarly, Increasing capital ratios decreases leverage which in turn increases the distance to default and decreases default risk. Hence, reducing asset risk and increasing capitalization ratios are substitutes. Moreover, these effects are not specific to banks. They apply to any corporate as well.

Banks are, however, different from regular corporates in the sense that they are exposed to the risk of bank runs. If fundamentals are expected to be poor in the future and there are costs to the premature liquidation of long-term investments, depositors may prematurely run on the bank. This may result in 1.) premature defaults due to liquidation proceeds being insufficient to repay redemptions, or 2.)

eventual defaults because premature redemptions generated liquidation losses that are sufficiently large to compromise the repayment of depositors that did not redeem prematurely, irrespective of whether the long-term investment projects yield a high or low return. The prospect of situation 1.) would discourage depositors from running, while the prospect of situation 2.) would induce them to run.

We set up a bank run model along the lines of Goldstein and Pauzner (2005), but with two extensions. First, we explicitly model equity as a residual claim on the bank's balance sheet. As a result, equity serves as loss absorbtion capacity for depositors. Second, we allow for a range of different portfolio risk levels for banks that are disclosed in full to the public. As in almost all bank run models, the long term investments on the asset side have a binary payoff distribution.

An increase in asset risk (holding expected returns constant) would increase structural default risk and hence would make situation 2.) more likely and stimulate runs. Yet, higher risk also comes with higher returns in good scenarios, which serve as loss absorbtion capacity in situation 2.) and hence are a substitute for equity. Which of the two effects of increased risk taking dominates depends on the capitalization ratio of the bank. When the capitalization ratio is low, the marginal value of an additional unit of loss absorption capacity is high and increased risk taking will lower the probability of bank runs. When the capitalization ratio is high, the marginal value of an additional unit of loss absorption capacity is low and increased risk taking will increase the probability of bank runs.

Note that in the absence of bank runs, there would not be any mitigating effect of higher returns in good states on default risk since the higher returns would only materialize in the good state, in which they are irrelevant for avoiding default risk. Because bank runs can generate problems with regards to solvency even when a bank survives the initial run and the good state of nature materializes, higher returns in the good state of nature can help to prevent runs and thereby default risk.

Our model also yields relevant policy implications. Most prominently, it shows that the standard reaction of shedding risk upon entering distress may be counterproductive and stimulate rather than prevent runs. This is true in particular for financial institutions, such as investment banks, that are largely funded by uninsured deposits (or equivalents, see Gorton and Metrick (2012)). In those scenarios, increased risk-taking might help, but only if depositors are aware of that. In view of implicit bail-out guarantees, regulators are unlikely to let this happen. To the extent that they are, the observed risk-shifting upon entering distress may actually not only benefit the shareholders, but even depositors through staving runs. The most effective way to stave runs is to beef up equity, possibly in combination with shedding risk. However, due to debt overhang problems, this may be hard in practice, unless a government commits to doing this. As a result, we provide additional rationale for (implicit) bail-out guarantees to stave runs.

We contribute to several strands of literature. First and foremost, we contribute to the literature on bank runs. This literature has typically modeled the liability side of banks as deposits only Diamond and Dybvig (1983), Goldstein and Pauzner (2005). As a result, depositors that do not run are residual claimants. This setup is ineffective for investigating the effect of equity and other sources of loss absorption capacity on the prevalence of bank runs. Interestingly, there is one paper in this literature, Moreno and Takalo (2016), that also finds that the higher expected return that is associated with increased risk taking may lead to reduced run risk. However, the mechanism through which this result materializes is different. In Moreno and Takalo (2016) depositors that do not run become residual claimants and hence benefit from speculative returns. These speculative returns are also relevant in scenarios in which no depositors withdraw their deposits. Moreover, in their setup, it is impossible to analyze the effect of equity as loss absorption capacity and any interactions between equity buffers and increased risk taking.

We also contribute to the literature on the governance role of bank run risk. The risk of bank runs has been show to be effective in disciplining equity holders and bank managers from embezzlement of and absconding with investment proceeds (Calomiris and Kahn (1991)), and to stop the manager from creating hold-up problem by renegotiating his compensation package (Diamond and Rajan (2000) and Diamond and Rajan (2001)). Run risk has also been show to limit or prevent risk shifting by managers and equity holders at the expense of depositors Cheng and Milbradt (2012). Yet, this comes at the risk of excessive premature liquidation Cheng and Milbradt (2012), Wagner (2009). Our results show that when discipline with regards to risk shifting is most relevant for banks, namely when capitalization ratios are high, run risk is decreasing rather than increasing in risk taking. This result goes against the result by Eisenbach (2017) who uses a general equilibrium model with correlated asset returns to show that market discipline through runs is not very effective in upturns and excessively costly in downturns. In our model runs are less likely to materialize in downturns when capitalization ratios have suffered and are low.

Our results are unique to a Global Games implementation of bank runs. Global Games, as developed by Morris and Shin (2006), Rochet and Vives (2004), and Goldstein and Pauzner (2005) are a useful extension to the bank runs literature started by Diamond and Dybvig (1983) since these models manage to obtain unique equilibria rather than multiple equilibria. As a result, one can analyze the objective probability of a bank run materializing. Since we link the probability of a bank run to loss absorption capacity, we are constrained to working in a Global games framework.

3.2 Setup

In this section we outline the setup of the model. The model is very similar to the one by Goldstein and Pauzner (2005). We extend their model in two ways: we explicitly model equity on the balance sheet and we allow for a menu of risk choices to the bank. For comparability, we have tried to keep notation as similar as possible to that in Goldstein and Pauzner (2005).

3.2.1 The agents

The model is a two-period $(1, 2)$ model that involves a continuum of depositors and an equity holder. Abstracting from any agency problems, "the bank manager" has her interest perfectly aligned with the equity holder thus is not modeled as a separate agent. The capital structure of the bank is exogenous and fixed. The shareholders contribute $e \geq 0$ amount of equity, while the depositors have a total deposit amount 1 in the bank. As a result, the balance sheet has a total size of

$e+1$ and the leverage is $\rho \equiv \frac{1}{e+1}$. We assume that the depositors are of mass 1 and each of them contribute a unit amount of the deposit. The dispersed depositor base gives rise to the possibility of runs resulting from coordination failure. We index the depositors with $i \in [0, 1]$.

Like in Goldstein and Pauzner (2005), there are two types of depositors. A fraction $0 < \lambda < 1$ of depositors are "impatient" and the rest $1 - \lambda$ are patient. Impatient depositors can consume c_1 only in period 1 to obtain linear risk neutral utility c_1. Patient depositors can consume in either period 1 or period 2 to obtain utility $c_1 + c_2$.

To extend the Goldstein and Pauzner (2005) setup, we include in our model an equity holder. She can only consume in period 2. This late consumption assumption is consistent with the low seniority of equity claims: the equity holder cannot be paid before the patient depositors. The equity holder is risk neutral and maximizes her expected profit.

3.2.2 The asset

The bank has access to assets that yield a higher return long run. The universe of assets available can be thought of as a continuum of loans. At maturity, the asset pays off $1 + \mu$ if it is successful, and zero if it fails. The *probability of success* is θ.

Based on the risk return tradeoff, we further assume that assets with higher μ are more risky: θ is drawn from a uniform distribution on $[1 - \mu, 1]$. High values of μ on one hand represent higher upside returns, but on the other hand represent that the asset's success probability can become very low. μ is the loan specific "asset risk" that we focus on and it is exogenous. This specification of μ allows for the consideration of risk return tradeoff of the bank assets in global game models. It is similar to that introduced by Moreno and Takalo (2016), who investigate bank transparency instead and do not include bank equity in their model.

The above random payoffs are generated only if the asset is held by the bank until period 2. If instead liquidated in period 1, the asset generates instead a value $0 < l < 1$.

3.2.3 The deposit contract

In period 1, each depositor $i \in [0, 1]$ can choose one of the two actions: withdrawal or rollover. If she withdraws in date 1, her contract promises gross return $r_1 \geq 1$. On the other hand, if she rolls over her deposit, the contract promises the interest rate

$$r_1 < r_2 < 1 + \mu \tag{3.1}$$

in period 2. The second inequality must hold, otherwise, the period 2 promised return is so high, that even if the bank survives period 1 and the underlying project pays off, the remaining creditor will always get the entire loan payoff, leaving the equity holder nothing in every state of the world. We assume that both r_1 and r_2 are exogenous.

The return r_1 or r_2 is paid in full in the respective period only if the bank has enough cash in that period. Otherwise the we refer to it as "bankruptcy", in which case entitled depositors receive a *pro rata* share of the cash that the bank has at that moment. [1]

In addition, in period 2, the depositors that did not withdraw in period 1 does not get paid anything if the realized loan payoff is 0.

To make the problem interesting, we assume that if all depositors withdraw in period 1, then even if the whole loan is liquidated, the liquidation proceeds are insufficient for full repayment:

$$(1 + e)l < r_1. \tag{3.2}$$

Otherwise, bankruptcy can never happen, which removes the incentive to run.

[1] We'll discuss under what circumstance the bank has or do not have enough cash for repayment in the next subsection.

3.2.4 The information structure

In the start of period 1, the depositors observe the innate asset risk μ of the loan. However, they do not observe the realized repayment probability θ. Instead, at the start of period 1, each depositor i receives an i.i.d. noisy signal $\theta_i = \theta + \epsilon_i$ of θ. The noise ϵ_i is uniformly distributed over $[-\epsilon, \epsilon]$, where ϵ represents the precision of the signal. Following the global game approach in Goldstein and Pauzner (2005), adding this noise allows us to pin down the unique equilibrium of this game. If there are multiple equilibria, it is impossible to assign a probability to runs, thus impossible to study how risk choice affect the equity holder's profit by affecting the run probability. Other parameters such as l and e are common knowledge throughout the game.

3.2.5 The time line

The game of this economy is set up to have two periods:

In period 1, first, the random variable θ realizes. Next, each depositor i learns about whether she is a patient or impatient type. Then all depositors receive noisy signals of θ and decide whether to roll over their deposits. The bank liquidates part of or the whole loan to meet the withdrawal demand. If the liquidation proceeds are not enough to repay everybody, the bank goes bankrupt. In that case, each withdrawing depositor gets a *pro rata* share of the liquidation proceeds, whereas those that rolled over and the equity holder gets nothing. Otherwise, the bank carries on the investment with the non-liquidated portion of the asset, and the game continues.

In period 2, the asset either yields a high or a zero return. The unliquidated part of the asset pays off $1 + \mu$ with probability θ, in which case the bank pays the remaining depositors. Each depositor is paid r_2, if the total payoff is enough to do so, in which case the equity holder keeps the residual profit, if any. If the total payoff is not enough to repay all remaining depositors, the bank goes bankrupt and the depositors that did not withdraw receive *pro rata* shares from the final payoff.

Figure 3.1. Model Timing

$t = 1$

- The success probability θ realizes.
- Each depositor learns about her own type being patient or impatient.
- Each depositor receives a signal θ_i of the repayment probability.
- Each depositor decides whether she want to roll over or withdraw. Those that withdraw gets paid. Bankruptcy may occur, in which case the game ends.

$t = 2$

- The asset succeeds or fails. The asset yields return $1 + \mu$ with probability θ and 0 otherwise. The asset return is distributed to the remaining depositors and any remainder to the equity holder.

3.3 Bankruptcy and runs

Before defining the equilibria, we first make some preliminary analyses of the game. In particular, we determine how the bank can go bankrupt in our setup, and how the payoff of a depositor is affected by his rollover/withdrawal actions as well as the bankruptcy status of the bank. These factors are crucial for the depositors incentive to run and thus are necessary foundations of the equilibrium analysis.

In period 1, each depositor i learns about whether she is a patient or impatient type and then observes her signal θ_i. Each depositor then has the option to withdraw her deposit from the bank. Let $n \leq 1$ denote the endogenous fraction of depositors who choose to withdraw. Since the impatient depositors have to consume in period 1, it is optimal for them to withdraw in period 1 regardless of θ_i they observe. This implies the following condition on n:

$$n \geq \lambda \tag{3.3}$$

In period 1, bankruptcy occurs if the fraction n of withdrawing depositors is so large that even if the entire asset is liquidated, the proceeds are insufficient for full interest payment r_1. Mathematically, this condition is represented by the following condition:

$$(e+1)l < nr_1 \tag{3.4}$$

Condition (3.4) is equivalent to:

$$n > \dot{n} \equiv \frac{(e+1)l}{r_1}, \tag{3.5}$$

where $\dot{n}$ represents the critical mass that triggers period 1 bankruptcy.

Under condition (3.4), the game ends, each withdrawer gets a *pro rata* return $\frac{(1+e)l}{n}$, while the depositors that choose to wait and the equity holder receive 0.

If there are fewer withdrawers than the critical fraction $\dot{n}$, the bank only has to liquidate $\frac{nr_1}{l} < 1+e$ amount of asset to meet the demand of the n depositors. The rest $(1+e-\frac{nr_1}{l})$ of the asset carries on the investment and yields a random return in period 2: the period 2 total asset payoff is $(1+\mu)\left(1+e-\frac{nr_1}{l}\right)$ with probability θ and 0 otherwise.

Even if bankruptcy does not occur in period 1 and the investment continues to period 2, the bank may still go bankrupt under two circumstances. First, period 2 bankruptcy occurs if the project underlying the asset fails. The probability of this happening conditional on the game continuing to period 2 is $1-\theta$. Second, period 2 bankruptcy may also occur, even if the asset succeeds (with probability θ), if sufficiently many depositors chose to withdraw in period 1. In that case, even though period 1 withdrawals can be fully accommodated, the liquidation costs are so large that there is insufficient value left for depositors that did not withdraw early.

$$\left(1+e-\frac{nr_1}{l}\right)(1+\mu) < (1-n)r_2 \tag{3.6}$$

Condition (3.6) is equivalent to the following condition:

$$n > \tilde{n} \equiv \frac{(1+\mu)(e+1)-r_2}{\frac{(1+\mu)r_1}{l}-r_2}, \tag{3.7}$$

where $\tilde{n}$ represents the critical mass that triggers a run-induced period 2 bankruptcy. It can be shown that $\tilde{n} < \dot{n}$. The intuition for this is that the bank only survives till period 2 if n is sufficiently small.

To make the analysis interesting, we assume that the fraction λ of impatient depositors is small enough that unavoidable liquidation losses due to impatient depositors do not guarantee bankruptcy in period 1. The following condition ensures this to be the case:

$$(1+e)l > \lambda r_1 \tag{3.8}$$

In addition, we also make an additional assumption to ensure that unavoidable liquidation losses due to impatient depositors do not guarantee bankruptcy in period 1. The following condition ensures this to be the case:

$$(1+\mu)\left(1+e-\frac{\lambda r_1}{l}\right) > (1-\lambda)r_2 \tag{3.9}$$

The left hand side of this inequality is the payoff of the remaining asset in period 2 when the project succeeds, if all patient depositors wait until period 2 instead of withdrawing in period 1, whereas the right hand side is the promised principal plus interest for these patient depositors that waited until period 2. [2]

In case period 2 bankruptcy occurs under condition (3.6), the loan return is allocated *pro rata* so that each remaining depositor gets $\frac{(1+\mu)\left(1+e-\frac{nr_1}{l}\right)}{1-n}$ with probability θ while the equity holder is left with nothing. Otherwise each remaining depositor gets r_2 with probability θ and the equity holder get the residual with the same probability and nothing otherwise.

We summarize the payoff to a depositor, as functions of both her own decision to wait/withdraw, as well as the aggregate withdrawal fraction of her peers, n in Table 3.1.

[2] After algebraic transformation, assumption (3.9) can be alternatively written as: $\mu > \frac{(1-\lambda)r_2}{1+e-\frac{\lambda r_1}{l}} - 1 \equiv \underline{\mu}$, a condition that puts a restriction on the lower bound of μ. Assumption (3.9) can also be re-written as: $\frac{(1+\mu)(e+1)-r_2}{\frac{(1+\mu)r_1}{l}-r_2} > \lambda$, ensuring that condition (3.3) and condition (3.7) are

Table 3.1. Expected payoff patterns resulting from depositor withdrawal and roll-over decisions

withdraw in period	$\lambda \leq n \leq \tilde{n}$	$\tilde{n} < n < \dot{n}$	$\dot{n} \leq n \leq 1$
1	r_1	r_1	$\frac{(e+1)l}{n}$
2	θr_2	$\theta \frac{(1+\mu)\left(e+1-\frac{nr_1}{l}\right)}{1-n}$	0

3.4 Equilibrium

3.4.1 The depositor's problem

In period 2, each depositor $i \in [0,1]$ determines whether to roll over or withdraw her deposit. Each patient depositor seeks to maximize her expected consumption $c_1 + c_2$. Therefore it's in her best interest to wait only if the payoff from roll over exceeds that of withdrawal, or equivalently, the utility differential $v(\theta, n(\theta))$ between rollover and withdrawal is positive. Based on the payoff from rollover and withdrawal we derived in Section 3.3 and summarized in Table 3.1, $v(\theta, n(\theta))$ should have the formula:

$$v(\theta, n(\theta)) \equiv \begin{cases} \theta r_2 - r_1 & \text{if } \lambda \leq n(\theta) \leq \tilde{n} \\ \theta \dfrac{(1+\mu)\left(e+1-\frac{n(\theta)r_1}{l}\right)}{(1-n(\theta))} - r_1 & \text{if } \tilde{n} < n(\theta) < \dot{n} \\ 0 - \dfrac{(e+1)l}{n(\theta)} & \text{if } \dot{n} \leq n(\theta) \leq 1, \end{cases} \tag{3.10}$$

where critical values $\dot{n}$ and $\tilde{n}$ are defined by conditions (3.5) and (3.7) respectively.

But since in our game the depositors do not observe θ or n, they are not able to compute $v(\theta, n(\theta))$ directly to inform their decision. The best a depositor can do is to estimate the θ and $n(\theta)$ based on her signal θ_i and compute the expected value of $v(\theta, n(\theta))$ accordingly:

$$\Delta(\theta_i) \equiv \mathbb{E}\left(v(\theta, n(\theta))|\theta_i\right) \tag{3.11}$$

not redundant.

Then she rolls over if and only if $\Delta(\theta_i)$ is positive.

The next step of the thought process of any depositor is to have an idea of the decision rule of all other depositors, so that she knows the functional form of $n(\theta)$ in terms of θ. We address this problem now in the Subsection 3.4.2.

3.4.2 Equilibrium definition

In this paper, we restrict attention to a special type of equilibria: the ones in which the decision rules of the depositors are "threshold strategy": each depositor i withdraws whenever her signal is lower than a certain threshold θ_i^*, and roll over if the signal is higher than this threshold. A threshold strategy is consistent with the behavior of depositors in the real world: they tend to withdraw their deposit if they believe the chance of asset success is low enough, and will also withdraw if their belief of the success probability is even lower.

In addition, since depositors are *ex ante* identical, we only consider symmetric equilibria in which they apply the same threshold: $\theta_i^* \equiv \theta^*$ for any $i \in [0, 1]$.

In summary, we define the equilibrium as:

Definition 5 *A symmetric, threshold equilibrium of the game is characterized by a run threshold θ^* of the depositors, such that, in period 2, if all other depositors apply the threshold θ^*, depositor i also find it optimal to also apply threshold θ^*, in the sense of maximizing $\mathbb{E}\left(v(\theta, n(\theta))|\theta_i\right)$.*

3.4.3 Equilibrium uniqueness

As can be seen from Subsection 3.4.1 that the signal θ_i is informative to depositor i in two ways: not only does it inform the depositor about the probability of asset successful, but also allows her to refine her expectation of the distribution of the other depositors' signals, thereby inferring the fraction of withdrawers $n(\theta_i)$.

Similar to Goldstein and Pauzner (2005), we assume that there exist extremely good (bad) states of θ, such that it is optimal for depositors to wait (withdraw), respectively, based on the good (bad) probability alone, regardless of what other depositors do. We term the regions of such extremely bad and extremely good states "upper (lower) dominance region".

"Upper dominance region" is the range of θ such that the asset prospect is so good such that no depositor would withdraw even if all other fellow depositors withdraw. Similar to Goldstein and Pauzner (2005), we construct this region by assuming that for any $\theta \geq \bar{\theta} \equiv 1 - 3\epsilon$, such that for any $\theta \in [\bar{\theta}, 1]$, the asset success probability is 1, and the asset generate its return $1 + \mu$ in period 1 instead of period 2. In other words, when $\theta \geq \bar{\theta}$, to meet interim withdrawal demand, instead of having to liquidate part of the long term asset for repayment, the bank can instead satisfy the redemption request with the asset return that has been delivered ahead of schedule. These extra assumptions can be interpreted that assets with high success probability tend to mature earlier as well. This construction underlies the proof of equilibrium uniqueness in Goldstein and Pauzner (2005).

Under these added assumptions, when $\bar{\theta} + 2\epsilon \leq \theta < 1$, then for any depositor i, her signal $\theta_i > \bar{\theta} + \epsilon$. As a result of her knowledge of the size of the noise, she knows that θ has to be bigger than $\bar{\theta}$. Consequently, even if all the other depositors withdraw in period 1, depositor i believes there is still $(1 + \mu)(1 + e) - r_1$ amount of cash left, which by assumption 3.1 is more than $(r_2 - r_1)(1 + e)$. Considering her own infinite small size, this amount is more enough to repay depositor i with r_2 if she choose to wait. Therefore indeed waiting is a dominant action of hers regardless of the actions of everybody else. Recall that this argument holds for any depositor i. Thus under these assumptions, all depositors would choose to wait when $\theta > \bar{\theta} + 2\epsilon$.

Though the existence of a neighborhood $[\bar{\theta}, 1]$ is necessary for the proof of equilibrium uniqueness[3], the size of this neighborhood can be infinitely small as the noise in the signal vanishes ($\epsilon \to 0$), since $\bar{\theta} \equiv 1 - 3\epsilon$. Hence, we can come arbitrarily close to a setting with perfect information. As a result, when computing the equilibrium quantities and comparative statics, we do not have to worry about the complications introduced by this assumption. All the results obtained from Section 3.3 and Subsection 3.4.1 still apply.

Similarly, "lower dominance region" is the range of θ such that the asset prospect is so bad such that every depositor would withdraw even if all other fellow depositors

[3] Goldstein and Pauzner (2005)'s argument

wait. This regions is constructed by finding a $\underline{\theta}$, such that if a depositor i believes that $\theta \leq \underline{\theta}$ with certainty, then she would withdraw in period 1 even if all other depositors choose to wait. We found this $\underline{\theta}$ as:

$$\underline{\theta} \equiv \frac{r_1}{r_2} \tag{3.12}$$

When a depositor believes that $\theta \leq \underline{\theta}$, she realizes that even if everybody else wait, the payoff of herself waiting is $\theta r_2 \leq \underline{\theta} r_2 = r_1$, but the payoff of withdrawing r_1 dominates waiting. To convince her that $\theta \leq \underline{\theta}$ for sure, her signal must be $\theta_i \leq \underline{\theta} - \epsilon$. And to make sure all depositors receive such signals, the realized state θ must satisfy: $\theta \leq \underline{\theta} - 2\epsilon$. However, recall that by our setup, $\theta \in [1 - \mu, 1]$, thus to ensure such θ exists, we need the assumption that:

$$1 - \mu \leq \frac{r_1}{r_2} - 2\epsilon. \tag{3.13}$$

This assumption puts a upper bound on the values of μ.

With the additional assumptions that allow for the construction of the upper and lower dominance region, there exists a unique symmetric threshold equilibrium.

Proposition 4 *This game has a unique symmetric threshold equilibrium.*

Proof. See Appendix. □

From now on we write the unique common threshold of patient depositors as θ^*. Now that we have proved the equilibrium uniqueness, we can derive outcome variables as a function of θ^* and be sure that they have unique values. For example, for each realization of the state variable θ, we can compute the fraction n of withdrawers by:

$$n(\theta, \theta^*) = \lambda + (1 - \lambda) \cdot \text{prob}[\theta + \epsilon_i < \theta^*] \tag{3.14}$$

Since ϵ_i is uniformly distributed on the interval $[-\epsilon, \epsilon]$, n can be evaluated according to the following corollary:

Corollary 6 *Given asset risk μ and equity e, the proportion of agents that withdraw before maturity depends only on the realized state of the asset's prospect. It is in the form:*

$$n(\theta, \theta^*(e,\mu)) = \begin{cases} 1 & \textit{if} \quad \theta < \theta^* - \epsilon \\ \lambda + (1-\lambda)\min\left(\left(\frac{1}{2} + \frac{\theta^*(e,\mu) - \theta}{2\epsilon}\right), 1\right) & \textit{if} \quad \theta^* - \epsilon \leq \theta \leq \theta^* + \epsilon \\ \lambda & \textit{if} \quad \theta > \theta^* + \epsilon \end{cases} \tag{3.15}$$

From this point on, like in Goldstein and Pauzner (2005), we derive most of our results at the limit when $\epsilon \to 0$. As $\epsilon \to 0$, by construction $\bar{\theta} \to 1$, thus the irregularities of the θ value and the payoff vanish as well, which allows us to use the preliminaries derived in Section 3.3 and Subsection 3.4.1 for our derivations from this point on.

Since we are interested in how asset risk μ affect the probability of bankruptcy of period 1 and period 2, first derive an analytical expression for the probability of bankruptcy π. Similar to Goldstein and Pauzner (2005), the run probability is intricately related to the equilibrium threshold $\theta^*(e,\mu)$. Therefore, we first give the analytical form of $\theta^*(e,\mu)$.

Lemma 7 *As the error in signal ϵ vanishes, equilibrium threshold has the formula:*

$$\theta^* \xrightarrow{\epsilon \to 0} \frac{lk(e)}{z(e,\mu)}, \tag{3.16}$$

where to condense the notation, functions $k(e)$ and $z(e,\mu)$ are explicit functional forms of the parameters defined as:

$$k(e) \equiv l(e+1)\left(1 - \log\left(\frac{l(e+1)}{r_1}\right)\right) - \lambda r_1 \tag{3.17}$$

and

$$z(\mu, e) \equiv -\big(r_1 - l(e+1)\big)\,(\mu+1)\log\left(\frac{(\mu+1)r_1}{(\mu+1)r_1 - lr_2}\right) + (1-\lambda)\,lr_2 \qquad (3.18)$$

Proof. See Appendix. □

In Goldstein and Pauzner (2005), the equilibrium threshold *is* the probability of bankruptcy. That is because state variable θ is uniformly distributed on interval $[0, 1]$. But θ in our paper is uniformly distributed on $[1-\mu, 1]$ instead, so we need a minor algebraic modification of θ^* to obtain the probability of bankruptcy.

As discussed earlier, the bank in our model can go bankruptcy in either period 1 or period 2. The question is Which probability to consider? It turns out that this is a non-issue, because the two probabilities converge as the signal noise vanishes.

Lemma 8 *As the error in signal ϵ vanishes, the probability that the bank goes bankrupt in period* 1 *is the same as the probability that it goes bankrupt in period* 2*. This unique probability, named π has the form:*

$$\pi \equiv prob(\theta \leq \theta^*) \xrightarrow{\epsilon \to 0} \frac{\frac{lk(e)}{z(e,\mu)} - 1 + \mu}{\mu}, \qquad (3.19)$$

Proof. See Appendix. □

The intuition behind the coincidence of period 1 and period 2 bankruptcy probability is that, as the depositors get infinitely accurate information, they would preemptively withdraw in period 1 once they learn that bank will go bankrupt in period 2, leading to period 1 bankruptcy. They are not able to do this when the signal is noisy enough, since in that case the depositors that received signal with positively biased noise would not believe that the bankruptcy will occur in period 2, and even the depositors that receive a signal that happens to be the same as the true θ are not sure about the signal either. In summary, the noise put a damper on the preemptive withdrawal in period 1. In light of this result, from now on, we refer to period 1 and period 2 bankruptcy collectively as “bank runs”, and π is the “run probability”.

One should notice that since most other bank run models do not feature equity, but have non-withdrawing depositors as the residual claimants, this result is unique to our setting.

3.5 Comparative statics

Taking derivative of the equilibrium run probability π with respective to the loss absorption capital e and risk profile μ allows us to investigate how the likelihood of runs change with respect to these two parameters.

First we derive how equity buffers affect the run probability π. As one might expect, the the larger the capital buffer, the less likely the runs, since equity reduces the likelihood of run-induced defaults in period 2 and thereby mitigates the incentive to run.

Proposition 5 *As $\epsilon \to 0$, run probability π strictly decreases in capital amount e*

Proof. See Appendix. □

We then continue by deriving the main result of our paper concerning the comparative static of the run probability with respect to the asset risk.

Since π is a function of the run threshold θ^*, we first derive the sensitivity of θ^* with respect to μ:

Lemma 9

$$\frac{\mathrm{d}}{\mathrm{d}\mu}\theta^*(\mu) < 0 \tag{3.20}$$

Proof. See Appendix. □

The intuition for this result is as follows. μ affects the threshold θ^* via two channels. On one hand, a larger μ is destabilizing, that is, it leads to higher θ^*. This is because it heighten the tipping point withdrawer fraction $\tilde{n}$ above which the bank becomes insolvent in period 2 thereby those that wait receive payoff less than the promised r_2, making runs more likely. On the other hand, increasing μ can also discourage runs by increasing the payoff to those that wait when the bank goes bankrupt in period 2: $\frac{(1+e-\frac{nr_1}{l})(1+\mu)}{1-n}$. This positive effect is due to the period 2 payoff when the asset payoff is not enough to pay r_2, and the depositor effectively becomes the residual claimant. Since any increase in period 2 payoff would lead to fewer runs, this second channel is a stabilizing effect of higher asset risk. Lemma 9 demonstrates that overall this stabilizing effect is larger in magnitude than the destabilizing effect mentioned earlier in this paragraph.

But this is only one side of the coin. Note that θ^* does not capture all the negative effects of higher asset risk. High asset risk means that the lower bound of the bad states θ is lower, as $\theta \in [1-\mu, 1]$, and hence, there would be more scenarios in which the asset payoff equals zero. This effect is captured by π, as can be seen from the $-(1-\mu)$ term in the numerator in formula (3.19). To weigh the cost of high risk against the benefit of high return in the good state, we analyze the mathematical form of the derivative of π with respect to μ. It turns out that when equity e is low enough, the stabilizing effect dominates; higher asset risk can in fact lead to less runs. The reason is that when equity is very low, the marginal benefit of an additional unit of period 2 loss absorption capacity (provided by higher good state returns) is very high and exceeds the negative effect of a high propensity of states with zero payoff.

Proposition 6 *As $\epsilon \to 0$, there exists an upper bound of capital level $\bar{e}(\mu)$, such that, for any $e \in \left(\frac{\lambda r_1}{l} - 1, \bar{e}\right)$, run probability π decreases in asset risk μ. In particular, such relationship may occur for $\mu > \frac{1}{2}$*

Proof. See Appendix. □

The interpretation of this result on a more detailed level is as follows. The sign of $\frac{\partial \pi}{\partial \mu}$ depends on whether the stabilizing effect of higher μ, represented by the first term in the formula of π dominates the destabilizing effect, represented by the second term. Let us first look at the first term. As discussed earlier, higher μ stands for both higher failure risk and higher period 2 returns in case of success. Higher equity level, on the other hand, indicates that equity holders provide insurance for the depositors. Thus higher e and the higher conditional return aspect of higher μ substitute each other in attracting the depositors to wait until asset maturity. This is why the sensitivity of the first term, θ^*, to μ can get very big when e is smaller. The second term of π representing the destabilizing effect, however, is not a function of e thus its negative effect turns out be offset by the stabilizing effect when e is small enough.

It is important to note that this interesting result may occur for $\mu > \frac{1}{2}$. One can show that the socially optimal risk choice equals $\frac{1}{2}$.[4] When $\mu > \frac{1}{2}$, aggregate welfare decreases in μ, meaning that the asset starts to become negative marginal NPV, if μ increases further. So this aspect means that run probability may decrease in asset risk, and consequently runs may encourage risk taking, even when such risk taking is NPV destroying.

[4]One would need to maximize $(1+\mu)(1-\frac{1}{2}\mu)$

Chapter 4

CEO evaluation frequencies and innovation

Abstract

I focus on an important aspect of corporate governance that has been challenging to empirically identify – the frequency at which the CEO performance is evaluated by the shareholders. Because trial and error is a time-consuming process but is crucial to produce important innovations, evaluation frequency is an important determinant of innovation, an insight provided by Manso (2011). Taking advantage of the 2011 SEC rule change that requires all public US companies to vote on the "say on pay" frequency, I apply sharp regression discontinuity design (RDD) to show that firms with less frequent evaluations generate more innovation in the long run. Furthermore, my evidence suggests that the underlying mechanism is that low evaluation frequency allow costly "explorations" in a broad spectrum of technological fields in the beginning. Later on, the firm benefits from the exploration by "exploiting" the knowledge acquired during exploration, as shown by my results. The patent "truncation problems" that traditionally limit investigations of *recent* (here 2011 onward) innovation are ameliorated by the collection and usage of patent *application* data.

4.1 Introduction

Corporate governance has first order effect on firm innovation. This notion is supported by evidence in various contexts[1], and underlies the JOBS (Jumpstart Our Business Startups) Act. Among various aspects of corporate governance, CEO performance evaluation frequency is of particular importance, as predicted by theories. Stein (1988) points out that the short term performance pressure resulted from very frequent evaluations may result in CEO Short-Termism, including the abandonment of innovation. Recently, Manso (2011) shows that the incentive scheme that best motivates innovation tolerates short term failure and rewards long term success. High value innovation is often accompanied by great uncertainty. Its completion requires “exploration” as the first stage, which serves to acquire intellectual capital but does not generate measurable short term performance. Its great benefit is actualized in the “exploitation” stage when the firm is able to produce tangible innovation outputs (patents, etc.). But when the evaluation frequency (inverse of the “frequency” introduced above) is too high, the shareholders of the firm are unable to determine whether the temporary lack of performance is the result of exploration or shirking. In equilibrium, the friction caused by the short term shareholders oversight may discourage the CEO from exploring, which means the firm is unable to pursue high value innovation. The competing “quiet life theory”[2] by Bertrand and Mullainathan (2003) makes the effect of evaluation frequency on innovation ultimately an empirical question. But so far it has been difficult to measure let alone identify the evaluation frequency directly[3]. The lack of direct evidence makes the impact of evaluation frequency on innovation controversial. For example, this issue is a subject of debate in the 2016 US presidential election[4].

[1] Notable examples are Seru (2014) and Bernstein (2015).

[2] Without short term evaluation the CEO would simply shirk and not innovate.

[3] Natural experiments usually generate exogenous variations in the *intensity* of governance but rarely in the *time length* of evaluation frequency. The closest proxy for the evaluation frequency have been active institutional ownership and public ownership in general. But these proxies are not able to exclude channels other than the frequency effect. The empirical studies on this issue are reviewed by Kerr and Nanda (2014) and Chemmanur and Fulghieri (2014).

[4] For two opposing views expressed in the Wall Street Journals, see: “The Imaginary Problem of Corporate Short-Termism.” (Roe (2015)) and “Clinton Gets It Right on Short-Termism.”(Galston (2015))

In this study, I identify the CEO evaluation frequency by exploring a unique SEC rule change made in 2011. All public firms in the US were required to hold a shareholder vote in 2011 on the *frequency* of the"say on pay"(SOP) votes. The outcome of this vote determined how frequently the shareholders would discuss and vote to approve/disapprove the executive compensation proposal. To identify the causal effect of interest, I apply regression discontinuity design (RDD) to compare firms whose shareholders voted for a lower frequency of SOP by a small margin to those that voted for a higher frequency by a small margin. Therefore the firms that lie close to either side of the vote share threshold can be considered identical in all aspects but the choice of the evaluation frequencies. I use SOP as a proxy for corporate governance in general because its effect extends beyond executive compensation. The SOP usually occurs as the last session of the shareholder annual meetings, by which time the shareholders have heard all items on the agenda and can use the SOP as a tool to punish or reward the CEO for all governance issues. As a result, CEOs of the companies that determined to hold such votes once every three years can be considered evaluated less frequently than those that do the votes every year. Furthermore, unlike most other shareholder votes, the vote on SOP frequency is not manipulated by the management. In addition, it is implemented 99.7% of the time in my sample. These two features allow for a clean *sharp* regression discontinuity design (RDD), eliminating concerns of endogeneity.

To address a long lasting challenge in the measurement of innovation, I apply a *new* methodology and collect new data. Previously, due to "truncation problems" with the patent data, researchers have only been able to evaluate an innovation at least 10 years after the patent is granted (Lerner and Seru (2015)). An analogous problem is also found in the academic publication process. When evaluating a researcher's *recent* productivity, one counts not only his published articles, but also his pre-publication working papers, since most of his recent work is still in this stage. Similarly, the pre-grant patent application data is crucial in the case of corporate innovation. However, it has not been collected into one database or matched with company identifiers such as CUSIPs. This makes *timely* (here, a shock in 2011) evaluation of innovation difficult. Thus, I collect patent *application* data via large scale data collecting and show through backtesting that by adding the application data, I am able to evaluate innovations as recent as three years with accuracy and precision.

I assemble a sample of 2247 firms that held the say on pay frequency vote and have patenting activities. In 2011, all the firms held[5] the SOP frequency vote in their annual meeting to determine the voting frequency of the SOPs. The options given by the SEC were either "every year" ("distracted firms") or "once every three years" ("undistracted firms"). The first installment of the SOP was held in 2011, and the second installment was to occur in 2012 for the distracted firms and in 2014 for the undistracted firms. Thus, the undistracted firms had a three-year undistracted period from 2011 to 2014, while the distracted firms did not. I apply RDD by comparing the innovation of these two groups of firms year by year throughout the undistracted period (2011 to 2014) for the undistracted firms. Any difference in innovation between these two groups during these years is a result of their different evaluation frequency (one year versus three years).

Consistent with the predictions made by Manso (2011), I find a dynamic relationship between the frequency of SOP voting and the value of innovation outputs. The undistracted firms produce patents at similar quantities and scientific values (measured by the number of citations received) in the first year of the undistracted period, but in the third year, this trend is reversed. Indeed, at this point in the undistracted period, the undistracted firms apply for more patents that garner more citations in the following 4 years than the distracted firms. Therefore in the end, the undistracted firms surpass the distracted firms, and this does not take into account the potentially greater benefits it may experience after the three year period. These results suggest that having the SOP meeting less frequently leads firms to make more valuable innovations in the long run.

What leads to this dynamic effect? To shed light on the underlying mechanism, I compare the firms' emphasis on exploration versus exploitation year by year. I find that in the first year, the undistracted firms are more likely to explore less

[5] A small number of firms held it in 2013. I will discuss this in detail in section 4.3.

familiar technological fields, measured by the "Exploitation Index" as well as the Herfendahl diversity of the cited fields (measures proposed by Trajtenberg et al. (1997)). This result suggests that without short term pressure, firms are able to spend time exploring technological fields that they have less expertise in previously. Taken together with the previous results, these initial exploration may explain the increased ability of the undistracted firms to produce more valuable patents over the years. It also confirms the empirical prediction by Manso (2011) that suggests that exploration manifests its benefit only in the long term and is tolerated only by firms with low evaluation frequencies. What is more, the possible change of technological focus suggests a intrafirm Schumpeterian creative destruction, which on aggregate leads to progress in technology.

As already explained throughout the above text, this paper makes two contributions. Not only is it the first study to empirically identify the causality of CEO evaluation frequency on innovation, but it develops a new data set and methodology to enable the study of recent innovations.

This paper relates to three strands of literature. First, it joins the debate on the role of corporate governance and ownership structure in innovation. The empirical evidence on this issue is divided. For example, Atanassov (2013) finds that innovation falls in the states that pass antitakeover laws while Chemmanur and Tian (2018) find that innovation is higher in firms with more antitakeover defenses. Similarly, Acharya and Krishnamurthy (2009) find that lenient bankruptcy laws promote innovation, suggestive of a greater willingness to take on risky projects with stronger downside protection. However, Mann (2018) finds that stronger creditor rights are associated with greater innovation by firms in his sample. My findings establish that the effect of corporate governance depends on the evaluation frequency.

In particular, it fills the gap between the theoretical and empirical works on the relationship between CEO incentives and firm innovation. As predicted by Manso (2011), there are two necessary conditions of CEO incentive structure that stimulate innovation: short term tolerance of failure and long term reward for success. Without the tolerance for early failure, CEOs do not make risky but potentially valuable innovations, without reward for long-term success, CEOs will not have an incentive to innovate and simply shirk. Thus it is important to test both of them simultaneously. Gonzalez-Uribe and Xu (2017) look at the impact of CEO contract length on innovation, which only captures short term tolerance for failure but not long term reward for success. A CEO's job is secure for the length of her contract, but the tenure itself is not designed to provide rewards for potential success. In my setup, however, these two necessary conditions can be tested together. Low say on pay *frequency* is a commitment not to cut the pay in case of short term failure, whereas say on pay itself serves as a tool to effectively increase the pay of the CEO in case of long term success, by approving the proposal for a raise.

Second, with regards to the underlying mechanism of this differing effect, this study relates to the literature on the trade-off between two types of innovations: exploration and exploitation. Exploration refers to the process in which the innovator experiments with various potential fields or strategies to identify the most promising one. In contrast, exploitations build on the expertise and knowledge acquired from explorations. Recent findings highlight the important role of exploration in generating fundamental innovations. (Kerr et al. (2014) and Manso (2011)) In particular, the theoretical model proposed by Manso (2011) predicts that the most effective contract to motivate innovation tolerates short term failures and rewards long term success. This study tests this prediction with rich dynamism enabled by novel measures of exploration and the time length specific identification strategy.

Third, by using say on pay as a proxy for corporate governance, this paper relates to the literature of the effect of say on pay votes. For example, Cuñat et al. (2012) uses RDD to study the stock market reactions to say on pay vote results. In contrast, my study is about the vote on SOP frequency not SOP vote itself. The variation in frequency is necessary to identify the governance horizon. I also supplement this literature by revealing dynamic actions the firms take between two consecutive SOPs.

The rest of the paper proceeds as follows. Section 4.2 describes the data and measures. Section 4.3 demonstrates the validity of the identification strategy. Section 4.4 provides the main results. Section 4.5 reports the robustness tests conducted.

Section 4.6 discusses the main limitations of this paper. Section 4.7 concludes.

4.2 Data and measures

4.2.1 Patent Data

4.2.1.1 Truncation problems with the patent data

This paper introduces the usage of patent application data into the study of corporate innovation. Traditionally, researchers have relied only on granted patent data. But due to "truncation problems" inherent in this data, a firm's innovation output can only be assessed 10 years after it is produced which makes it impossible to evaluate impacts of recent events affecting innovation. With the addition of application data, however, I show that innovation can be measured with reasonable accuracy, even for events that have occurred as recent as 2013. In this section, I will explain what I mean by "truncation problems" and describe how patent application data can solve them.

The formal patenting process begins with the filing of an application. Patent status is awarded or denied based on an investigation into the originality of the invention, which takes 34 months on average. Because of this extended processing time, a company's most recent inventions typically remain in the application stage and do not appear in the patent database. In this sense, the patent count, which represents a company's innovation, appears truncated in the database, as demonstrated by the left panel of figure C.3. In the figure, the x-axis represents the year in question and the y-axis the average number of patents per firm that were *applied* in the year in question and granted by the end time of the database. The red dots represent the number of patents counted in 2006, whereas the black dots represent the patents counted in 2016. Hall et al. (2005) demonstrate that inventions that are granted patent status receive their status within ten years after the initial application. Therefore for the same firm at the same filing year, the black line represents the eventual accurate patent count, whereas the red line represents what researchers could construct with data available in 2006. The widening gap between the two plots appears as if the firms' innovative activities declined as the sampling period came to an end, but this truncation is actually an artifact. The company's recent patent patent applications have not been granted patent status yet, and thus are not collected by the patent database.

A related truncation problem exists in regards to citations that a patent receives, which are often used as a measure of the scientific value of the patent. Patents garner citations over time, and a large number of citations they receive in the patent's infancy come from other applications that have not been awarded patent status; though, citations are also received from other patents. In addition, a patent starts to receive citations when it is still an "application", from both other applications and other patents. In total, there are four sources of citations that a firm's inventions receive in a certain year (Figure C.2), but without the application data, researchers are only able to observe those from the first source (arrowhead A). This is especially an issue for the recent innovative activities, since most of the innovation outputs are still in the "application" stage thus missing from the database. The right panel of Figure C.3 illustrates the truncation problem with the citations. One can control for the time fixed effect to eliminate resulted biases, but the small number of citations and patents counted make the measurement very noisy. Traditionally, researches have waited at least 10 years after a patent application was filed to assess its importance, by which time the small sample problem has been compensated by the accumulated citations over the years.

4.2.1.2 Adding the application data

There are two challenges in using the application data. First, the application filings warehoused by Google Patents were in the form of around six million individual webpages rather than one dataset, and the company names, if provided, are not standardized and may include typos. For example, the company "Hewlett-Packard" has as many as 34 name variants, including "HP", "HP INC." and "Hewett-Packerd". Second, these patents are not matched with company identifiers such as the CUSIP, making it difficult to merge them with financial databases.

I overcome these challenges using Python scripts to perform web crawling to obtain patent and application information. I then match the names with CUSIPs using the pre-existing matches provided by the NBER patent database (constructed by Hall et al. (2005)), assisted by the disambiguation dataset provided by PatentsView. I also use manual matching to overcome incomplete names from the patent database. Further details on the process are described in the appendix.

To confirm that the added application data allows me to accurately measure recent innovation, I conducted “backtesting”. This methodology has been used primarily in asset pricing studies to test whether an investment strategy will be profitable in the future by assessing whether it *would have* been profitable had it been applied in the past. An important caveat lies in that the only information that can be used in the testing is the information available at the time the testing is intended to emulate. I use the same principle to conduct backtesting of the innovation measures in three steps. First, for all the firms, for the years $2001-2006$, I construct measures both with the application data (the “supplemented measure”) and without (the “raw measure”). Additionally, only the information available in 2006 was used. In other words, any application that was granted after 2006 is counted as an application rather than a patent. Similarly, any citations that occurred after 2006 are not counted. Then for the same firms of the same years $(2001-2006)$, I construct the innovation measures using only patent data, but using the information that is available as of 2016. In other words, I count the applications that eventually get patent status as well as citations these patents received after 2006. The measure constructed in this step, which allows for a 10-year accumulation period is what have been used in the literature. It has the disadvantage of a long wait, but should be precise and accurate. Thus it serves as the benchmark against which the supplemented and raw measures are evaluated. Finally I compute the Pearson correlation coefficients between the supplemented and raw measures and this benchmark. The difference between these two correlations signifies how accurate the supplemented measure (including the application data) is compared to the raw measure (without the application data) that this paper creates is relative to the raw measure. If this advantage is big enough, we can be confident to extrapolate that, the supplemented measures that I compute for the year 2011, with the data available as of 2016, accurately predict the citations and patent counts that this firm will receive in the long run, say, by 2026.

The backtesting shows that the supplemented measure computed with the added application data features less bias and less noise compared to the raw measurement that does not include application data. This improvement in accuracy is demonstrated in Figure C.4 by the alleviated truncations (green dots) for both patent counts (upper panel) and citation counts (lower panel), relative to their raw counterparts (red dots). Aside from accuracy, the assessment of precision is reported in Table C.1, which tabulates the Pearson correlation coefficients between the two tested measures and the benchmark. As the literature rightly worries, the raw measure correlates poorly with the true value when the patent is very recent, which is represented by its low correlation with the true value for all five tested years, especially the last three. In contrast, the supplemented measures correlate very well with the benchmark, especially for the first three years. Repeating this backtesting with information sets available in $2007, ..., 2015$ respectively (results not shown), it can be shown that the "complete measure" computed for the $2001-2003$ using the information available as of 2006 is as accurate as the raw measure computed as late as 2011. This makes us confident that the supplemented measures computed in this paper for the years $2010-2013$ is accurate. The measure for 2014 is noisier but still would have a high accuracy.

4.2.2 Measures of innovation

In this subsection I describe rigorously how the supplemented measures are computed with the added application data, and why they are less biased and less noisy.

4.2.2.1 "Supplemented measures" of patent counts and patent citations

Before explaining these two supplemented measures, I first describe the measures that have been used in the literature, which I refer to as the "raw measures". These measures are precise only when they are computed for patents and applications filed at least 10 years before the end of the sample period of a studyLerner and Seru (2015). They serve as the benchmark to which the supplemented measures are compared and contrasted.

"Raw patent count" $P_{j,t}$ is simply the total number of patents that firm j filed in year t, that are granted patent status as of year T, the last year of the available patent data:

$$P_{j,t} \equiv \sum_{p \in \mathbb{P}_{j,t}} 1 \tag{3.1}$$

,where $\mathbb{P}_{j,t}$ represents the set of granted patents that firm j filed in year t [6]. This measure has been used to represent the *quantity* of the innovation output that a firm produces in a certain year. But, it suffers from the truncation problem, since as t is very close to the end of the sample period T, most patents are still at the stage of "applications", and are thus accounted for by this measure.

"Raw average citation count" $\bar{C}^{P,P}_{j,t}$ is the average normalized number of citations that each patent in $\mathbb{P}_{j,t}$ receives from other granted patents as of the end of the sampling period T:

$$\bar{C}^{P,P}_{j,t} \equiv \frac{\sum_{p \in \mathbb{P}_{j,t}} \frac{C^{P,P}_{p}}{\bar{C}^{P,P}_{f,t}}}{P_{j,t}} \tag{3.2}$$

, where "P, P" super scripts represent the fact that the citations counted are only those from granted patents to granted patents. Thus, $C^{P,P}_{p}$ represents the citations that patent p receives from other patents. $\bar{C}^{P,P}_{f,t}$ represents the average number of citations that are received by other patents filed in the same technological field in the same year. The division by this cohort average serves to control for the substantial cross field and time series variation of the patent citations Lerner and Seru (2015).

As alluded to in the previous section, the patent-to-patent citations counted by this measure only account for one of the four types of citations, which is represented by the A type in Figure C.2. As t approaches the end of the sampling period, most of the patents are still applications, which results in this measure ignoring a large

[6] In this paper, I use "Holy Roman bold" font to represent sets.

part of total citations.

In summary, since they are constructed using only the patent data, the raw measures suffer from small sample problems when constructed for recent patent and applications, and this reduces the measure's statistical precision.

Because of the limitation inherent in using only patent data, I add application data to my analytic repertoire. However, patent applications should not be given the same weight as granted patents in quantifying innovation. After all, only 56% of all applications are eventually granted patent status. This implies that on average, applications are of less value than granted patents. Thus, to the extent that the number of eventually granted patents is the most accurate measure of innovation, the supplemented measure would generate a positive bias if the applications were used without any discount. This is analogous to the fact that when evaluating someone's research output, her working paper is less of a signal than her published papers, *ceteris paribus.*

To implement this discount to the applications, I derive the "application discount factor" $\beta_{\tau,f}$. It is the conditional probability that any patent application that hasn't been approved in τ years will be approved eventually. It is estimated by:

$$\beta_{f,\tau} = \Pr(\mathscr{A} \cup \mathscr{B} | \bar{\mathscr{A}}) = \frac{\Pr(\mathscr{B})}{1 - \Pr(\mathscr{A})} \tag{3.3}$$

$$= \frac{\sum_{l=\tau+1}^{\infty} g_{\tilde{L}}^{f}(l)}{1 - \sum_{l=0}^{\tau} g_{\tilde{L}}^{f}(l)} \tag{3.4}$$

, where $\mathscr{A}$, $\mathscr{B}$ and $\mathscr{C}$ represent three disjoint events. $\mathscr{A}$ is that a patent application filed in year t would be approved as of year T, $\mathscr{B}$ is that it would be approved after year T, and $\mathscr{C}$ is that it would never be approved eventually. Thus by Bayes' rule, formula 3.3 represents the conditional probability that any patent application that hasn't been approved in τ years will be approved eventually. This probability is estimated by formula 3.4, where $g^f(l)$ is the number of patents in technological field f that are granted l years after the application was filed. The estimation assumes different application-grant lag distribution for patents in different technological field, a phenomenon reported in Hall et al. (2005). Any patent that shows up in the patent database must have been approved prior to T, then by this formula, it has a discount factor that equals 1.

With the discount factors applied to applications, I am able to compute the unbiased supplemented version of the two raw measures.

The supplemented measure for patent count is defined as the total number of patents and applications a firm j filed in year t, with the applications discounted:

$$PA_{j,t} \equiv \sum_{p \in \mathbb{P}_{j,t}} 1 + \sum_{a \in \mathbb{A}_{j,t}} \beta_{\mathbf{F}(a),t} \tag{3.5}$$

$$= \underbrace{P_{j,t}}_{\substack{\text{raw}\\ \text{patent}\\ \text{count}}} + \underbrace{\sum_{a \in \mathbb{A}_{j,t}} \beta_{\mathbf{F}(a),t}}_{\substack{\text{discounted}\\ \text{application}\\ \text{count}}} \tag{3.6}$$

,where $\mathbf{F}(a)$ is the technological field to which the application a belongs [7]. $\mathbb{A}_{j,t}$ is the set of applications filed by firm j in year t but have not been approved as of T. Each application is not counted as one like the patents, but is discounted by the probability ($\beta_{\mathbf{F}(a),t}$) that it will eventually turn into a patent, given that it hasn't made it as of T. As a result, it can be shown that this measure converges to the benchmark. In contrast, the raw measures would underestimate the benchmark for recent patents because it is composed only of the first term in formula 3.6. The undiscounted supplemented measure, on the other hand, would overestimate it because it does not apply the discount factors that are smaller than one. This convergence translates into the accuracy of the supplemented measure that has been illustrated in Figure C.3.

The "supplemented average citation count measure" is defined as the average number of citations made to each patent or application that a firm filed at year t, normalized by the mean citations to other patents or applications in the same cohort:

[7]In this paper, upright bold letters represent operators that take a patent or an application as the argument, and generate a certain characteristic of that patent or application.

$$\bar{C}_{j,t} \equiv \frac{1}{PA_{j,t}} \left[\underbrace{\sum_{p\in\mathbb{P}_{j,t}} \frac{C_p^{P,P}}{\bar{C}_{\mathbf{F}(p),t}^{P,P}}}_{\equiv \bar{C}_{j,t}^{P,P};\quad A} + \underbrace{\sum_{p\in\mathbb{P}_{j,t}} \frac{\beta_{\mathbf{F}(c),\mathbf{T}(c)} C_p^{P,A}}{\bar{C}_{\mathbf{F}(p),t}^{P,A}}}_{B} + \underbrace{\sum_{a\in\mathbb{A}_{j,t}} \beta_{\mathbf{F}(a),t} \frac{C_a^{A,P}}{\bar{C}_{\mathbf{F}(a),t}^{A,P}}}_{C} \right.$$
$$\left. + \underbrace{\sum_{a\in\mathbb{A}_{j,t}} \beta_{\mathbf{F}(a),t} \frac{\beta_{\mathbf{F}(c),\mathbf{T}(c)} C_a^{A,A}}{\bar{C}_{\mathbf{F}(a),t}^{A,A}}}_{D} \right] \tag{3.7}$$

Similar to the supplemented measure for patent count, this measure can be shown to converge with the “long waited measure for patent count”. In contrast, the raw measures would underestimate the “long waited measure” for recent patents, because it is composed only of the term A in formula 3.7. The terms B, C, and D correspond to the other three sources of citations in Figure C.2 that are not utilized in the raw measure. The undiscounted supplemented measure, on the other hand, would overestimate it because it doesn’t apply the discount factors that are smaller than one. This convergence translates into the accuracy of the supplemented measure illustrated in Figure C.3.

4.2.2.2 Measures to differentiate exploration and exploitation

In the innovation literature, exploration usually refers to the endeavor of exploring an area new to the *innovator*. I identify the intellectual heritage of a patents by the citation it gives to previously granted patents, as patent citations have been shown to represent knowledge flow (Jaffe et al. (2000)). Just like academic papers have to cite previous papers that built the base of the field, patents give citations to previous patents that laid the foundation for the technological field. In my data, for each patent, we can identify which previous patents it has cited. In addition, we can identify the technological fields of each of the cited patents, as well as the patent in question. As a result, for any patent, I am able to compute its “Exploitation

Index" (EI)[8] as how many of its citations are given to existing patents that are in the same technological field as that of itself. The less a patent cites previous work of the same field, the more it is exploring innovative areas. For example, the first patent ever filed for electric cars probably cited a large number of patents in the field of electric generator while the patent in question is in the field of motor vehicle itself. This patent would have a low exploitation index by our definition above. We then compute the average of the EI of all patents a firm files in a certain year to represent how (not) explorative that firm is in that year.

While the EI measures the intellectual *proximity* of a patent to its predecessors, another important aspect of exploration is how many different fields the firm in question have synthesized to produce a certain patent. Highly innovative patents are usually build on previous works of more than one field. As a result, the more aggressive a company is exploring, the more diverse it patents citations are. Therefore the second measure for exploration is the diversity of technological fields to which a patent gives citations in year t [9]:

$$Herf_j = 1 - \sum_{f=1}^{7} \theta_{f,j}^2 \tag{3.8}$$

where $f = 1, 2, ..., 7$ is the seven-category division of technological fields proposed by Hall et al. (2001). This categorization has been shown in the literature to capture the bulk of cross field variations in patenting activities. Since for each patent j, I can identify which previous patents it has cited as well as the technological fields of each of the cited patents, I am able to measure the technological diversity of the patents that patent j, represented by $\theta_{f,j}$. The value of this measure increases as the fields a firm engages in get more diversified. Similar to the treatment to EI discussed above, I then compute the average of the $Herf_j$ of all patents a firm files in a certain year to represent how (not) explorative that firm is in that year.

[8]This measure was proposed by Trajtenberg et al. (1997).
[9]This measure was also proposed by Trajtenberg et al. (1997).

4.2.3 Voting data

Implemented in the spirit of the Dodd-Frank Act, on Jan. 25th., 2011, the SEC issued a new rule requiring that all reporting companies have their shareholders vote on the *frequency* of the "say on pay" (SOP) meetings to be held in the future. The outcome of this vote determines how frequent the Say On Pay votes are to be held in the subsequent years. The first SOP vote is also required to be held in 2011, at the same time with the vote on its frequency.(See C.1 for a rough time line.) These SOP meetings would be held as a session in the annual meetings, during which the shareholders vote to decide whether to approve the executive compensation packages proposed by the management.[10] The companies are given three choices of the frequency: once per year, once every two years and once every three years.[11] The SEC also required that this voting take place in the first annual meeting after the rule change. The companies with market capitalizations smaller than 75 million USD were the only ones that were granted a later deadline, which was set in their annual meetings in 2013. Each company is also required to disclose the voting result and whether it will be implemented in its $8-K$ filings. I obtained the SOP frequency voting results from the ISS, which includes the year of the voting, the share of votes for each of the three choices and whether the voting result was implemented. I also obtained from ISS the information regarding the SOP meetings themselves, including the meeting date and the vote share for and against the compensation plan proposed by the management.

In addition, I obtained information pertaining to the company fundamentals from Compustat as well as stock prices and market index prices from CRSP.

4.3 The identification strategy

It is challenging to establish the causal relationship between the $frequency$ of corporate governance and corporate innovation. There might be underlying factors,

[10] It's important to clarify that there are two types of voting involved here, the voting to occur in the SOP meetings and the voting beforehand on the frequency of these meetings. The second type of voting provides me with the RDD identification, and the first type is involved in producing only one set of results. The first type is used in Cunat el al 2015 but not the second type. I will refer to the first type of voting as "SOP voting" and the second type as "SOP *frequency* voting".

[11] I address later how RDD can be applied to voting with three possible outcomes.

such as the innovation strategy of the company, that drive both variables. In addition, the shareholders may anticipate the innovation intensity of the firm and choose governance frequency accordingly. As mentioned in the introduction, these underlying factors, most of which relate to the propensity to innovate, are not observable, and thus, are difficult to control for.

To overcome this challenge and prove causality, this paper exploits the 2011 SEC mandated SOP frequency voting as an exogenous source of variation in the governance frequency. I apply regression discontinuity design(RDD) methodology, effectively comparing the innovation outcomes of the firms whose shareholders voted for "once per year" with a small margin over 50% to those that are marginally in favor of "once every three years." As long as these two groups of firms are identical in all other aspects, we can attribute all later differences in innovation to the different SOP frequencies adopted by the firms. This identifying assumption is subject to rigorous falsification tests and found no evidence of it being violated, as is demonstrated later in the paper. In particular, necessary for this methodology is the assumption that the firms on the left and right side tight neighborhood around the cutoff are identical in important firm characteristics and *ex ante* innovation variables. This assumption is tested by applying the RDD to the firm characteristics and the *ex ante* innovation variables. For this and other reasons listed below, the regulatory mandate exploited provides a unique opportunity for clean identification.

4.3.1 No manipulation

The SOP frequency vote is rarely manipulated by management. If the management was able to influence votes cast by the shareholders, the "everything else equal" assumption would be violated. This is because the firms that narrowly voted for "once every three years" would have been more likely to possess characteristics that make this voting outcome appealing to the management so that management would try hard to sway the votes to obtain this result.[12] Empirical testing shows that

[12]Imbens and Lemieux (2008) shows that the RDD is valid for identification as long as the management cannot manipulate the outcome *perfectly*. But in that case the interpretation of the causal effect would be less straightforward and its estimate less accurate. The SOP voting, which has no manipulation at all, provides cleaner identification.

this concern is not valid. Had manipulation occurred, there would be a discontinuity at the 50% cutoff in the vote share distribution. Figure C.5 and Figure C.6 suggest that this is not the case. [13]

To evaluate the probability of manipulation rigorously, I applied the test of discontinuity in distribution developed by McCrary (2008). The point estimate of the discontinuity is 0.1109 with a standard error of 0.1556, thereby making it insignificant. Figure C.6 plots the probability density estimated by this test and the discontinuity is shown to be almost nonexistent.

The absence of manipulation observed here is very rare in management initiated shareholder voting. In the US, most voting is manipulated by the management, as reported in Yermack (2010). The absence of manipulation in the SOP frequency voting is probably due to the pressure for prudent governance in the immediate aftermath of the 2008 financial crisis.

Theoretically, the absence of manipulation alone guarantees the satisfaction of the "everything else the same" assumption in a large sample asymptotically Imbens and Lemieux (2008). I nonetheless conducted falsification tests, known as the "balance checks"(proposed by Roberts and Whited (2013)) to try to find potential violations of this assumption. I did not find any. These in combination are strong evidence that there exists little if any selection bias. The results of the balance checks are presented in the results section.

4.3.2 Near 100% implementation

The "cause" in the causal relationship of interest is the *implemented* frequency of the SOP rather than the mere voting frequency. If the voting result was not always implemented, I would have needed to adjust the directly estimated effect size by the implementation rate, introducing noise. But strikingly, the implementation of the SOP frequency voting results are almost 100%, allowing for an intuitive and precise *sharp* RDD. This feature of near perfect implementation is illustrated by figure C.7.

[13] The vote share in this figure appears asymmetric, representing the fact that in most firms, the shareholders voted for "once per year". This, however. does not affect the validity of the RDD in any way. This asymmetry is probably a result of the elevated consciousness of governance after the financial crisis.

A near perfect implementation is unique to this particular voting occurrence, as previous work has found that the management only implements a fraction of these non-binding voting results(Yermack (2010)). This is again probably due to the stringent application of corporate governance in the wake of the 2008 crisis.

4.3.3 Exogeneity of the participation and voting time

As explained earlier, all the firms that report to the SEC are required to hold this voting procedure, and none of them have the luxury of doing so on their own timeline. All the “big” firms were required to vote in their annual meeting in 2011, while all “small” firms had to vote in their 2013 meetings, with the criteria for each category (“big or small”)set arbitrarily by the SEC. Had the participation been voluntary, the causal effect measured would have been the “local average” for the firms motivated enough to hold such a voting, decreasing the external validity of the identification. If the companies had had the freedom to choose the time of the voting, the problem would have been even more complicated because it is well known that on the aggregate level, innovations come in waves over time (Lerner and Seru (2015)). Thus, it is possible that the companies might time the SOP frequency voting to accommodate these innovation waves. Trying to overcome this problem by adding time fixed effects would have biased the effect size estimation as the byproduct by abandoning meaningful time series variations as pointed out by Hall et al. (2005).

Fortunately, this paper does not suffer from these concerns. Due to the mandated participation and timing of this voting, all relevant variations can be controlled for. This exogenous participation and timing is rare for shareholder voting, because most shareholder voting is called not randomly, usually if and when needed.

4.3.4 SOP frequency representing governance frequency in general

“Say on pay” meetings are about more than just “pay”. They have been found by previous papers to exert profound influence on corporate governance in general. These SOP sessions are part of the shareholders’ annual meetings, during which other important issues are communicated and discussed. Typically, the SOP session is the last item on the agenda. The timing of the vote makes the SOP session a powerful tool of the shareholders when negotiating issues beyond executive compensation. After listening to reports from management about past performance and future plans, the shareholders can use the SOP vote to punish or reward the management for all issues discussed in the annual meeting. Furthermore, what is at stake for the management is arguably what they care the most: their compensation. It must be noted that the result of the vote is not binding[14], so the management can adopt pay scheme that deviates from the shareholders' vote. But Cuñat et al. (2012) shows that they nonetheless "get the message" and would modify behaviors shareholders signal disapproval of. Ultimately, the frequency of the SOP vote substantially represents the frequency of the governance actions that the shareholders get to take on the management.

4.3.5 RDD applicable to voting with three possible outcomes

RDD is usually applied to votes with only two choices on the ballot. In the case of SOP frequency vote the shareholders have three choices: once per year, once every two years and once every three years. In the paper, I argue that the identification assumptions carry over to the majority vote cases with three possible outcomes, as long as the one of the voting result is not used in the analysis, and the vote shares used are the ones for the remaining two choices. Loss of observations presents the only downside to this method, but in this case less than 5% of the companies voted in favor of "once every two years" so abandoning this outcome is not a significant loss. The other two options typically receive most of the votes in my sample, an outcome common to the plural majority votes.

4.3.6 External validity

RDD is known to have very strong internal validity but relatively weak external validity. However, I argue that this does not apply to the RDD with vote share as the running variable.

RDD would have had limited external validity had the running variable be a fundamental characteristic of the firm. For example, if the treatment is applied only

[14] The SOP frequency voting are effectively binding, as explained above, but not necessarily the SOP voting.

to firms larger than a size threshold, then the RDD estimate is only locally valid for the companies with sizes very close to the threshold. Vote share, however, is not a fundamental feature of the company, so the firms close by the vote share threshold may be anywhere in the spectrum of firm characteristics. These factors lessen our concern over the external validity issue of this identification strategy.

4.4 Results

In this section, I first introduce the specification of the regression discontinuity design (RDD) methodology that is used in generating all following results. Then I conduct the balance checks which aim to further validate the internal validity of the identification. Furthermore, I report the first set of results concerning innovation outcome in terms of patent quantity and quality. Finally as an attempt to elucidate the underlying mechanism, I report the results on the exploration - exploitation dynamics.

4.4.1 Regression discontinuity design: specification

Throughout the analyses, I use regression discontinuity design (RDD) to estimate the following varying weight regression, based on the methodology developed by Imbens and Kalyanaraman (2012):

$$Y_j = \alpha + \beta I3yr_j + P_l(v, c) + P_r(v, c) + \gamma I2011_j + \epsilon_j \tag{3.9}$$

Aside from the variables already defined in the previous subsection, $P_l(v, c)$ is a flexible polynomial function for observations on the left-hand side of the threshold c; $Pr(v, c)$ is a flexible polynomial function for observations on the right-hand side of the threshold c with different polynomial orders; v is the total vote share (percentage of votes in favor). This equation is estimated with more weight put on the observations closer to the cutoff. This ensures that the analysis relies mostly on the observations near the continuity, which leads to a *ceteris paribus* comparison. Following the convention of the literature, I report throughout the results section the polynomial specification of the fourth order, but the results are not substantially different when other orders are applied.

Indicator variable $I2011_j$ controls the difference in timing of when the SOP vote takes place. Self selection of the participation time can result in endogeneity that is difficult to control because the innovation comes in waves and the self selection may cater to this time trend. This concern is not necessary in this setup, because the SEC mandates that companies with market capitalization larger than 75 million dollars hold the SOP frequency vote in their annual meetings in 2011 while those smaller than 75 million vote in 2013. This arbitrary assignment of treatment time makes the control simple because the variable that needs control is well defined and observable. Aside from this covariate representing heterogenous in the vote mechanism itself, the RDD identification assumptions validated in section 4.3 ensure that the firms whose shareholders vote in favor of "once every three years" by a small margin of votes is otherwise identical (on average) to firms whose shareholders vote in favor of "once per year" by a small margin. The only difference between these two groups of firms is that the undistracted firms would have the SOP once every three years while the second group would have it every year. Thus, all differences in innovative outcomes in subsequent years can be attributed to their different governance frequency.

4.4.2 Balance check

Theoretically, the absence of manipulation (shown in Section 4.3) alone guarantees the satisfaction of the "everything else the same" assumption in a large sample asymptotically Imbens and Lemieux (2008). I nonetheless conducted falsification tests, known as the "balance checks"(Roberts and Whited (2013)) to try to find potential violations of this assumption. I did not find any. The balance checks are conducted via the comparison important *ex ante*(at year 2011) firm characteristics[15] between the treatment and control groups to provide additional validation, as reported in Table C.3 and Figure C.8[16]. Should there be any selection bias, that is, should the firms in the control and treatment groups differ in aspects that make their innovation propensity different, the comparison would have yielded significant

[15]These firm characteristics are computed in the same way as Chemmanur and Tian (2018).

[16]Note that the balance checks are conducted via the RDD, just like the main results, as the RDD are methodologies to compare two groups at their border.

differences of the variables tested. Yet the discontinuity at cutoff is not significant for all firm characteristics. This set of results show that firm characteristics that have been shown to affect innovation do not differ significantly across the two groups. Similar balance checks can be done on the outcome variables of this paper: I further and show that the various innovation measures of interest do not differ significantly across the treatment and the control *before*[17] the treatment took place. (Table C.4 and Figure C.9) Taken together, this is strong evidence that the treatment and the control groups differ only in whether the treatment is received.

4.4.3 Innovation quantity and quality

I find that compared with the control group, firms that hold say on pay (SOP) votes once every three years generate similar amount of patents in the first year but patents with greater quantity and quality in the third year.

The quantity of the innovation output measured by the patent count[18] is the most salient output of the innovation activity. Table C.5 reports the aforementioned comparison in patent quantity. I find that undistracted firms produce similar number of patents compared with the distracted companies in the first year after the SOP frequency vote. The first installment of the SOP is required by SEC to be held immediately after the SOP frequency vote, therefore this year is also the first year in the three-year "undistracted period" for the firms that have SOPs once every three years. In the third year, however, the difference emerges. The right column of Table C.5 shows that in the third year, the undistracted firms produce 1.333 more patents than the control group. Furthermore, this difference is robust with alternative orders of polynomials in equation 3.9. This difference can also be seen in the rightmost panel of Figure C.10. The sharp discontinuity at the vote cut-off represents the effect size of the low frequency of the SOP. The fact that the fitted line breaks out of the confidence interval of the boundary bin average of the opposite side is the demonstration that the continuity is significant.

[17] 2010 values are taken for these variables.

[18] As discussed in the data section, in my sample there are not only granted patents but also pre-grant patent applications. In this section I use the word "patent" to represent the combination of both.

The quality of a patent is measured based on the citations[19] from other patents that it would receive in the future. It has been shown by previous studies (such as Hall et al. (2005) and Lerner and Seru (2015)) that citations of a patent not indicate its scientific impact, but also correlate strongly with its private economic value to the firm. It has also been found that the majority of patents submitted for any year are worthless, while the remaining few have high values (Nordhaus (1989)). As a result, citation counts have been used in a complimentary fashion to patent counts in the study of innovation.

Table C.6 and Figure C.11 report the comparison in patent citation count. Undistracted companies produce patents with similar quality compared with the distracted companies in the first year after the SOP frequency vote. In the third year, however, the difference emerges.

This drastic time varying effect also demonstrates the importance of a clean identification of the treatment time in innovation research. Since the firms can change their output of innovation in matters of two years, any measurement error in the treatment time that is larger than a year effectively produces a moving average of the time varying effects with opposite signs. This may explain why the existent studies find contradictory results when quantifying this relationship.

4.4.4 Exploration - exploitation dynamics

How are the undistracted firms able to produce more patents if given more time? Why are they unable to do so in the first year of the cycle? Theories provided by Manso (2011) and Kerr et al. (2014) predict that the results in subsection 4.4.3 can be explained by the roles of exploration and exploitation, two integral stages necessary to generate innovations with great value. “Exploration” surveys all possible solutions to a problem or all possible areas to do R&D in. The information collected in the exploration stage guides the innovator’s direction in the “exploration” stage. They argue that the exploration stage is especially necessary for fundamental innovations that venture into unchartered territories. But exploration is expensive

[19] As discussed in the data section, I designed a scheme to discount citations from a pre-grant application. In this section by the word “citation” I refer to the citation number corrected by this scheme.

and its progress is hard to measure in the short term. These features may lead to friction in corporate governance, especially when the frequency of the governance is too high to appreciate the long term benefit of exploration. Therefore a possible explanation of the previous results suggests that the undistracted firms do not have short term performance pressure and are thus able to participate in exploration as the first stage. The short term drop in patent value may reflect the time lag of the benefit of exploration. To validate this mechanism, I search for evidence of more explorative activities from the undistracted firms in the first year of the undistracted period.

The results suggest that the undistracted firms are more explorative in the first year of the undistracted period compared with the distracted firms. They file for patents that cite significantly less of the work of the same field (Figure C.12 and Table C.7).

Using the second measure of exploration, we confirm the above finding (Figure C.13 and Table C.8). Firms with lower monitoring frequency cites a more diverse spectrum of prior patents in the first year of the 3-year cycle, indicating a more interdisciplinary type of exploration. Interestingly, in the second and third year, these low-frequency firms start to do less exploration and more exploitation, evidenced by the higher exploitation index. This suggests that these firms enter into the exploitation stage when they cash in on the information collected in the previous stage. The breadth of technology that the undistracted firms explore then drops to the same level as the control firm suggesting that the undistracted firms regained focus.

The citation diversity Herfindahl index is taken the time series first difference to further control for the firm specific inherent difference in terms of exploratitiveness that is not resulted from the treatment. The findings confirm the preceding findings.

4.5 Robustness tests

In this subsection, I subject the above results to a battery of robustness tests, varying the specifications of RDD. These alternative specifications yield similar results as reported above. For example, instead of using the optimally solved bandwidth for the RDD, I applied fixed bandwidth and obtained qualitatively similar results (Table

C.9). There has been debates which type of bandwidth setup is more internally valid and I did the above test to show that our results are not flukes given rise by the type I errors of the dynamic generation of bandwidth. In addition, as a falsification test, I used alternative cutoffs 30% of the vote share as a dummy parameter to the RDD and did not find any systematically significant result. 50% vote share is what is necessary to pass a vote, whereas 30% does not have such consequence. Thus should we observe that the RDD results are significant even if we apply 30%, then we know that the identification itself is problematic and the previous results must have been spurious. But this is not the case based on the results. (Table C.10) Furthermore, all our results are robust with alternative orders of polynomials in equation 3.9.

4.6 Limitations

Aside from the limited external validity discussed in 4.3, a problem common to all RDDs, the main limitation of this paper is that the setup does not allow us to investigate the long term effect (longer than 3 years) of the frequency shock. The reason is that this SEC policy is applicable on a continual basis. In 2014, 3 years after the determination of the frequency, the clock has been reset and restarted for another 3-year cycle. For researchers who want to look into the 4th year and 5th year impact of the 2011 event may be tempted to analyze the innovation outcome of the years 2014 and 2015, respectively. But the problem is that these years are not only subject to the influence of the 3-year cycle that have finished, but also that of the second 3-year cycle that started in 2014 and is ongoing. To put it another way, the year 2014 is not only the 4th year after the start of the first 3-year cycle, but also the 1st year of the second 3-year cycle. So any effect found in those years will be a combination of two sources of causal relationships that cannot be teased apart from each other.

Another aspect that future work can make improvement in is adding two more firm characteristics to the balance checks: institutional ownership and shareholder information asymmetry proxied by analyst forecast dispersion. Both factors have been shown to affect innovation (Aghion et al. (2013) and Hall and Lerner (2010)) so should be included in the balance checks. The author is currently unable to do so due to data access constraints. However, it's important to note that even if these two firm characteristics differ between the treatment and the control, one can solve the problem by including those two controls in the estimation of the RDD.

4.7 Conclusion

This paper provides strong evidence that commitment to governance with long term performance evaluation period encourages exploration which leads to innovation with less economic value in the short term, but innovation with more economic value and scientific value in the long term. In addition, exploration demonstrates permanent effect of shifting the technological focus of a firm to the more promising fields in the future.

These findings suggests the importance of matching governance frequencies with the natural periodicity of the innovation activities that the firms meant to undertake. In sub-section 4.4.4 we observe shifting from exploration to exploitation in the second and third year of the three year evaluation period. Though the outcome is superior than SOPs once a year, this shift is not necessarily optimal. It is possible that the shift to exploitation is still too early. This results advocates the SEC's allowance of options of even lower frequency of the SOP. A concern might be that an option of lower frequency may open the Pandora's box for the management to manipulate the shareholders into choosing a frequency that is too low for their best interest. But the absence of manipulation in the 2011 and 2013 votes is reassuring.

The SOP frequency votes suggests a decentralized way of assigning appropriate time length for governance horizon. A one size fit all governance frequency is obviously not optimal.(Kerr and Nanda (2014)) Mandating shareholders to vote on this issue serves to utilize the information the shareholders have about the idiosyncracy of the innovation that the firm is conducting. To the extent that the shareholders can rationally determine *ex ante* the appropriate governance horizon, voting serves as a commitment device. Therefore testing whether the SOP frequency votes choose the optimal frequency for each firm can be a fruitful area of future research.

Summary

The three chapters of this PhD thesis explore three ways in which financial miscoordination can happen. In the classical economic framework pioneered by works of Adam Smith, each agent only needs to optimize her own welfare. But in the end by virtue of the invisible hand, the resources are allocated in a way that is socially optimal. In this thesis, I research on instances when the invisible hand fails to motivate the players to act optimally for the common good of the society.

The thesis first explore in the context of trading speed competition. In the trading platforms, there are not only mutually beneficial strategic complementarities between a maker and a taker, but there are also zero-sum strategic substitutions between any two makers or any two takers. Thus an arms-race like speed competition is a likely outcome amongst the makers or amongst the takers. In a theoretical model, I demonstrate that whether the cost outweighs the benefit depends crucially on how the marginal contribution of trading speed to the gains from trade changes with trading speed itself. I show that when it is decreasing, the market discipline by the technological expenses borne by the traders themselves is not sufficient to curb the speed competition – the traders would over-invest in speed and exhibit "arms-race behavior". In the end of this chapter, I provide a micro-foundation for this decreasing gain from trade of trading speed, based on the classical Merton portfolio re-balancing problem.

Then I did my investigation in the banking context. Given limited liability, equity holders have incentives to engage in "risk shifting": she would take on excessive risk so that she's the one that benefits when the realized outcome is good, but her creditors are the ones to suffer the downside losses. In the context of commercial banks, the market discipline in place to address this conflict of interest is by ways of short-term, demandable debts. I show theoretically that higher risk taken by the bank may discourage runs when the bank equity buffer is low. I demonstrate that this

is a result of a common fact in finance: the higher the asset risk, the higher also is the asset upside payoff. When the bank goes bankrupt, the depositors effectively becomes the residual claimant of the asset at its maturity. Thus a higher asset risk means the asset would pay more when it succeeds, which benefits the depositors if they do not run and hold their deposits until maturity. This theory explains well the risk management practice of counting high payoff assets as substitutes for capital amongst commercial banks.

In the last part of the thesis, I move to a corporate governance context and set out to empirically test a long held question: do frequent evaluation of the CEO hurt or encourage firm innovation? I show that if given the opportunity to monitor the CEO too frequently, shareholders discipline may make the CEO reluctant to undertake R&D projects that are costly in the short term but valuable in the long term. I establish the causal relationship between evaluation frequency and corporate innovation by comparing the innovation outcomes of companies with different frequencies of SOP. To address the endogeneity problem, I take advantage of a special rule imposed by the SEC: the frequency of evaluation itself has to be determined for each company by shareholder votes. This allows me to conduct "regression discontinuity design" analysis to demonstrate that shareholder discipline may stifle innovation when the market is not able to accurately valuate innovations. This result speaks to the internet frenzy in late 1990s when tech companies are valuated much higher than their true value.

Samenvatting

De drie hoofstukken van deze PhD scriptie verkennen drie manieren waarin financiële mis-coördinatie kan gebeuren. In de klassieke economische framework gepionierd door het werk van Adam Smith, hoeft elke agent alleen maar haar eigen welvaart te optimaliseren. Maar door wijze van de onzichtbare hand worden alle hulpmiddelen gealloceerd op een manier dat optimaal is voor de maatschappij. In deze scriptie onderzoek ik gevallen waar de onzichtbare hand faalt om de spelers te motiveren om optimaal op te treden voor het algemeen belang van de maatschappij.

Deze scriptie verkent eerst de context van competitie voor verhandel snelheid. Op verhandel platforms zijn er niet alleen wederzijdse strategische contemplariteiten tussen een maker en een nemer, maar er zijn ook nul-som strategische substituties tussen ieder twee makers en ieder twee nemers. Dus een wapenwedloop voor snelheid is een aannemelijk uitkomst tussen de makers of tussen de nemers. In een theoretische model demonstreer ik dat de situatie wanneer kosten compenseren voor de profijten cruciaal af hangt van hoe de marginale contributie van verhandelsnelheid op de verhandelwinsten verandert met verhandelsnelheid zichzelf. Ik toon aan dat als het afnemend is, dan is de market discipline door de technologische kosten gedragen door de verhandelaars zichzelf niet voldoende om competitie voor snelheid in te tomen – de verhandelaars zullen over-investeren in snelheid en "wapenwedloop gedrag" vertonen.

Daarna deed ik mijn onderzoek in de context van banken. Vanwege beperkte aansprakelijkheid hebben aandeelhouders een drijfveer om "risico-verschuiving" te beoefenen: zij zullen excessief risico nemen om te profiteren wanneer de gerealiseerde uitkomst goed is, maar de crediteuren zijn degene die de nadelige verliezen verdragen. In de context van commerciële banken zijn er market disciplines in plaats om deze conflict van interesses te adresseren door middel van korte-termijn opeisbare schulden. Ik toon theoretisch aan dat het hoger risico nemen van de bank een bank run kan

ontmoedigen wanneer het vermogens buffer van de bank laag is. Ik demonstreer dat dit het resultaat is van een veelvoorkomend feit in financiën: hoe hoger de risico van het activum, hoe hoger is ook de voordelige vergoeding. Wanneer de bank failliet gaat, dan worden de depositohouders daadwerkelijk de overgebleven eiser van de activum op het einde van de looptijd. Dus een hoger activum risico betekent dat het activum meer uitkeert wanneer het succesvol is, wat dus gunstig is voor de depositohouders als zij niet terugtrekken en hun deposito's houden tot het einde van de looptijd. Deze theorie verklaart goed de risico management praktijk om hoog vergoedende activum te tellen als substitutie voor kapitaal onder commerciële banken.

In het laatste gedeelte van de scriptie schuif ik naar een ondernemingsbestuur context en test ik empirisch een lang gestelde vraag: lijdt of stimuleert bedrijfsinnovatie onder frequente evaluatie van de CEO? Ik laat zien dat als er gelegenheden gegeven zijn om de CEO te frequent te monitoren, dan zal aandeelhouder discipline de CEO terughoudend maken om R&D projecten te ondernemen die kostelijk zijn op de korte-termijn maar waardevol op de lange-termijn. Ik breng een causale verband tot stand tussen evaluatie frequentie en ondernemingsinnovatie door innovatie uitkomsten van bedrijven met verschillende SOP frequenties te vergelijken. Om de endogeniteit te adresseren maak ik gebruik van een speciale regeling opgelegd door de SEC: de frequentie van evaluatie zichzelf moet vastgesteld worden voor iedere bedrijf door de stemmen van de aandeelhouder. Dit laat mij toe om een "regression discontinuity design" analyse uit te voeren om te demonstreren dat aandeelhouders discipline innovatie kan onderdrukken wanneer de markt niet capabel is om nauwkeurig de innovaties te waarderen. Deze resultaat spreekt de internet waanzin aan op het einde van de jaren 90 toen de technologische bedrijven veel hoger gewaardeerd waren dan hun echte waardering.

Appendices

Appendix A

Trading speed competition: Can the arms race go too far?

A.1 Gains from Trade: Micro-Foundations

A.1.1 An Extended Setup

In our extended framework, portfolio rebalancing is the main motivation for trading. We build our analysis on the Merton (1969) style portfolio selection problem in which risk-averse takers trade due to a re-balancing need resulting from market price changes. One could alternatively look at a dynamic hedging problem for a non-linear derivative position where the taker is risk averse to hedging error and can only trade at random points in time; the intuition in such a problem would run along the same lines. In turn, the makers in our extended model are risk neutral, are rewarded a share of the gains from trade from the takers and serve to provide liquidity. Overall, this new setup extends from the one in Subsection 2.3.1, integrating the mechanism generating the need for trading.

We now assume that in an economy of infinite time horizon, there are two assets available for trade. There is a risk free asset with constantly compounded rate of r:

$$dX_0(t) = rX_0(t)dt, \tag{A.1.1}$$

and a risky asset whose price change follows a geometric Brownian motion:

$$dX_1(t) = \alpha X_1(t)dt + \sigma X_1(t)dB(t), \tag{A.1.2}$$

where $B(t)$ is a Brownian motion. Only the risky asset is traded (in a LOB).

We allow continuous quantities of limit orders in the order book to be traded on, after which the rest of the limit order is withdrawn by assumption. This way, the book fully accommodates the desire to continuously rebalance as a result of price changes. By assumption, if the LOB is filled, there is always enough volume available to accommodate any trading demand of a single trader.[1] Whenever a maker arrives to an empty LOB, she can fill it to "full", meaning that she puts in a limit buy order as well as a sell order, both of infinite size. The taker has an opportunity to trade when she arrives to the LOB and sees that the LOB is full (i.e., it has been filled by a maker and not yet been hit by another taker). Whenever she sees this opportunity, the taker will trade almost surely in equilibrium, since the price of the risky asset is continuously changing and therefore, her portfolio deviates from the optimal portfolio with probability 1.

We maintain the assumptions in Section 3.2 about the monitoring intensities of the makers and the takers. Thus for any taker j, the occurrences of the joint event that "she reaches the book *and* the book is filled at the time" follow a Poisson Point Process (PPP): $(s_1, s_2, \dots)$, henceforth referred to as this taker's "trading times". Note that a certain taker's trading time is discrete and stochastic.

Let $P(t)$ denote the Poisson process from which the aforementioned PPP is generated. Thus the differential:

$$dP(t) = \begin{cases} 1 & \text{when} \quad \exists n, \text{ such that:} \quad s_n \in [t, t + dt), \\ 0 & \text{otherwise.} \end{cases}$$

And the intensity parameter of this parameter, henceforth referred to as her "trading frequency" can be computed:

$$\lambda_j = \frac{\tau_j \bar{\mu}}{\bar{\mu} + \bar{\tau}}. \tag{A.1.3}$$

In this section, we take intensities as given, as they are chosen before trading

[1] Restricting the available volume in the LOB would only strengthen our results. To see this, with infrequent trading, the required trading size will on average be larger than with frequent trading as the volatility of realized distortions from optimal portfolio weights since the last trade is larger. As a result, the constraint on available depth becomes less and less binding, implying declining marginal GFT.

starts. Hence, we can abstract from the multitude of agents of the previous sections and focus on one of the takers, say j. For this representative taker, let $W_0(t)$ and $W_1(t)$ be the amount of her wealth invested in the risk free and the risky securities respectively at time t. $Q(t)$ is the amount of wealth that she transfers from the risky asset to the risk free asset at her trading time $t = s_n$, $n = 1, 2, \ldots$. Let $C(t)$ be the process of her consumption flow. The consumption of the taker is allowed to happen continuously.

We assume that there is a cost for each transaction paid by the taker to the maker involved. This fee equals a fixed proportion π_m of the GFT $g(P(t))$ of this trade with the trade size $P(t)$. We further assume that whenever trade execution happens, the rest of the depth (i.e., the remaining limit orders) is automatically canceled. This simplifying assumption implies that the maker responsible for these orders does not have to monitor again to cancel them.

To generate a need for portfolio rebalancing, we assume the takers to be risk averse with log utility and CRRA with relative risk aversion parameter δ.[2]

A.1.2 Endogenizing Gains from Trade

In order to endogenize the gains from trade, we first define the problem of taker j, the solution of which pins down all the quantities in this model.

We assume that each taker maximizes the expected sum of her time-discounted consumption trajectory under inter-temporal budget constraints and the no-bankruptcy constraint.

Definition 10 *Given her trading frequency λ_j, the **"taker's problem"** is each taker j's problem to choose the quantity of portfolio re-balancing $\{Q(t)\}_{t=s_n}$[3] and the continuous consumption process $\{C(t)\}_{t\geq 0}$ to solve:*

$$\max_{\{C(t),Q(t)\}_t} \mathbb{E} \int_0^\infty e^{-\rho t} U(C(t))dt \equiv J(x, y, t = 0|\lambda_j, \pi_m), \tag{A.1.4}$$

[2]Our results can be generated for any utility function with risk aversion. The log utility function happens to provide us with a high degree of tractability.

[3]Sometimes we write $\{Q(t)\}_{t=s_n}$ simply as $\{Q(t)\}_{t\geq 0}$, since the value of $Q(t)$ doesn't matter when $dP(t) = 0$.

subject to:

$$W_0(0) = x, \tag{A.1.5}$$

$$W_1(0) = y, \tag{A.1.6}$$

$$dW_0(t) = W_0(t)\frac{dX_0(t)}{X_0(t)} - C(t)dt + (Q(t) - \pi_m g(Q(t)))dP(t), \tag{A.1.7}$$

$$dW_1(t) = W_1(t)\frac{dX_1(t)}{X_1(t)} - Q(t)dP(t), \tag{A.1.8}$$

$$0 \leq W_0(t) + W_1(t), \tag{A.1.9}$$

where

$$g(Q(t)|\lambda_j, \pi_m) \equiv J(W_0(t) + Q(t), W_1(t) - Q(t), t|\lambda_j, \pi_m) - J(W_0(t), W_1(t), t|\lambda_j, \pi_m). \tag{A.1.10}$$

Equations (A.1.5) and (A.1.6) are initial conditions. Equations (A.1.7) and (A.1.8) are the laws of motions for the risk free security and the risky security, respectively. Finally, inequality (A.1.9) is the no-bankruptcy condition for taker j.

The term $-Q(t)dP(t)$ in equation (A.1.8) and the term $Q(t)dP(t)$ in equation (A.1.7) represent that wealth of amount $Q(t)$[4] is transferred from the risky security to the risk free security at time $t = s_n$, $n = 1, 2, \ldots$.

In equation A.1.10, the gain from trade function $g(Q(t))$ is defined as the value function (defined in (A.1.4)) when this trade is carried out at quantity $Q(t)$, subtracted by the value function when this trade is not carried out. Correspondingly, term $\pi_m g(Q(t))dP(t)$ in equation (A.1.7) represents that this transaction $Q(t)$ costs the taker $\pi_m g(Q(t)))$, which is a proportion π_m of the gain from this trade: $g(Q(t)))$. As we explain later, this assumption of the transaction cost being proportional to the gain from a trade is crucial for the tractability of our model.

Next, we define the "expected marginal GFT", the variable that we endogenize in this section.

[4] $Q(t)$ is allowed to be negative in which case the transfer is the other way around.

Definition 11 *The* ***"expected marginal GFT"****, is defined as:*

$$G(\lambda_j, \pi_m) \equiv \mathbb{E}_{t=s_n} g(Q^*(W_0(s_n), W_1(s_n)), s_n | \lambda_j, \pi_m), \tag{A.1.11}$$

where $Q^*(W_0(s_n), W_1(s_n))$ is the second argument of the solution to the problem in Definition 10.

By the properties of the conditional expectations, for a given taker j, every one of her trade has the identical *ex ante* expected gain. It equals the expectation of the gain from trade function $g(Q(t))$ (defined in equation (A.1.10)), evaluated at the optimal trade size: $Q^*(W_0(s_n), W_1(s_n), t)$.

The goal of this section is to identify, for any taker j, the monotonicity of her expected GFT for a trade: $G(\lambda_j, \pi_m)$ with respect to her trading frequency λ_j.

A.1.3 Asymptotically Diminishing Expected GFT for a Trade

As shown by the review of Pham (2009), a closed-form solution is not currently possible for the type of stochastic control problem as in Definition 10.[5] But we are able to bypass the need for an analytic solution of this problem, and instead identify the monotonicity of the expected GFT for a trade directly.

To achieve this, we first simplify the problem by showing that its solution is to equivalent to that of a similar problem *without* transaction costs.

Proposition 7 ***(sufficient condition of the monotonicity of the "expected marginal GFT")*** *To show that $G(\lambda_j, \pi_m)$ defined in Definition 11 decreases in λ_j, it suffices to show that the following quantity is concave in λ_j:*

$$J(W_0(0), W_1(0), 0 | \lambda_j, \pi_m = 0).$$

Proof. See appendix. □

Simply put, this quantity is the value function of problem 10 when there is no

[5]Our problem 10 differs from the classic portfolio re-balancing models, such as Merton (1969), in two important ways. First, the trade can only occur at discrete random time points; second, there is transaction costs when the trade happens.In this aspect, our model is therefore related to the models of dynamic equilibria with transaction costs, such as Constantinides (1986), but they do not have the restriction that the trades can only happen at discrete random time points.

transaction cost. Intuitively, this result crucially follows from the assumption that for any trade, the transaction cost is proportional to the gain from this trade. In this case, at any of the trading times s_n, $n = 1, 2, \ldots$, the actual objective function of the taker is simply a linear transformation of the objective function of the corresponding transaction-cost-free problem.

Based on the Proposition 7, it remains only to show that $J(W_0(0), W_1(0), 0|\lambda_j, \pi_m = 0)$ [6] is concave in λ_j. To achieve this, we first write down the Bellman Equation of the problem without transaction cost.

Equation (A.1.4) can be restated in dynamic programming form to apply the Bellman Principle of Optimality:

$$\begin{aligned}
&J[W_0(t_0), W_1(t_0), t_0|\lambda_j] = \\
&\quad \max_{C(s), W_1(s)} E_{t_0}\left[\int_{t_0}^{t} e^{-\rho s} U[C(s)]ds) + J[W_0(t), W_1(t), t|\lambda_j]|W_0(t_0), W_1(t_0)\right].
\end{aligned} \tag{A.1.12}$$

Then by Taylor's theorem and the mean value theorem for integrals, there exist $\bar{t} \in [t_0, t]$, such that the above equation can be restated as:

$$\begin{aligned}
&J[W_0(t_0), W_1(t_0), t_0|\lambda_j] = \\
&\max_{C(t), W_1(t)} E(t_0)\left[e^{-\rho \bar{t}} U[C(\bar{t})]l + J + \frac{\partial J}{\partial W_0(t_0)}(W_0(t) - W_0(t_0))\right. \\
&+ \frac{\partial J}{\partial W_1(t_0)}(W_1(t) - W_1(t_0)) + \frac{1}{2}\frac{\partial^2 J}{\partial W_1(t_0)^2}(W_1(t) - W_1(t_0))^2 \\
&\left. + \lambda_j [J(W_0(t_0) + Q(t_0), W_1(t_0) - Q(t_0), t_0|\lambda_j) - J(W_0(t_0), W_1(t_0))] + o(l)\right].
\end{aligned} \tag{A.1.13}$$

where $l \equiv t - t_0$ is the latency at time t_0. The last term is due to the possibility of the jump of the portfolio processes.

We now take the $\mathbb{E}_{t_0}$ operator onto each term, and eliminate $J[W_0(t_0), W_1(t_0), t_0|\lambda_j] = \mathbb{E}_{t_0} J[W_0(t_0), W_1(t_0), t_0|\lambda_j]$ from both sides, then evaluate $\mathbb{E}_{t_0}[W_0(t) - W_0(t_0), t_0|\lambda_j]$,

[6] Henceforth written as: $J(W_0(0), W_1(0), 0|\lambda_j)$

$\mathbb{E}_{t_0}[W_1(t) - W_1(t_0)]$ and $\mathbb{E}_{t_0}[W_1(t) - W_1(t_0)]^2$, and finally take the limit as $l \to 0$ we get:

$$\begin{aligned} 0 = \max_{C(t), W_1(t)} \Big[& U[C(t)] - \rho J[W_0(t), W_1(t), t_0|\lambda_j] + \frac{\partial J}{\partial W_0(t)}(rW_0(t) - C(t)) \\ & + \frac{\partial J}{\partial W_1(t)} \alpha W_1(t) + \frac{1}{2}\sigma^2 W_1^2(t) \frac{\partial^2 J}{\partial W_1^2(t)} \\ & + \lambda_j [J(W_0(t) + Q(t), W_1(t) - Q(t), t_0|\lambda_j) - I(W_0(t), W_1(t))] \Big] \end{aligned} \quad \text{(A.1.14)}$$

Using the technique of asymptotic expansion of the solution as in Roger and Zane (2002), we are able to show the asymptotic concavity of the expected gain from a trade in the trading speed, as $l \to 0$.

Proposition 8 ***(expected aggregate GFT without transaction cost)*** *For each taker, her value function $J[W_0(0), W_1(0), 0|\lambda_j]$ is a concave function of λ_j to the 2nd order of the inverse of the trading frequency λ_j, [7], that is, as the trading frequency λ_j is high enough that $\frac{1}{\lambda_j^2}$ is negligible.*

Proof. See appendix. □

Our reasoning in the proof is based on asymptotic expansion. The very high speed of the HFTs warrants our assumptions of λ_j getting very large in the current trading environment.

The concavity comes from the nature of the Geometric Brownian price change in combination with risk aversion. With a Geometric Brownian Motion, the variance of returns is proportional to time elapsed. The expected MGFT is proportional to the expected variance reduction due to trading, and hence by transitivity to time elapsed. Since the average time span between trades declines with transaction intensity, expected MGFT are decreasing in transaction intensity.

A.2 Figures

[7] While fixing the main arguments W_0, W_1.

Figure A.1. The figures present the regions for under- and over-monitoring for a number of examples. We set the marginal cost coefficient of the makers equals that of the takers: $\beta = \gamma$ and normalize the expected marginal GFT $G(\cdot) = G_0 = 1$. The graphs display plots of over-monitoring, that is the equilibrium intensity minus the first best intensity, of either side (row 1 for the maker intensity and row 2 for taker intensity), as a function of the maker's share of the total gains from trade: π_m and the marginal cost coefficient $\beta = \gamma$. The flat surface is the 0 plane. The region above this plane corresponds to over-monitoring, while that under under it corresponds to under-monitoring. Each column of the figure is for one combination of (M, N).

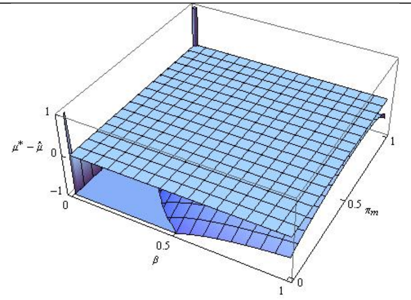
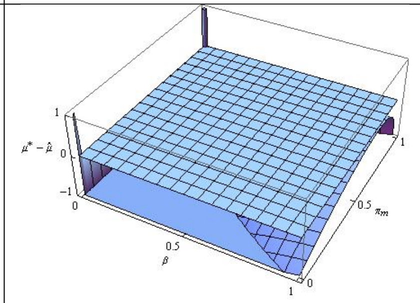
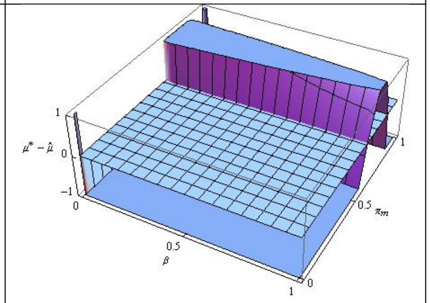
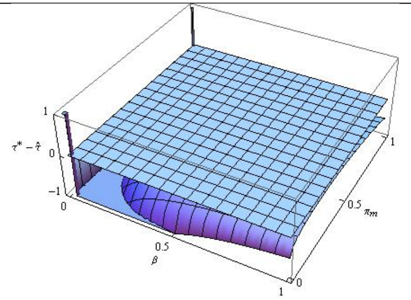
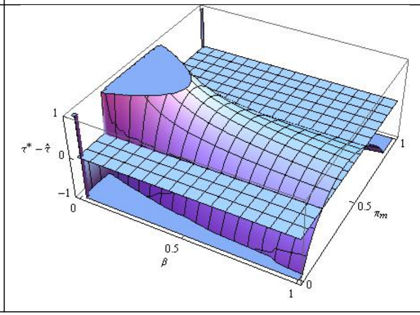
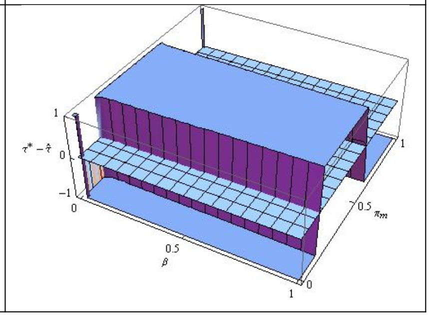
side of market	combinations of the numbers of the makers (M) and that of the takers (N)		
	$M = 2,\ N = 2$	$M = 2,\ N = 20$	$M = 50,\ N = 250$
Maker side:			
Taker side:			

Figure A.2. The Figure is equivalent to Figure A.1, but with $G(R) = -R + 1$.

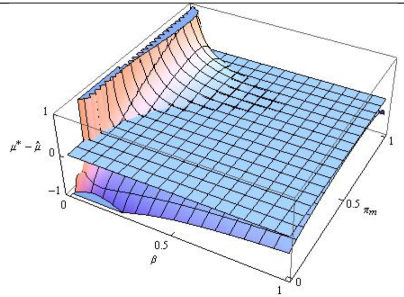
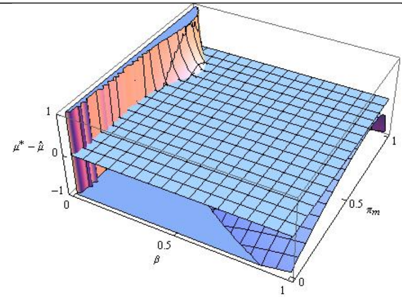
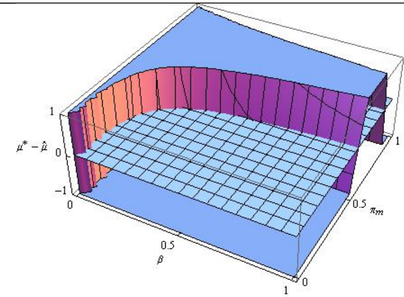
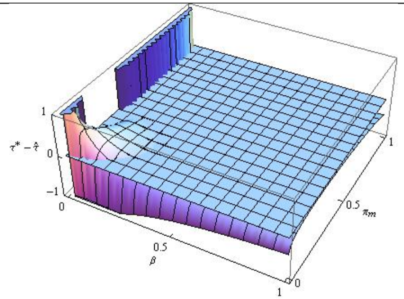
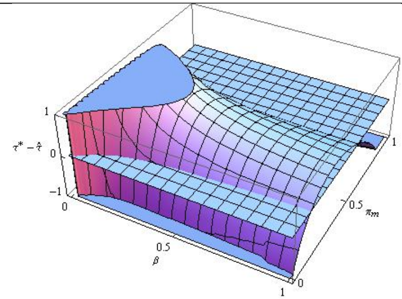
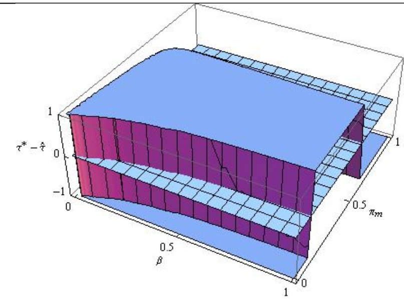
side of market	combinations of the numbers of the makers (M) and that of the takers (N)		
	$M = 2,\ N = 2$	$M = 2,\ N = 20$	$M = 50,\ N = 250$
Maker side:			
Taker side:			

Figure A.3. The Figure is equivalent to Figure A.1, but with $G(R) = -0.2R + 1$.

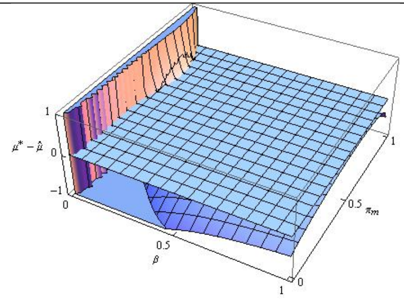
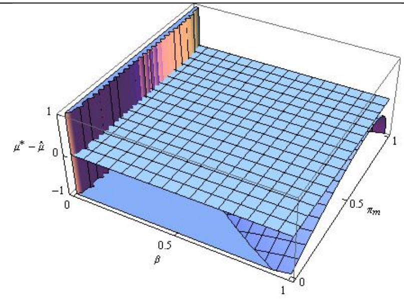
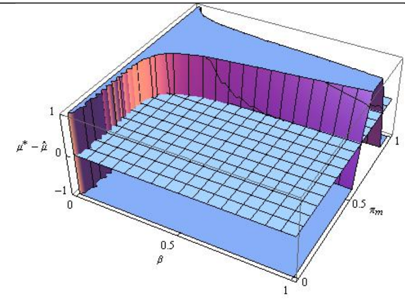
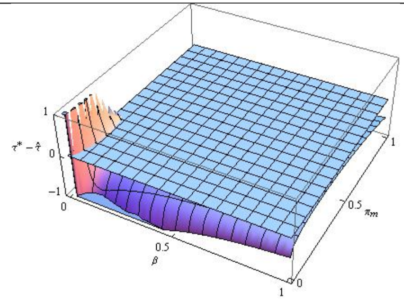
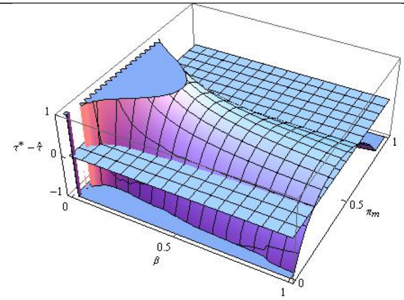
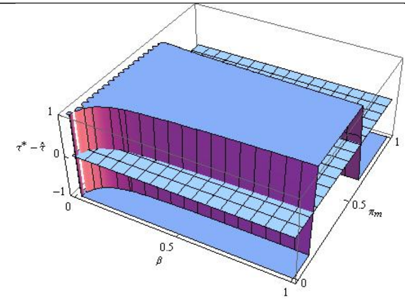
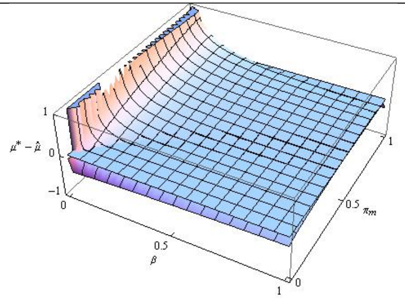
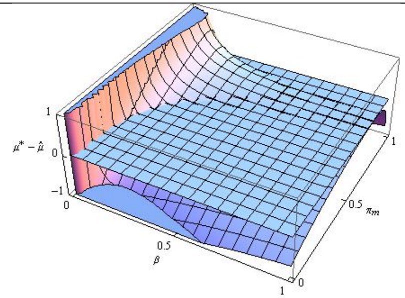
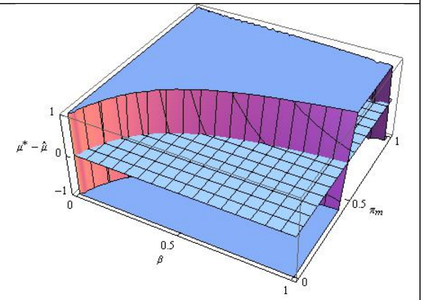
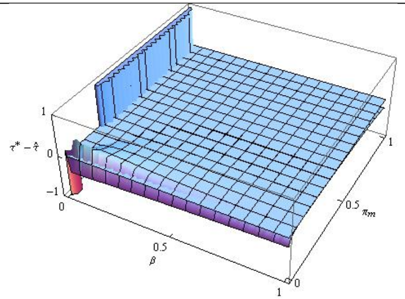
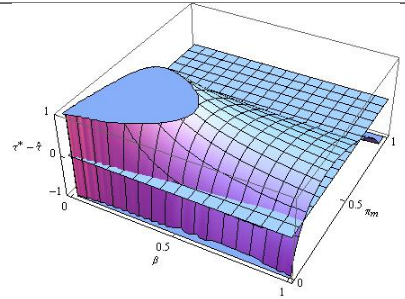
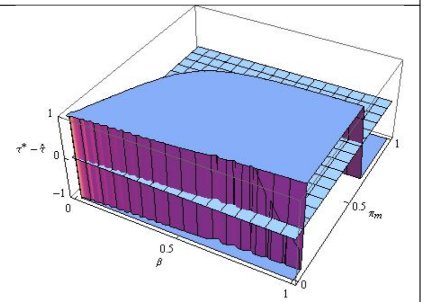
side of market	combinations of the numbers of the makers (M) and that of the takers (N)		
	$M = 2,\ N = 2$	$M = 2,\ N = 20$	$M = 50,\ N = 250$
Maker side:			
Taker side:			

Figure A.4. The Figure is equivalent to Figure A.1, but with $G(R) = -5R + 1$.

side of market	combinations of the numbers of the makers (M) and that of the takers (N)		
	$M = 2,\ N = 2$	$M = 2,\ N = 20$	$M = 50,\ N = 250$
Maker side:			
Taker side:			

A.3 Proofs

Proof. [Proof of Proposition 1] We derive some mathematical results:

$$\frac{\partial\left(\frac{\tau_j\bar{\mu}}{\bar{\mu}+\bar{\tau}}\right)}{\partial\mu_i} = \frac{\tau_j\bar{\tau}}{(\bar{\mu}+\bar{\tau})^2}, \tag{A.3.15}$$

$$\frac{\partial\left(\frac{\tau_j\bar{\mu}}{\bar{\mu}+\bar{\tau}}\right)}{\partial\tau_j} = \frac{\bar{\mu}^2+\bar{\mu}(\bar{\tau}-\tau_j)}{(\bar{\mu}+\bar{\tau})^2}, \tag{A.3.16}$$

$$\frac{\partial\left(\frac{\tau_j\bar{\mu}}{\bar{\mu}+\bar{\tau}}\right)}{\partial\tau_k} = -\frac{\tau_j\bar{\mu}}{(\bar{\mu}+\bar{\tau})^2}, \qquad \forall k \neq j. \tag{A.3.17}$$

For the SPP of Definition 1, we derive first order conditions with respect to $\mu_i, i = 1, 2, \ldots, M$ and $\tau_j, j = 1, 2, \ldots, N$. The FOC w.r.t. μ_i is given by:

$$\beta\mu_i = \sum_{j=1}^{N} G\left(\frac{\tau_j\bar{\mu}}{\bar{\mu}+\bar{\tau}}\right)\frac{\tau_j\bar{\tau}}{(\bar{\mu}+\bar{\tau})^2} + \sum_{j=1}^{N} G'\left(\frac{\tau_j\bar{\mu}}{\bar{\mu}+\bar{\tau}}\right)\frac{\tau_j^2\bar{\mu}\bar{\tau}}{(\bar{\mu}+\bar{\tau})^3}, \tag{A.3.18}$$

And the FOC w.r.t. τ_j is given by :

$$\begin{aligned}\gamma\tau_j = {} & G\left(\frac{\tau_j\bar{\mu}}{\bar{\mu}+\bar{\tau}}\right)\frac{\bar{\mu}^2+\bar{\mu}(\bar{\tau}-\tau_j)}{(\bar{\mu}+\bar{\tau})^2} + G'\left(\frac{\tau_j\bar{\mu}}{\bar{\mu}+\bar{\tau}}\right)\frac{\bar{\mu}^2+\bar{\mu}(\bar{\tau}-\tau_j)}{(\bar{\mu}+\bar{\tau})^2}\frac{\tau_j\bar{\mu}}{\bar{\mu}+\bar{\tau}} \\ & - \frac{\tau_j\bar{\mu}}{(\bar{\mu}+\bar{\tau})^2}\sum_{k\neq j} G\left(\frac{\tau_k\bar{\mu}}{\bar{\mu}+\bar{\tau}}\right) - \frac{\tau_j\bar{\mu}}{(\bar{\mu}+\bar{\tau})^2}\sum_{k\neq j} G'\left(\frac{\tau_k\bar{\mu}}{\bar{\mu}+\bar{\tau}}\right)\frac{\tau_k\bar{\mu}}{\bar{\mu}+\bar{\tau}}.\end{aligned} \tag{A.3.19}$$

Due to the symmetry of the problem, the optimal monitoring intensity for all makers should be equal, so are those of the takers. Thus we plug into (A.3.18) and (A.3.19) the symmetry conditions:

$$\begin{aligned}&\mu_i = \hat{\mu} = \frac{\bar{\mu}}{M}, \quad \text{for all } i = 1, 2, \ldots, M; \\ &\tau_j = \hat{\tau} = \frac{\bar{\tau}}{N}, \quad \text{for all } j = 1, 2, \ldots, N.\end{aligned} \tag{A.3.20}$$

The FOCs now simplify to

$$\beta\hat{\mu} = G\left(\frac{M\hat{\mu}\hat{\tau}}{M\hat{\mu}+N\hat{\tau}}\right)\frac{N^2\hat{\tau}^2}{(M\hat{\mu}+N\hat{\tau})^2} + G'\left(\frac{M\hat{\mu}\hat{\tau}}{M\hat{\mu}+\hat{\tau}}\right)\frac{M\hat{\mu}N^2\hat{\tau}^3}{(M\hat{\mu}+N\hat{\tau})^3}, \tag{A.3.21}$$

$$\gamma\hat{\tau} = G\left(\frac{\hat{\tau}M\hat{\mu}}{M\hat{\mu}+N\hat{\tau}}\right)\frac{M^2\hat{\mu}^2}{(M\hat{\mu}+N\hat{\tau})^2} + G'\left(\frac{M\hat{\mu}\hat{\tau}}{M\hat{\mu}+N\hat{\tau}}\right)\frac{M^3\hat{\mu}^3\hat{\tau}}{(M\hat{\mu}+N\hat{\tau})^3}. \tag{A.3.22}$$

Similar to Foucault et al. (2013), we divide (A.3.21) by (A.3.22) and rewrite to get:

$$\frac{\hat{\mu}}{\hat{\tau}} = \left(\frac{\gamma}{\beta}\frac{N^2}{M^2}\right)^{\frac{1}{3}}. \tag{A.3.23}$$

We now set $G(\cdot) = G_0$. This implies that $G'(\cdot) = 0$. Moreover, we slightly rewrite (A.3.23) to

$$\hat{\mu} = \left(\frac{\gamma}{\beta}\frac{N^2}{M^2}\right)^{\frac{1}{3}}\hat{\tau}. \tag{A.3.24}$$

Substituting (A.3.24) into (A.3.21) and (A.3.22) and slightly rewriting, we get:

$$\hat{\mu} = \frac{N^2}{(M\hat{r}+N)^2}\frac{G_0}{\beta}, \tag{A.3.25}$$

$$\hat{\tau} = \frac{M^2r^2}{(M\hat{r}+N)^2}\frac{G_0}{\gamma}, \tag{A.3.26}$$

where

$$\hat{r} = \left(\frac{\gamma}{\beta}\frac{N^2}{M^2}\right)^{\frac{1}{3}}. \tag{A.3.27}$$

Total welfare is obtained by substituting first best intensities into the welfare function:

$$\hat{\Pi} = \frac{M\hat{\mu}\cdot N\hat{\tau}}{M\hat{\mu}+N\hat{\tau}}G_0 - \frac{\beta M\hat{\mu}^2}{2} - \frac{\gamma N\hat{\tau}^2}{2}. \tag{A.3.28}$$

□

Proof. [Proof of Proposition 2] From Definition 2, we obtain FOCs for (individual) makers and takers. The FOC for maker I is given by

$$\beta\mu_i = \pi_m \frac{\bar{\mu}+\bar{\tau}-\mu_i}{(\bar{\mu}+\bar{\tau})^2} \sum_{j=1}^{N} G\left(\frac{\tau_j\bar{\mu}}{\bar{\mu}+\bar{\tau}}\right)\tau_j + \pi_m \frac{\mu_i\bar{\tau}}{(\bar{\mu}+\bar{\tau})^3} \sum_{j=1}^{N} G'\left(\frac{\tau_j\bar{\mu}}{\bar{\mu}+\bar{\tau}}\right)\tau_j^2, \tag{A.3.29}$$

and for taker j by

$$\gamma\tau_j = \pi_t \frac{\bar{\mu}^2+\bar{\mu}(\bar{\tau}-\tau_j)}{(\bar{\mu}+\bar{\tau})^2} \left[G\left(\frac{\tau_j\bar{\mu}}{\bar{\mu}+\bar{\tau}}\right) + G'\left(\frac{\tau_j\bar{\mu}}{\bar{\mu}+\bar{\tau}}\right)\frac{\tau_j\bar{\mu}}{\bar{\mu}+\bar{\tau}}\right]. \tag{A.3.30}$$

To solve for a symmetric equilibrium, we plug (A.3.20) into (A.3.29) and (A.3.30) and get:

$$\beta\hat{\mu} = \pi_m \frac{N^2\hat{\tau}^2+N(M-1)\hat{\mu}\hat{\tau}}{(M\hat{\mu}+N\hat{\tau})^2} G\left(\frac{\hat{\tau}M\hat{\mu}}{M\hat{\mu}+N\hat{\tau}}\right) + \pi_m \frac{\hat{\mu}N^2\hat{\tau}^3}{(M\hat{\mu}+N\hat{\tau})^3} G'\left(\frac{\hat{\tau}M\hat{\mu}}{M\hat{\mu}+N\hat{\tau}}\right), \tag{A.3.31}$$

$$\gamma\hat{\tau} = \pi_t \frac{M^2\hat{\mu}^2+M(N-1)\hat{\mu}\hat{\tau}}{(M\hat{\mu}+N\hat{\tau})^2} \left[G\left(\frac{\hat{\tau}M\hat{\mu}}{M\hat{\mu}+N\hat{\tau}}\right) + \frac{\hat{\tau}M\hat{\mu}}{M\hat{\mu}+N\hat{\tau}} G'\left(\frac{\hat{\tau}M\hat{\mu}}{M\hat{\mu}+N\hat{\tau}}\right)\right]. \tag{A.3.32}$$

We now set $G(\cdot) = G_0$. This implies that $G'(\cdot) = 0$. Substituting into (A.3.31) and (A.3.32), we get:

$$\mu^* = \frac{N(\tau^*)^2+N(M-1)\mu^*\tau^*}{(M\mu^*+N\tau^*)^2}\frac{\pi_m G_0}{\beta}, \tag{A.3.33}$$

$$\tau^* = \frac{M^2(\mu^*)^2+M(N-1)\mu^*\tau^*}{(M\mu^*+N\tau^*)^2}\frac{\pi_t G_0}{\gamma}. \tag{A.3.34}$$

Dividing (A.3.33) by (A.3.34), we get

$$\frac{\mu^*}{\tau^*} = \frac{N(\tau^*)^2+N(M-1)\mu^*\tau^*}{M^2(\mu^*)^2+M(N-1)\mu^*\tau^*}\frac{\pi_m\gamma}{\pi_t\beta}. \tag{A.3.35}$$

Rewriting yields

$$0 = M^2(\mu^*)^3 + M(N-1)(\mu^*)^2\tau^* - z(N(\tau^*)^3 + N(M-1)\mu^*(\tau^*)2), \quad \text{(A.3.36)}$$

where $z = \frac{\gamma}{\beta}\frac{\pi_m}{\pi_t}$. Now we multiplying through with $\frac{MN}{N\tau^*}$. We now have that (A.3.36) is equivalent to

$$Nr^3 + (N-1)r^2 - (M-1)zr - Mz = 0, \quad \text{(A.3.37)}$$

where $r = \frac{M\mu^*}{N\tau^*}$. Let r^* be the positive real root to (A.3.37). We must have that $M\mu^* = r^*N\tau^*$. Substituting this into (A.3.33) and (A.3.34), we get

$$\mu^* = \frac{M + (M-1)r^*}{M(1+r^*)^2}\frac{\pi_m}{\beta}, \quad \text{(A.3.38)}$$

$$\tau^* = \frac{N(r^*)^2 + (N-1)r^*}{N(1+r^*)^2}\frac{\pi_t}{\gamma}. \quad \text{(A.3.39)}$$

These solutions are similar to Foucault et al 2013 A6-A11.

EQ ”total welfare”, defined as the sum of all agents’ profit is given by:

$$\Pi^* = \frac{M\mu^* \cdot N\tau^*}{M\mu^* + N\tau^*}G_0 - \frac{\beta M(\mu^*)^2}{2} - \frac{\gamma N(\tau^*)^2}{2}. \quad \text{(A.3.40)}$$

□

Proof. [Proof of Lemma 3] The proof directly follows from substituting $G(\cdot) = G_0$ into equations (A.3.21), (A.3.22), (A.3.31), and (A.3.32). □

Proof. [Proof of Lemma 4] The proof directly follows from equations (A.3.21), (A.3.22), (A.3.31), and (A.3.32). □

Proof. [Proof of Proposition 3] Since $\mu_i, \tau_j \geq 0 \forall i, j$, we have that the terms in (2.11) to (2.14) involving G' have the same sign as G'. Keeping everything else constant, the other terms are as in the case with $G' = 0$. The term involving G' in (2.12) is in absolute terms smaller or equal than in (2.11) because $\frac{\mu_i}{\bar{\mu}} \leq 1$, leading to an under-reaction compared to first best. The term involving G' in (2.14) is in absolute terms larger or equal than in (2.13) because $\frac{\bar{\mu}^2(\bar{\mu}+\tau_{-j})}{\bar{\mu}^3} \geq 1$, leading to an over-reaction compared to first best. This term is multiplied with π_m for makers and with $1 - \pi_m$ for takers, which proves the interaction effect. $\square$

Proof. [Proof of proposition 7] We first express $J(W_0(0), W_1(0), 0|\lambda_j, \pi_m)$ recursively.

By definition,

$$
\begin{aligned}
& J(W_0(0), W_1(0), 0|\lambda_j, \pi_m) \\
= & \max_{\{C(s)\}_{0\leq s\leq s_1}} \mathbb{E}_{t=0} \int_0^{s_1} e^{-\rho s} U(C(s))ds \\
& + \mathbb{E}_{t=0} e^{-\rho s_1} \left[(1-\pi_m) g(Q^*(s_1)) + J(W_0(s_1), W_1(s_1), s_1|\lambda_j, \pi_m)\right] \\
= & \max_{\{C(s)\}_{0\leq s\leq s_1}} \mathbb{E}_{t=0} \int_0^{s_1} e^{-\rho s} U(C(s))ds \\
& + \mathbb{E}_{t=0} e^{-\rho s_1} (1-\pi_m) g(Q^*(s_1)) \\
& + \mathbb{E}_{t=0} \max_{\{C(s)\}_{s_1\leq s\leq s_2}} \mathbb{E}_{t=s_1} \int_{s_1}^{s_2} e^{-\rho s_1} e^{-\rho s} U(C(s))ds \\
& + \mathbb{E}_{t=0} \mathbb{E}_{t=s_1} e^{-\rho s_2} J(W_0(s_2), W_1(s_2), s_2|\lambda_j, \pi_m) \\
& + \mathbb{E}_{t=0} \mathbb{E}_{t=s_1} e^{-\rho s_2} (1-\pi_m) g(Q^*(s_2)) \\
& \text{(Iterate the above process by } N-1 \text{ times...;)} \\
& \xrightarrow{n\to+\infty} \\
& \sum_{n=1,2,\dots} \max_{\{C(s)\}_{s_{n-1}\leq s\leq s_n}} \mathbb{E}_{t=0} \int_{s_{n-1}}^{s_n} e^{-\rho s_{n-1}} e^{-\rho s} U(C(s))ds \\
& + 0 \cdot \mathbb{E}_{t=0} J(W_0(s_N), W_1(s_N), s_N|\lambda_j, \pi_m) \\
& + \mathbb{E}_{t=0} \sum_{n=1,2,\dots} e^{-\rho s_n} (1-\pi_m) g(Q^*(s_n)) \\
= & J(W_0(0), W_1(0), 0|0, \pi_m) \\
& + (1-\pi_m) G(\lambda_j, \pi_m) \mathbb{E}_{t=0} \sum_{n=1,2,\dots} e^{-\rho s_n} \\
& \text{(by definition: } G(\lambda_j, \pi_m) \equiv \mathbb{E}_{t=0} g(Q^*(s_n)))
\end{aligned}
$$

Evaluate the second term, we have:

$$
\begin{aligned}
&(1-\pi_m)G(\lambda_j,\pi_m)\mathbb{E}_{t=0}\sum_{n=1,2,\ldots} e^{-\rho s_n}\\
=&(1-\pi_m)G(\lambda_j,\pi_m)\sum_{n=1,2,\ldots}\mathbb{E}_{t=0}e^{-\rho s_n}\\
=&(1-\pi_m)G(\lambda_j,\pi_m)\sum_{n=1,2,\ldots}\int_0^\infty e^{-\rho s_n}\frac{\lambda_j}{\Gamma(n)}s_n^{n-1}e^{-\lambda_j s_n}ds_n\\
&\text{(We plug in the pdf of } Gamma(n,\lambda) \text{ distribution)}\\
=&(1-\pi_m)G(\lambda_j,\pi_m)\sum_{n=1,2,\ldots}\left(\frac{\lambda_j}{\rho+\lambda_j}\right)^n\\
=&\frac{1-\pi_m}{\rho}G(\lambda_j,\pi_m)\lambda_j
\end{aligned}
$$

Thus:

$$
J(W_0(0),W_1(0),0|\lambda_j,\pi_m) = J(W_0(0),W_1(0),0|0,\pi_m) + \frac{1-\pi_m}{\rho}G(\lambda_j,\pi_m)\lambda_j \tag{A.3.41}
$$

We now prove that if $J(W_0(0),W_1(0),0|\lambda_j,\pi_m)$ is concave in λ_j for any $\pi_m > 0$, then $G(\lambda_j,\pi_m)$ decreases in λ_j, for any $\pi_m > 0$. The proof is by contradiction. Suppose that $G(\lambda_j,\pi_m)$ does not monotonically decrease in λ_j, then by the continuity of $G(\lambda_j,\pi_m)$ in λ_j, there must exist an interval of λ_j inside $[0,+\infty)$ on which $G(\lambda_j,\pi_m)\lambda_j$ is (locally) convex in λ_j, a contradiction to the concavity of $J(W_0(0),W_1(0),0|\lambda_j,\pi_m)$.

Finally, we prove that if $J(W_0(0),W_1(0),0|\lambda_j,\pi_m=0)$ is concave in λ_j, then so is $J(W_0(0),W_1(0),0|\lambda_j,\pi_m)$. Since A.3.41 applies for any $1\geq\pi_m\geq 0$, Then:

$$
J(W_0(0),W_1(0),0|\lambda_j,0) = J(W_0(0),W_1(0),0|0,\pi_m) + \frac{1}{\rho}G(\lambda_j,\pi_m)\lambda_j \tag{A.3.42}
$$

Notice that $J(W_0(0),W_1(0),0|0,\pi_m)$ is flat with respect to λ_j, thus the concavity of $J(W_0(0),W_1(0),0|\lambda_j,0)$ implies the concavity of $\frac{1}{\rho}G(\lambda_j,\pi_m)\lambda_j$, which in turn implies the concavity of $J(W_0(0),W_1(0),0|\lambda_j,\pi_m)$. □

Proof. [Proof of proposition 8] A closed-form solution is not currently possible even for the problem without transaction cost Pham (2009), but the problem without transaction cost allows for an asymptotic Taylor expansion of the solution, as in Roger and Zane (2002). First, we define $l \equiv \lambda_j^{-1}$ is the latency of the trades for taker j, where the subscript is suppressed on the left hand side. We apply a Taylor expansion around $l = 0$ and get that $J[x, y, 0|\lambda_j, 0]$ can be expressed by that of the Merton (1969) problem with continuous trading, adjusted by higher order terms.

$$J[x, y, 0|\lambda_j, 0] = a^{-\delta}U(x + y) + \sum_{n=1}^{\infty} b_n \frac{l^n}{n!}, \tag{A.3.43}$$

where $a^{-\delta}U(x + y)$ is the value function of the Merton (1969) problem, and $a^{-\delta} \equiv \frac{1}{\delta}\left(\rho + (\delta - 1)\left[r + \frac{(\alpha-r)^2}{2\delta\sigma^2}\right]\right)$ is a coefficient entirely composed of exogenous parameters. Naturally, $\sum_{n=1}^{\infty} b_n \frac{l^n}{n!} < 0$, such that $J[x, y, 0|\lambda_j, 0] < a^{-\delta}U(x + y)$ for $l > 0$. Hence, the (asymptotic) shape of $J[x, y, 0|\lambda_j, 0]$ depends on $b_n, n \geq 1$.

Following Roger and Zane (2002), we can solve for:

$$b_1 = -\frac{\sigma^6\delta^3 w^2(1 - w)^2}{2\sigma^2\delta\rho + (\delta - 1)[2\delta\sigma^2 r + (\alpha - r)^2]}, \tag{A.3.44}$$

where $w \equiv \frac{\alpha-r}{\sigma^2\delta}$ is the equilibrium portfolio weight of the risky asset in the Merton (1969) portfolio problem, which is entirely composed of exogenous parameters. As long as the risk-aversion parameter δ is large enough (and for sure when $\delta > 1$), we have $b_1 < 0$.Assuming that latency l is very small, a fact very much true in the context of HFTs, terms $b_2, b_3, \ldots$ are negligible and we have

$$J[x, y, 0|\lambda_j, 0] \approx a^{-\delta}U(x + y) + b_1 l = a^{-\delta}U(x + y) + b_1\lambda_j^{-1}. \tag{A.3.45}$$

Taking the second order derivative we see that the value function is asymptotically concave in trading speed if δ is sufficiently large:

$$\frac{\partial^2 J[x,y,0|\lambda_j,0]}{\partial(\lambda_j)^2} = \frac{\partial^2(a^{-\delta}U(x+y)+b_1\lambda_j^{-1})}{\partial(\lambda_j)^2}, \tag{A.3.46}$$

$$= 2b_1\lambda_j^{-3} < 0, \text{for} \lambda_j \text{big enough.} \tag{A.3.47}$$

□

A.4 Notation Summary

Symbol	*Support*	*Description*
	Parameters	
M	$[1, \infty)$	the total number of makers
N	$[1, \infty)$	the total number of takers
G_0	$[0, \infty)$	marginal gain per trade in linear GFT setting
π_m	$[0, 1]$	share of gains from trade that is captured by the maker
π_t	$[0, 1]$	share of gains from trade that is captured by the taker
β	$[0, \infty)$	technological cost parameter for each maker
γ	$[0, \infty)$	technological cost parameter for each taker
r	$[0, \infty)$	the risk free rate
α, σ	$[0, \infty)$	the parameters of the Brownian motion
δ	$[0, 1) \cup (1, \infty)$	the representative taker's relative risk aversion.
	States of nature	
$X_0(t)$	$[0, \infty)$	the price trajectory of the riskless asset accumulation
$X_1(t)$	$[0, \infty)$	the price trajectory of the stock price
$B(t)$	$[0, \infty)$	Geometric Brownian Motion serving as control process of $S(t)$.
$P(t)$	$[0, \infty)$	representative taker's Poisson process recording transactions
	Indices	
i	$\{1, 2, .., M)\}$	maker i
j	$\{1, 2, .., N)\}$	taker j
t	$\{0, .., \infty\}$	time
s_n	$\{1, 2, .., \infty\}$	representative taker's trading time
	Decision variables	
μ_i	$[0, \infty)$	the intensity parameter of maker i.
τ_j	$[0, \infty)$	the intensity parameter of taker j.
$W_0(t)$	$[0, \infty)$	the wealth invested in the riskless asset at time t.
$W_1(t)$	$[0, \infty)$	the wealth invested in the stock at time t.
$Q(t)$	$(-\infty, \infty)$	the amount transferred from the stock to the riskless asset at time t (when trading is possible).
$C(t)$	$[0, \infty)$	the amount of consumption at time t.

Appendix B

Asset risk and bank runs

B.1 Auxiliary lemmas and proofs

In this section of the Appendix, we first present technical, auxiliary lemmas (with proofs) that help in proving the main results of the paper, but add little intuition. Next, we provide the proofs of the main results of the paper.

B.1.1 Auxiliary lemmas

The following lemmas are of purely technical purpose. All the properties of the functions of μ defined below are valid on the interval $\left(\frac{lr_2}{r_1} - 1, +\infty\right)$ unless otherwise specified.

Lemma 12 *Function*

$$f(\mu) = (\mu + 1)\log\left(\frac{(\mu + 1)r_1}{(\mu + 1)r_1 - lr_2}\right) \tag{B.1.1}$$

is strictly decreasing for $\mu \in \left(\frac{lr_2}{r_1} - 1, +\infty\right)$ *, and*

$$f(\mu) > \frac{lr_2}{r_1} \tag{B.1.2}$$

Proof. We first determine the monotonicity of $f(\mu)$:

The derivative of f is given by:

$$f'(\mu) = \frac{(\mu + 1)r_1 \log\left(\frac{(\mu+1)r_1}{(\mu+1)r_1 - lr_2}\right) - lr_2\left(\log\left(\frac{(\mu+1)r_1}{(\mu+1)r_1 - lr_2}\right) + 1\right)}{(\mu + 1)r_1 - lr_2} \tag{B.1.3}$$

The sign of f' is not obvious, so we determine its sign by analyzing instead the second derivative which has a simpler form:

$$f''(\mu) = \frac{l^2 r_2^2}{(\mu + 1)\left((\mu + 1)r_1 - lr_2\right)^2} > 0. \tag{B.1.4}$$

Therefore, $f'(\mu)$ is itself monotonically increasing on $(\frac{r_2}{r_1}l-1,+\infty)$.

Next, we find the limiting values of f'[1]:

$$f'(\mu) \xrightarrow{\mu \to \left(\frac{r_2}{r_1}l-1\right)^+} -\infty \tag{B.1.5}$$

$$f'(\mu) \xrightarrow{\mu \to +\infty} 0. \tag{B.1.6}$$

Due to the monotonicity of $f'(\mu)$ on $\left(\frac{r_2}{r_1}l-1,+\infty\right)$, we conclude that $f'(\mu)<0$ for any μ on the interval $\left(\frac{r_2}{r_1}l-1,+\infty\right)$, that is, f strictly decreases on this interval.

Given the monotonicity of f, the range of values of $f(\mu)$ can be found by determining the limiting values of f on the interval of interest.

$$f(\mu) \xrightarrow{\mu \to (\frac{r_2}{r_1}l-1)^+} +\infty \tag{B.1.7}$$

$$f(\mu) \xrightarrow{\mu \to +\infty} \frac{lr_2}{r_1} \tag{B.1.8}$$

Thus $f(\mu) > \frac{lr_2}{r_1}$ on the interval $\left(\frac{r_2}{r_1}l-1,+\infty\right)$.

□

Lemma 13 *The function*

$$h(\mu) = -(\mu+1)r_1 \log\left(\frac{(\mu+1)r_1}{(\mu+1)r_1 - lr_2}\right) + lr_2\left(\log\left(\frac{(\mu+1)r_1}{(\mu+1)r_1 - lr_2}\right)+1\right) \tag{B.1.9}$$

is strictly decreasing on $\left(\frac{lr_2}{r_1}-1,+\infty\right)$, *and*

$$0 < h(\mu) \le lr_2 \tag{B.1.10}$$

Proof.

[1] In this paper, we use notation $a(x) \xrightarrow{x_0^+}$ to denote the “right limit” of $a(x)$ at x_0. In other words, it is the value of $a(x)$, when x approaches x_0 from the right on the number axis. Similarly, we use $a(x) \xrightarrow{x_0^-}$ to denote the “left limit”. “Left limit” is not the same as the “right limit”, if function $a(x)$ is discontinuous at $x = x_0$. In this paper, however, when we find the left or right limit because it’s easier and sufficient for our proof. We do not imply that the function is discontinuous at that point.

For h:

$$h'(\mu) = \frac{lr_2 - r_1 f(\mu)}{\mu + 1} < 0, \tag{B.1.11}$$

which is a consequence of the range of $f(\mu)$ that was proved in (B.1.2).

Given the monotonicity of h, the range of its values can be found by determining its extreme values on the interval of interest.

$$h(\mu) \xrightarrow{\mu \to (\frac{r_2}{r_1} l - 1)^+} lr_2 \tag{B.1.12}$$

$$\xrightarrow{\mu \to +\infty} 0 \tag{B.1.13}$$

□

Lemma 14 *For $\mu \in \left(\frac{lr_2}{r_1} - 1, +\infty\right)$, function $z(\mu, e)$ strictly increases in μ and e. Furthermore, for $\mu \in \left(\underline{\mu}, +\infty\right)$ and $e \in \left(\frac{\lambda r_1}{l} + 1, \frac{r_1}{l} + 1\right)$, we have that:*

$$0 < z(\mu, e) < (1 - \lambda)\, lr_2 \tag{B.1.14}$$

Proof.

We first differentiate z w.r.t. μ to obtain

$$\frac{\partial z}{\partial \mu} \tag{B.1.15}$$

$$= \frac{\big(r_1 - l(e+1)\big)\left(-(\mu+1)r_1 \log\left(\frac{(\mu+1)r_1}{(\mu+1)r_1 - lr_2}\right) + lr_2\left(\log\left(\frac{(\mu+1)r_1}{(\mu+1)r_1 - lr_2}\right) + 1\right)\right)}{(\mu+1)r_1 - lr_2},$$

$$= \frac{\big(r_1 - l(e+1)\big)\, h(\mu)}{(\mu+1)r_1 - lr_2} > 0. \tag{B.1.16}$$

Next, we differentiate z with respect to e to obtain.

$$\frac{\partial z}{\partial e} = lf(\mu) > 0 \tag{B.1.17}$$

Given the monotonicity in μ, to show that $z(\mu, e) > 0$, it suffice to show that its value is positive at its lower bound. In the following derivation we show that it is indeed the case.

$$
\begin{aligned}
&z(\mu, e|\mu = \underline{\mu}) && \text{(B.1.18)}\\
&= \frac{(1-\lambda) l r_2 \left(l(e+1) \left(\log\left(\frac{(1-\lambda)r_1}{r_1 - l(e+1)}\right) + 1 \right) - r_1 \left(\lambda + \log\left(\frac{(1-\lambda)r_1}{r_1 - l(e+1)}\right) \right) \right)}{le + l - \lambda r_1}\\
&\overset{sign}{=} \left(l(e+1) - r_1\right) \log\left(\frac{(1-\lambda)r_1}{r_1 - l(e+1)}\right) + l(e+1) - r_1\lambda\\
&> \left(l(e+1) - r_1\right) \left(\frac{(1-\lambda)r_1}{r_1 - l(e+1)} - 1\right) + l(e+1) - r_1\lambda\\
&= 0 && \text{(B.1.19)}
\end{aligned}
$$

To prove that $z(\mu, e) < (1-\lambda)\, l r_2$, we have shown in (B.1.17) that z is increasing in e. Now we show that the value of z at the upper bound of e is $(1-\lambda)\, l r_2$.

$$
z(\mu, e) \xrightarrow{e \to \left(\frac{r_1}{l} - 1\right)^-} (1-\lambda) l r_2 \qquad \text{(B.1.20)}
$$

□

Lemma 15 $(1-\lambda)h(\mu) < z(e, \mu)$*, for* $\mu \in \left(\underline{\mu}, 1\right)$,

Proof.

The relative magnitude of $(1-\lambda)h(\mu)$ and $z(e, \mu)$ can be shown by the facts that $z(e, \mu) - (1-\lambda)h(\mu)$ increases in μ and that the differential is 0 at the lower bound of μ.

First, we determine the sign of $\frac{\partial\left(z(\mu)-(1-\lambda)h(\mu)\right)}{\partial\mu}$:

$$\begin{aligned}
&\frac{\partial\left(z(\mu)-(1-\lambda)h(\mu)\right)}{\partial\mu}\\
=&\frac{-l^2r_2\Big((e+1)\left(f(\mu)+\mu+1\right)-(1-\lambda)r_2\Big)-\lambda r_1^2(\mu+1)f(\mu)}{(\mu+1)\left((\mu+1)r_1-lr_2\right)}\\
&+\frac{lr_1\Big(\lambda r_2\left(f(\mu)+\mu+1\right)+(e+1)(\mu+1)f(\mu)\Big)}{(\mu+1)\left((\mu+1)r_1-lr_2\right)}\\
=&\frac{f(\mu)\left((\mu+1)r_1-lr_2\right)\left(l(e+1)-\lambda r_1\right)+(1-\lambda)l^2r_2^2-l(\mu+1)r_2\left(l(e+1)-\lambda r_1\right)}{(\mu+1)\left((\mu+1)r_1-lr_2\right)}\\
&\text{then by } f(\mu)>\frac{lr_2}{r_1}\text{, which is a result proved in Lemma 12, the above derivative is:}\\
>&\frac{\frac{lr_2}{r_1}\left((\mu+1)r_1-lr_2\right)\left(l(e+1)-\lambda r_1\right)+(1-\lambda)l^2r_2^2-l(\mu+1)r_2\left(l(e+1)-\lambda r_1\right)}{(\mu+1)\left((\mu+1)r_1-lr_2\right)}\\
=&\frac{l^2r_2^2\left(r_1-le-l\right)}{(\mu+1)r_1\left((\mu+1)r_1-lr_2\right)}\\
>&0
\end{aligned}\tag{B.1.21}$$

Now we show that the value of the differential is 0 at the lower bound of μ, which concludes the proof.

$$z(e,\mu)-(1-\lambda)h(\mu)\overset{\mu=\underline{\mu}}{=}0\tag{B.1.22}$$

□

Lemma 16 *For $e\in\left(\frac{\lambda r_1}{l}-1,\frac{r_1}{l}-1\right)$, function $k(e)$ strictly increases, and*

$$0<\lambda(1-\lambda)r_1<-\lambda r_1\log(\lambda)<k(e)<(1-\lambda)r_1\tag{B.1.23}$$

Proof.

For $e\in\left(\frac{\lambda r_1}{l}-1,\frac{r_1}{l}-1\right)$, differentiating k, yields

$$k'(e)=-l\log\left(\frac{l(e+1)}{r_1}\right)>0\tag{B.1.24}$$

Given the monotonicity of k, the range of values of $k(e)$ can be found by determining the extreme values of k on the boundaries of the support.

$$k(e) \xrightarrow{e \to (\frac{\lambda r_1}{l} - 1)^+} -\lambda r_1 \log(\lambda) > 0 \tag{B.1.25}$$

$$k(e) \xrightarrow{e \to (\frac{r_1}{l} - 1)^-} (1 - \lambda) r_1 > 0 \tag{B.1.26}$$

□

Proposition 4:

Proof.

Our proof has three stages. First, we assume equilibrium threshold strategies and derive the optimal common threshold. Next, we show that there is only one such threshold, proving uniqueness. Finally, we prove that given this unique optimal threshold, it is individually optimal to engage in these threshold strategies.

Consider a patient depositor that has her signal equalling θ_i. Then her expected utility differential defined by formula (3.11) should be a function of the behavior of other patient depositors represented by the common threshold θ' and should have the following form:

$$\Delta(\theta_i, \theta') \equiv \mathbb{E}\left(v(\theta, n(\theta, \theta'))|\theta_i\right) = \frac{1}{2\epsilon} \int_{\theta=\theta_i-\epsilon}^{\theta_i+\epsilon} \frac{1}{\mu} v(\theta, n(\theta, \theta')) d\theta, \tag{B.1.27}$$

where $\frac{1}{\mu}$ is the probability density function of θ, which is uniformly distributed on interval $[1-\mu, 1]$.

Since depositor i would roll over if her signal is above θ' and withdraw if below, then if her signal is θ', she should be indifferent between withdrawal and rollover, that is:

$$\Delta(\theta', \theta') = 0 \tag{B.1.28}$$

. $\Delta(\theta', \theta')$ can be considered a function of $\theta' \in [0, 1]$. For equilibrium existence, we need to prove that there is a θ' such that (B.1.28) holds.

With the assumption of the existence of upper dominance region and lower dominance region, we can show that there is θ' low enough such that $\Delta(\theta', \theta') < 0$ and it high enough such that $\Delta(\theta', \theta') > 0$. The existence of lower dominance region means that when the state is so slow that $\theta \leq \underline{\theta}$, the signal of all depositors are as a result low enough such that, no matter what the shared threshold θ' is, they would run.

That is: $\Delta(\theta_i, \theta') < 0$; and naturally for $\theta' \leq \underline{\theta}$, we have: $\Delta(\theta', \theta') < 0$. Using the similar argument, assuming the upper dominance region exist, we have: for $\theta' \geq \bar{\theta}$, we have: $\Delta(\theta', \theta') > 0$.

Then to prove the existence of an equilibrium it suffices to show that $\Delta(\theta', \theta')$ is continuous in θ'. This is indeed the case as the integrand $\frac{1}{\mu}v(\theta, n(\theta, \theta'))$ is bounded. $v(\theta, n(\theta, \theta'))$ is discontinuous at $\bar{\theta}$, but the continuity of the integrand is not necessary for the proof of the continuity of the integral as the discontinuity of the integrand is smoothed out by integration.

Furthermore, to show that there is only one symmetric threshold equilibrium, we only need to show that $\Delta(\theta', \theta')$ is increasing in θ'. This is indeed the case because when both the private signal and the threshold strategy increase by the same amount, the depositor's expectation of the withdrawal proportion of the entire depositor base n is unchanged, while the chance of paying off for waiting is higher. From here on, we write the θ' that satisfies the fix point condition above θ^*, which is a candidate of the equilibrium threshold.

Lastly, we show that θ^* is indeed an equilibrium. To this end, we need to show that for any $\theta_i < \theta^*$, $\Delta(\theta_i, \theta^*) < 0$[2], and for any $\theta_i > \theta^*$, $\Delta(\theta_i, \theta^*) > 0$. To show this, we first prove that function $v(\theta, n(\theta, \theta^*))$ crosses 0 only once:

Claim 1 *There is only one* θ_0[3] *such that* $v(\theta_0, n(\theta_0, \theta^*)) = 0$. *Furthermore, for any* $\theta < \theta_0$, $v(\theta, n(\theta, \theta^*)) < 0$ *and for any* $\theta > \theta_0$, $v(\theta, n(\theta, \theta^*)) > 0$.

Proof of Claim 1.1 *As can be seen from the formula* (3.10), $v(\theta, n(\theta, \theta^*)) > 0$, *if* $\lambda \leq n(\theta) \leq \tilde{n}$; *and* $v(\theta, n(\theta, \theta^*)) < 0$, *if* $\dot{n} \leq n(\theta) \leq 1$. *So it suffice to show that* $v(\theta, n(\theta, \theta^*)) > 0$ *is monotonic when* $\tilde{n} < n(\theta) < \dot{n}$. *This is indeed the case by taking derivative with respect to* θ *for* $\theta\frac{(1+\mu)\left(e+1-\frac{n(\theta)r_1}{l}\right)}{(1-n(\theta))} - r_1$, *applying chain rule and obtaining* $\frac{\partial n(\theta)}{\partial \theta}$ *by taking derivative for formula* (3.15).

Then compared with the value of $\Delta(\theta^*, \theta^*)$, for any $\theta < \theta^*$, the probability is shifted from the positive values of v to the negative values of v. As a result, $\Delta(\theta, \theta^*) < \Delta(\theta^*, \theta^*)$, since we've shown above that $\Delta(\theta^*, \theta^*) = 0$, we conclude: $\Delta(\theta, \theta^*) < 0$, for any $\theta < \theta^*$. A symmetric argument can be made for: $\Delta(\theta, \theta^*) > 0$, for any $\theta > \theta^*$. This concludes the entire proof.

□

Lemma 7

[2]Note that this is not implied by what we have proved above: for any $\theta' < \theta^*$, $\Delta(\theta', \theta') < 0$.

[3]As the root of v, θ_0 is usually different in value from θ^*, which is the root of the *expected* value of v.

Proof.

θ^* is defined by the indifference condition (B.1.28) when evaluated in θ^*. This means that for any depositor i, when all other depositors adopt θ^* as the withdraw threshold, and when depositor i happens to receive a signal at θ^*, then she is indifferent between withdrawing and waiting. To write this in full:

$$
\begin{aligned}
\Delta(\theta^*,\theta^*) &= \frac{1}{2\epsilon}\int_{\theta=\theta^*-\epsilon}^{\theta^*+\epsilon}\frac{1}{\mu}v(\theta,n(\theta,\theta^*))d\theta \\
&\text{(plug in values and bounds in Table 3.1:)} \\
&= \frac{1}{2\epsilon}\int_{\theta^*-\epsilon}^{\tilde{\theta}}\frac{1}{\mu}\left(0-\frac{(e+1)l}{n(\theta,\theta^*)}\right)d\theta \\
&+\frac{1}{2\epsilon}\int_{\tilde{\theta}}^{\dot{\theta}}\frac{1}{\mu}\left(\theta\frac{(1+\mu)\left(e+1-\frac{n(\theta,\theta^*)r_1}{l}\right)}{(1-n(\theta,\theta^*))}-r_1\right)d\theta \\
&+\frac{1}{2\epsilon}\int_{\dot{\theta}}^{\theta^*+\epsilon}\frac{1}{\mu}\left(\theta r_2-r_1\right)d\theta \\
&= 0
\end{aligned}
\tag{B.1.29}
$$

Based on the relationship between n and θ defined in formula (3.15), changing measure from θ to n, equation (B.1.29) can alternatively be written as:

$$
\begin{aligned}
\Delta(\theta^*,\theta^*) &= \int_{\dot{n}}^{1}\left(0-\frac{(e+1)l}{n}\right)dn \\
&+\int_{\tilde{n}}^{\dot{n}}\left(\theta\frac{(1+\mu)\left(e+1-\frac{nr_1}{l}\right)}{(1-n)}-r_1\right)dn \\
&+\int_{\lambda}^{\tilde{n}}\left(\theta r_2-r_1\right)dn \\
=&0
\end{aligned}
\tag{B.1.30}
$$

Similar with the solution strategy in Goldstein and Pauzner (2005), to simplify this formula further so we can pull θ out of the integrals in the second and third term in equation (B.1.30), we now show that θ is approximately constant and equal to θ^* for $n \in [\lambda, \dot{n}]$.

First, we express θ as a function of n. Reversely from formula (3.15), given a threshold θ^* common to all depositors, for a fraction of withdrawers $n \in (\lambda, 1)$, the realized state θ that led to this fraction must be:

$$\theta = \theta^* + \epsilon \left(1 - \frac{2n - 2\lambda}{1-\lambda}\right) \tag{B.1.31}$$

We can see that θ converges to θ^* for $n \in [\lambda, \dot{n}]$.

Then similar to Goldstein and Pauzner (2005), taking limit with respect to $\epsilon \to 0$ for both sides of equation (B.1.30), we get:[4]

$$\begin{aligned}\lim_{\epsilon \to 0} \Delta(\theta^*, \theta^*) = &\int_{\dot{n}}^{1} \left(0 - \frac{(e+1)l}{n}\right) dn \\ &+ \int_{\tilde{n}}^{\dot{n}} \left(\theta^* \frac{(1+\mu)\left(e + 1 - \frac{nr_1}{l}\right)}{(1-n)} - r_1\right) dn \\ &+ \int_{\lambda}^{\tilde{n}} \left(\theta^* r_2 - r_1\right) dn \\ =&0\end{aligned} \tag{B.1.32}$$

Solving for θ^*, we obtain:

$$\begin{aligned}\lim_{\epsilon \to 0} \theta^* &= \frac{\int_{\dot{n}}^{1} \frac{(e+1)l}{n} dn + \int_{\tilde{n}}^{\dot{n}} r_1 dn + \int_{\lambda}^{\tilde{n}} r_1 dn}{\int_{\tilde{n}}^{\dot{n}} \left[\frac{(1+\mu)\left(e+1-\frac{nr_1}{l}\right)}{(1-n)}\right] dn + \int_{\lambda}^{\tilde{n}} r_2 dn} \\ &\text{(Making integration of simple functions } \frac{1}{n}, \frac{1}{1-n} \text{ and 1, and plugging in} \\ &\text{formula (3.5) of } \dot{n} \text{ and formula (3.7) of } \tilde{n}\text{:)} \\ &= -\frac{l\left(l(e+1)\left(\log\left(\frac{r_1}{l(e+1)}\right) + 1\right) - \lambda r_1\right)}{\left(\left(r_1 - l(e+1)\right)(\mu+1)\log\left(\frac{(\mu+1)r_1}{(\mu+1)r_1 - lr_2}\right) - (1-\lambda) l r_2\right)} \end{aligned} \tag{B.1.33}$$

(For the ease of notation, replacing blocks of expressions with function $k(e)$ based on formula (3.17) and with function $z(\mu, e)$ based on (3.18):)

$$= \frac{lk(e)}{z(e,\mu)} \tag{B.1.34}$$

[4] In this paper, $\lim_{\text{condition}}$ and $\xrightarrow{\text{condition}}$ are used interchangeably.

□

Lemma 8

Proof. Since bankruptcy probability is intimately linked with the bankruptcy threshold of the state θ, we first compute the thresholds for period 1 and period 2 bankruptcy, respectively.

Period 1 bankruptcy is triggered when there are $\dot{n} \equiv \frac{(e+1)l}{r_1}$ (from condition (3.5) mass of depositors withdrawing. The quantity of withdrawers in period 1 depends on the realized state θ, since the depositors decide whether to withdraw based on their signals of $\theta_i = \theta + \epsilon_i$. In proving Lemma 7, we derived the formula (B.1.31) of θ as a function of n. This allows us to pin down the threshold $\dot{\theta}$ by plugging the formula of $\dot{n}$ into formula (B.1.31):

$$\dot{\theta} = \theta^* + \epsilon \left(1 - \frac{2\dot{n} - 2\lambda}{1-\lambda}\right) = \theta^* + \epsilon \left(1 - \frac{2\frac{(e+1)l}{r_1} - 2\lambda}{1-\lambda}\right) \tag{B.1.35}$$

Similarly, since $\tilde{n}$ is the critical mass for period 2 bankruptcy, we can plug its formula as in (3.7) into formula (B.1.31) to obtain the bankruptcy threshold of period 2:

$$\tilde{\theta} = \theta^* + \epsilon \left(1 - \frac{2\tilde{n} - 2\lambda}{1-\lambda}\right) = \theta^* + \epsilon \left(1 - \frac{2\frac{(1+\mu)(e+1)-r_2}{\frac{(1+\mu)r_1}{l} - r_2} - 2\lambda}{1-\lambda}\right) \tag{B.1.36}$$

Now we can see that:

$$\lim_{\epsilon \to 0} \dot{\theta} = \lim_{\epsilon \to 0} \tilde{\theta} = \theta^* \tag{B.1.37}$$

In other words, θ is approximately always equal to θ^* in the interval $n \in [\tilde{n}, \dot{n}]$.

Since θ is uniformly distributed on $[1-\mu, 1]$, the probability π_1 of period 1 bankruptcy can be expressed as a function of $\dot{\theta}$ as follows:

$$\pi_1 = \text{prob}(\theta \leq \dot{\theta}) = \frac{\dot{\theta} - (1-\mu)}{\mu}, \tag{B.1.38}$$

Similarly, period 2 bankruptcy probability π_2 can be written as:

$$\pi_2 = \text{prob}(\theta \leq \tilde{\theta}) = \frac{\tilde{\theta} - (1-\mu)}{\mu} \tag{B.1.39}$$

Then by equation (B.1.37), we have that the two probabilities converge:

$$\begin{aligned}
&\lim_{\epsilon \to 0} \pi_1 = \lim_{\epsilon \to 0} \pi_2 \\
=&\frac{\theta^* - (1-\mu)}{\mu} \\
\xrightarrow{\epsilon \to 0}&\frac{\frac{lk(e)}{z(e,\mu)} - 1 + \mu}{\mu} \equiv \pi,
\end{aligned} \tag{B.1.40}$$

where we use Lemma 7 for the last step. □

Proposition 5

Proof.

As $\epsilon \to 0$, run probability π has the formula as in (3.19) due to Lemma 8. Therefore, we have that

$$
\begin{aligned}
\frac{\partial \pi}{\partial e} &\xrightarrow{\epsilon \to 0} \frac{\partial}{\partial e}\frac{\theta^* - 1 + \mu}{\mu} = \frac{1}{\mu}\frac{\partial \theta^*}{\partial e} \\
&= \frac{l^2\left(\left(\left(\lambda - \log\left(\frac{r_1}{l(1+e)}\right)\right) r_1 - (1+e)l\right) f(\mu) + (1-\lambda) l r_2 \log\left(\frac{r_1}{l(1+e)}\right)\right)}{\mu z(\mu)^2} \\
&\overset{sign}{=} \left(\left(\lambda - \log\left(\frac{r_1}{l(1+e)}\right)\right) r_1 - (1+e)l\right) f(\mu) + (1-\lambda) l r_2 \log\left(\frac{r_1}{l(1+e)}\right) \\
&\text{(Using } f(\mu) > \frac{l r_2}{r_1} \text{ from Lemma 12, and the coefficient of } f(x) \text{ being negative} \\
&\text{because } \lambda r_1 < (1+e)l \text{ from assumption (3.8):)} \\
&< \left(\lambda - \log\left(\frac{r_1}{l(1+e)}\right)\right) l r_2 - (1+e)\frac{l^2 r_2}{r_1} + (1-\lambda) l r_2 \log\left(\frac{r_1}{l(1+e)}\right) \\
&= \lambda l r_2 - (1+e)\frac{l^2 r_2}{r_1} - \lambda l r_2 \log\left(\frac{r_1}{l(1+e)}\right) \\
&\text{(Using a property of logarithm: } \log(x) > 1 - \frac{1}{x} \text{ for } x > 0) \\
&< \lambda l r_2 - (1+e)\frac{l^2 r_2}{r_1} - \lambda l r_2 \left(1 - \frac{l(1+e)}{r_1}\right) \\
&= -(1+e)\frac{l^2 r_2}{r_1} + \lambda \frac{l^2 r_2 (1+e)}{r_1} \\
&< 0
\end{aligned} \tag{B.1.41}
$$

where $f(\mu)$ is defined as in Lemma 12.

□

Lemma 9

Proof. Taking derivative with respect to μ for θ^* as in formula (3.16):

$$\frac{\partial}{\partial\mu}\theta^*(\mu) = -\frac{l\left(r_1-(e+1)l\right)h(\mu)\left((e+1)l\left(1+\log\left(\frac{r_1}{(e+1)l}\right)\right)-\lambda r_1\right)}{\left((\mu+1)r_1-lr_2\right)\left((r_1-l(e+1))f(\mu)-l(1-\lambda)r_2\right)^2}$$

Here we replace blocks of expressions with functions k, h and z defined in Lemmas 16, 13 and 14, respectively:

$$= -\frac{l\left(r_1-l(e+1)\right)k(e)h(\mu)}{\left((\mu+1)r_1-lr_2\right)\left(z(e,\mu)\right)^2} \tag{B.1.42}$$

$$< 0, \tag{B.1.43}$$

where $r_1-l(e+1)>0$ is implied by (3.2); $(\mu+1)r_1-lr_2>0$ since $1+\mu>r_2$ and $r_1>l$ by assumption, and the signs of functions k, h and z are proved in Lemma 16, Lemma 13 and Lemma 14 respectively. □

Proposition 6

Proof. As $\epsilon\to 0$, run probability π has the formula as in (3.19).

First, we show that $\frac{\partial}{\partial e}\frac{\partial\pi}{\partial\mu}>0$, so that the task of proof is reduced to finding any interior point of e for which $\frac{\partial\pi}{\partial\mu}<0$, then any e below that point gives rise to $\frac{\partial\pi}{\partial\mu}<0$. Finally, we determine a loose upper bound of e that makes this happen.

Lemma 17 *For any combinations of parameters within their support,*

$$\frac{\partial}{\partial e}\frac{\partial\pi}{\partial\mu}>0 \tag{B.1.44}$$

Proof. We have that

$$\frac{\partial}{\partial e}\frac{\partial\pi}{\partial\mu} = \mu\frac{\partial}{\partial e}\frac{\frac{\mathrm{d}\theta^*(\mu)}{\mathrm{d}\mu}-\theta^*(\mu)+1}{\mu^2} \overset{sign}{=} \mu\frac{\partial}{\partial e}\frac{\mathrm{d}\theta^*(\mu)}{\mathrm{d}\mu}-\frac{\partial}{\partial e}\theta^*(\mu). \tag{B.1.45}$$

Next, we derive expressions for the two terms separately. To obtain the derivative of the first term with respect to e, we take derivative of $\frac{\mathrm{d}\theta^*(\mu)}{\mathrm{d}\mu}$ from (B.1.42) with respect to e:

$$\begin{aligned}
&\mu\frac{\partial}{\partial e}\frac{\mathrm{d}\theta^*(\mu)}{\mathrm{d}\mu}\\
=&-\mu\frac{\partial}{\partial e}\frac{l\left(r_1-(e+1)l\right)h(\mu)k(e)}{\left((\mu+1)r_1-lr_2\right)z(e,\mu)^2}\\
=&\mu\frac{l^2\left(r_1-(e+1)l\right)h(\mu)\left(2k(e)f(\mu)-\log\left(\frac{r_1}{l(1+e)}\right)z(e,\mu)\right)}{\left((\mu+1)r_1-lr_2\right)z(e,\mu)^3}
\end{aligned}\tag{B.1.46}$$

$r_1-l(e+1)>0$ is implied by (3.2); $(\mu+1)r_1-lr_2>0$ since $1+\mu>r_2$ and $r_1>l$ by assumption; and the signs of functions h and z are proved in Lemma 13 and Lemma 16 respectively, thus to show the above formula positive, it suffice to show:

$$\left(2k(e)f(\mu)-\log\left(\frac{r_1}{l(1+e)}\right)z(e,\mu)\right)>0\tag{B.1.47}$$

And condition (B.1.47) holds because:

$$\begin{aligned}
&k(e)f(\mu)-\log\left(\frac{r_1}{l(1+e)}\right)z(e,\mu)\\
=&\left(l(e+1)\left(1+\log\left(\frac{r_1}{l(e+1)}\right)\right)-\lambda r_1\right)f(\mu)\\
&-\log\left(\frac{r_1}{l(e+1)}\right)\left(-\left(r_1-l(e+1)\right)f(\mu)+(1-\lambda)lr_2\right)\\
=&\left(l(e+1)-\lambda r_1\right)f(\mu)+l(e+1)\log\left(\frac{r_1}{l(e+1)}\right)f(\mu)\\
&-l(e+1)\log\left(\frac{r_1}{l(e+1)}\right)f(\mu)+\log\left(\frac{r_1}{l(1+e)}\right)\left(r_1f(\mu)-(1-\lambda)lr_2\right)\\
&\text{(By Lemma 12:)}\\
>&\left(l(e+1)-\lambda r_1\right)f(\mu)+\log\left(\frac{r_1}{l(1+e)}\right)\left(r_1\frac{r_2}{r_1}l-(1-\lambda)lr_2\right)\\
=&\left(l(e+1)-\lambda r_1\right)f(\mu)+\log\left(\frac{r_1}{l(1+e)}\right)\lambda lr_2\\
>&0
\end{aligned}\tag{B.1.48}$$

Now to prove the lemma, we only need to show that the second term in formula (B.1.45) is also positive, that is: $\frac{\partial}{\partial e}\theta^*(\mu) < 0$. This is the case because:

$$\begin{aligned}
&\frac{\partial}{\partial e}\theta^*(\mu) = \frac{\partial}{\partial e}\frac{lk(e)}{z(e,\mu)} \\
=&\frac{l^2}{z(e,\mu)^3}\left(z(e,\mu)\left(k(e)f(\mu) - \log\left(\frac{r_1}{l(1+e)}\right)z(e,\mu)\right) + \frac{k(e)\mu h(\mu)z(e,\mu)}{(1+\mu)r_1 - lr_2}\right)
\end{aligned} \tag{B.1.49}$$

similar to the arguments made for the derivative of the first term, $k(e), h(\mu), z(e,\mu)$ and $(1+\mu)r_1 - lr_2$ are all positive. Thus to show the above formula positive, it also suffices to show (B.1.47), which we have already proved. This concludes the entire proof.

□

Now after proving Lemma 17, it remains to be shown that there is at least one e_0 within the support of e restricted by the assumptions made so far in this paper, such that when $e = e_0$, $\frac{\pi}{\partial\mu} < 0$. Then by Lemma 17, holding other parameters constant, for any $e \in \left(\frac{\lambda r_1}{l} - 1, e_0\right)$, we also have $\frac{\pi}{\partial\mu} < 0$. We show that such an e_0 exists by giving a numeric example:

It can be checked that, when $r_1 = 1.5, r_2 = 1.55, l = 0.7, \lambda = 0.5, e = 0.5, \mu = 0.9$, not only is $e < \bar{e}(\mu)$ satisfied, but also all the assumptions we've made so far: (3.2),(3.8), (3.9) and (3.13).

This numerical result also proves the existence for $\mu > \frac{1}{2}$, the societal optimal. □

B.2 Notation Summary

Symbol	*Support*	*Description*
	Parameters	
e	$[0, \infty)$	the equity amount of the bank
$\rho \equiv \frac{1}{1+e}$	$(0, 1)$	the leverage of the bank
l	$[0, 1)$	liquidation value of the asset
r_1	$(1, r_2)$	the return to the depositor if withdraw in period 1
r_2	$(r_1, 1+\mu)$	the return to the depositor if wait until period 2
μ	$[\frac{r_1}{r_2} - 2\epsilon, 1]$	asset risk
$1-\mu$	$[0, 1 - \frac{r_1}{r_2} + 2\epsilon]$	upside payoff of the asset
λ	$[0, 1]$	fraction of impatient depositors
	Indices	
i	$[0, 1]$	depositors
t	$\{0, 1, 2\}$	time periods
	States of nature	
θ	$[1-\mu, 1]$	asset success probability
ϵ_i	$[-\epsilon, \epsilon]$	the noise of depositor i's signal of θ
θ_i	$[1-\mu-\epsilon, 1+\epsilon]$	the noisy signal of θ for depositor i.
	Decision variables	
n	$[\lambda, 1]$	the fraction of depositors that withdraw
$\dot{n}$	$[\lambda, 1]$	smallest n to trigger $t = 1$ bankruptcy
$\tilde{n}$	$[\lambda, 1]$	smallest n to trigger $t = 2$ bankruptcy
$v()$	$[-1, 1]$	utility differential between rollover and withdraw
$\Delta(\theta_i)$	$[-1, 1]$	expected value of $v()$ given θ_i.
θ^*	$[1-\mu, 1]$	run threshold
$\bar{\theta}$	$\{1 - 3\epsilon\}$	upper dominance bound
$\underline{\theta}$	$\{\frac{r_1}{r_2}\}$	lower dominance bound

Appendix C

CEO evaluation frequencies and innovation

C.1 Figures

Figure C.1. Diagram of governance horizon

Usually corporate governance is interpreted as a mapping from the observable value change of a firm to rewards or punishments to be delivered to the management. "Governance horizon" (a.k.a. results evaluation horizon) is an implicit parameter of this mapping because an evaluation horizon needs to be specified for the computation of the change of the observable firm value . The curve in this diagram represents the exploration-exploitation stages of the innovation process. Short term governance horizon would result in the delivery of punishment due to the lack of short term return for the exploration.

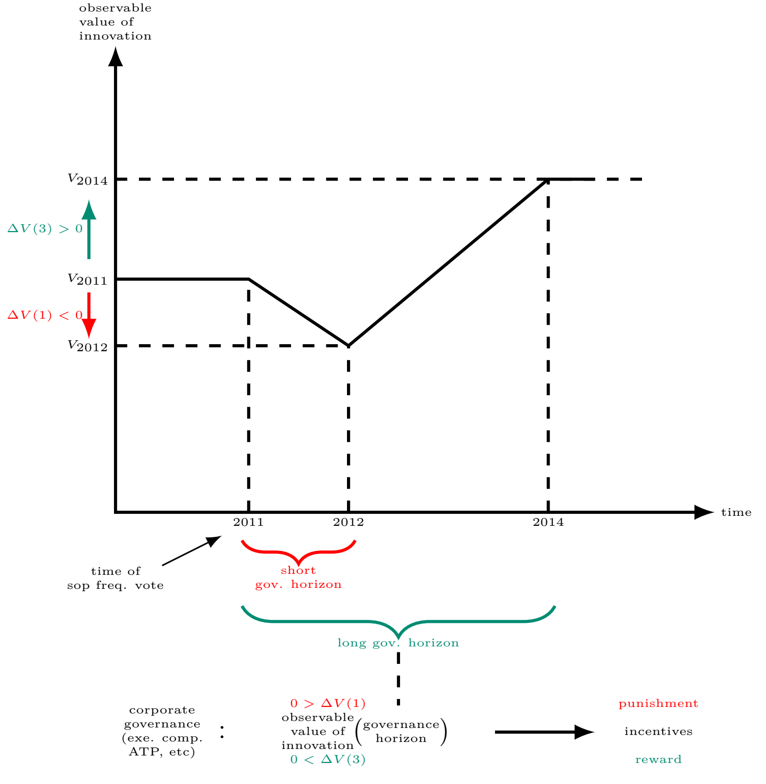


Figure C.2. Truncation problems of patent data - cause

This figure illustrates the cause of truncation problems. The solid boxes and solid arrowhead represent the information available in the patent database, and the dashed boxes and dashed arrowheads represent information only available from the application data. Each arrowhead represents one source of citations. For example, arrowhead B represents the citations made to pre-grant applications from granted patents. The patent data does not allow researchers to count applications for the patent count measure. This is the first type of truncation problem. It does not allow researchers to count sources B, C and D for the citation count measure. This is the second type of truncation problem.

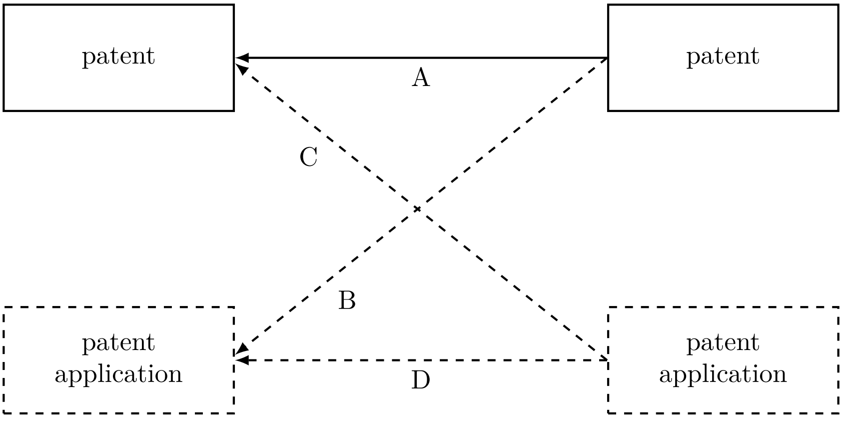
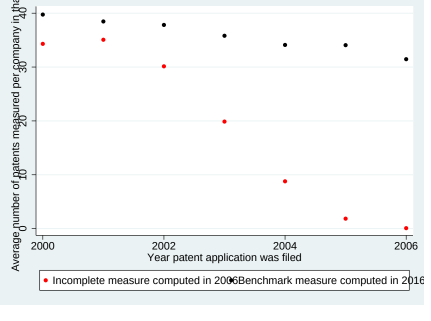


Figure C.3. Truncation problems of patent data - accuracy

This figure illustrates how truncation problems with patent data cause biases of the raw measures for evaluating recent innovations. In both panels, the x-axis represents the year the patent applications were filed. For the left panel, the y-axis represents the number of patents applied in year x that is granted as of 2016 (the black dots), or as of 2006 (the red dots). For the right panel, the y-axis represents the number of citations received by the patents that were filed in year x, and granted as of 2016 (the black dots), or as of 2006 (the red dots). Measures constructed after a ten year wait period suffers little from truncation problems, as most applications in figure C.2 are either granted patent status, or are effectively dormant at that point.(Lerner and Seru (2015)) Thus, the black dots represent the benchmark to which short term measures compete to estimate. Therefore the vertical gap between black dots and red dots represent how much the amount of bias in the raw measure. The figure demonstrates that the bias is substantial for recent innovations (from the perspective of a researcher in 2006), and the bias increases as year x approaches 2006. The detailed procedure of the computation of the raw measure is described in section 4.2

Panel A. A: Log patent count



Panel B. B: Ave. citations

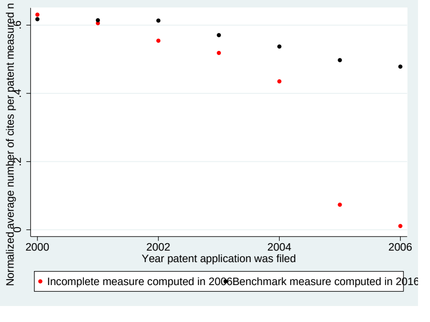


Figure C.4. Truncation problems solved by adding patent application data - accuracy

This figure illustrates how the truncation problem can be solved by including patent application data. The x, y axes, black dots and red dots are defined the same as figure C.3. In both panels, the green dots represent the two measures computed using both patent and application data[1] with the data available as of 2006 (called the "supplemented measures"). The vertical gap between each 2006 measure (red and green) and the benchmark represent how biased each measure is for patents filed in year x.

Panel A. C: Log patent count *Panel B. D: Ave. citations*

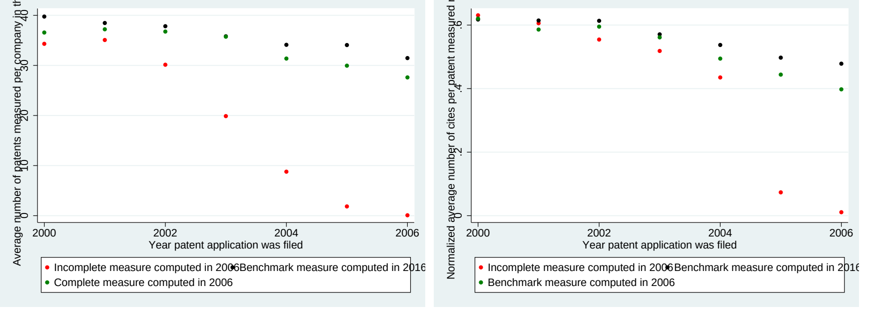


Table C.1. Truncation problems solved by adding patent application data - precision

This table quantifies and contrasts the *precision* of the raw and supplemented measures by their Pearson correlations with the benchmark measure. p-value is reported in the parenthesis.

Measure	**Correlation with the benchmark**					
Patent count						
	2001	2002	2003	2004	2005	2006
Suppl.	0.9944(0.00)	0.9888(0.00)	0.9948 (0.00)	0.9965(0.00)	0.9956(0.00)	0.9947(0.00)
Raw	0.9898(0.00)	0.9709(0.00)	0.9434 (0.00)	0.9051 (0.00)	0.8201 (0.00)	0.6765(0.00)
Average normalized citation count per patent						
	2001	2002	2003	2004	2005	2006
Suppl.	0.8646(0.00)	0.7718(0.00)	0.7304 (0.00)	0.6654 (0.00)	0.6155 (0.00)	0.5758 (0.00)
Raw	0.4533(0.02)	0.3048(0.02)	0.1625 (0.02)	0.0427 (0.08)	0.0046 (0.85)	0.0108 (0.68)

Table C.2. Summary statistics

This table reports the summary statistics for variables constructed based on the sample of U.S. public firms from 2011 to 2015. Panel A reports the summary statistics of innovation measures. Panel B reports the summary statistics of the percentage of votes that are for "once every three years". Panel C presents the year distribution of SOP freq. voting. Panel D reports the industry distribution of firms with shareholder proposals. Panel E reports the summary statistics of firm characteristics that have shown to affect innovations.

Panel A: Innovation measures

Variable	Mean	Median	Std. Dev.	N
Log Pat. Count	2.399	2.079	1.933	1292
Log Cit. Count	4.451	4.304	2.269	1041
Exploi. Index	0.402	0.401	0.228	1094
Cit. diversity	0.247	0.248	0.165	1094

Panel B: Percentage of Votes that are for "once every three years"

Variable	Mean	25^{th}	Median	75^{th}	Passage Rate	N
All SOP freq. votes	31.2%	7.3%	15.6%	54.3%	28.3%	4,884
Those filed patents	29.2%	7.7%	14.7%	48.2%	27.2%	2,247
Close-call SOP freq votes	47.5%	44.1%	49.7%	55.1%	51.5%	249

Panel C: Year Distribution of Shareholder Proposal Voting

Year	All SOP freq. votes	Those filed patents	Close-call SOP freq votes	Ave. firm size Mil.$
2011	3,328	1,673	162	1665
2012	13	8	1	660
2013	929	340	71	181
2014	272	85	9	424
2015	204	75	9	440
2016	138	64	4	NA
Total	4,884	2,247	249	1271

Panel D: Industry distribution of firms that voted on SOP frequency

SIC	Description	All proposals	ATP-related Proposals	Close-call ATP Proposals
0	Agriculture	11	5	1
1	Mining	266	55	3
2	Light manufacturing	668	458	73
3	Heavy manufacturing	976	781	91
4	Transportation	326	115	10
5	Wholesale trade	342	102	4
6	Finance	1,016	148	10
7	Services	489	272	23
8	Health services	162	64	13
9	Public Administration	15	7	3

Panel E: Distribution of firm characteristics

Variable	Mean	Median	Std. Dev.	
Assets (Mil. $)	13162	649.64	93329	1577
R&D/Assets (%)	0.129	0.066	0.204	1258
PPE/Assets (%)	0.170	0.114	0.173	1574
Leverage	0.156	0.086	0.261	1569
CapEx/Assets (%)	0.034	0.023	0.039	1574
Industry Herf.	0.242	0.181	0.190	1579
Tobin's Q	2.962	1.779	8.441	1568
ROA	0.017	0.106	0.345	1589
KZ Index	6.604	-0.120	198.059	1546

Figure C.5. Internal validity - manipulation - density of vote shares

This figure plots the histogram of the distribution of the percentage of votes for "once every three years" SOP in our sample across 40 equally-spaced bins. The x-axis is the percentage of votes favoring a "once every three years" SOP frequency. The y-axis represents the fraction of firms whose shareholders votes in favor of this. Votes for "once every two years" are not counted. SOP frequency voting results are obtained from the ISS.

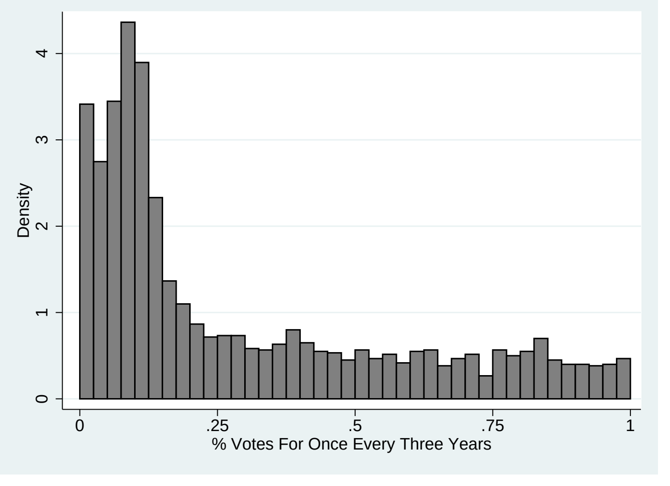


Figure C.6. Internal validity - manipulation - McCrary test

This figure reports the result of the McCrary test (McCrary (2008)) of discontinuity of vote share density at 50%. The x-axis represents the percentage of votes favoring "once every three years" SOP. The y-axis represents the density estimate. The dots depict the density estimate, and the solid line represents the fitted density function of the x (the vote share) with a 95% confidence interval around the fitted line. The votes in favor of "once every two years" are not counted. SOP frequency voting results are obtained from the ISS.

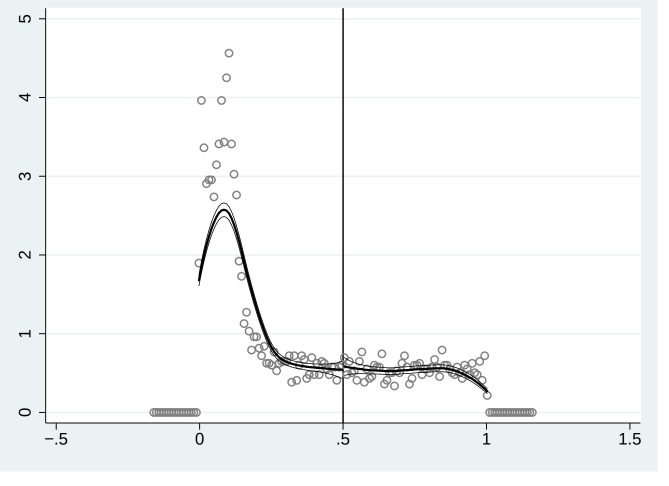


Figure C.7. Internal validity - implementation of voting results

This figure plots the implementation of the SOP frequency voting results. The x-axis is the percentage of votes favoring "once every three years" SOP. The y-axis is the SOP frequency implemented, where "1" represents "once per year" and "3" represents "once every three years". Each dot represents the SOP frequency voting of one firm in my sample. The votes in favor of "once every two years" are not counted. SOP frequency voting results are obtained from the ISS. Implementation data is collected by ISS from SEC 8-K filings.

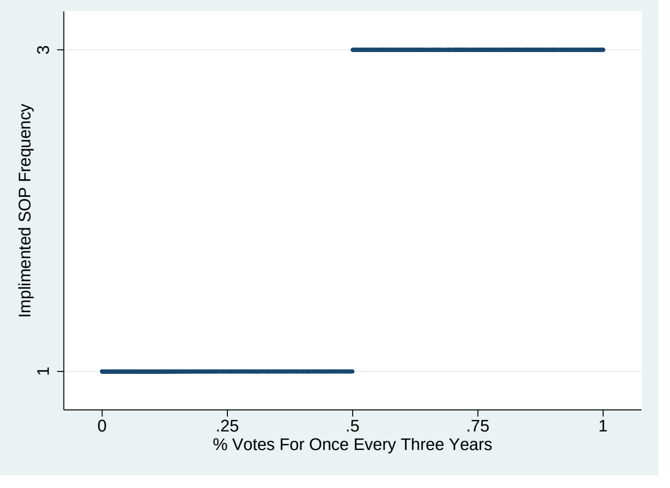


Table C.3. Balance checks - absence of discontinuity of firm characteristics

This table supports the validity of the RDD design by showing that *ex ante* (year 2010) values of firm characteristics do not have discontinuities around the 50% vote share. The variables are defined in the same way as Chemmanur and Tian (2018). I use local polynomial with local correctional term with an optimal bandwidth to estimate the discontinuous jump in the ex-ante outcome, following Imbens and Kalyanaraman (2012). Standard errors are presented in brackets. */**/*** indicates significant at the 10%, 5% and 1% levels, respectively.

	Assets (Mil$)	R&D/Assets	PPE/Assets
β	-1324.5	-0.014	0.003
	[3622.3]	[0.114]	[0.041]
Obs	2308	1683	2298
	Leverage	CAPX/Assets	Industry HHI
β	0.013	0.001	0.031
	[0.037]	[0.008]	[0.049]
Obs	2298	2305	2313
	Tobin's Q	ROA	KZ Index
β	-1.535	-0.034	-25.8
	[1.802]	[0.104]	[26.1]
Obs	2298	2336	2243

Figure C.8. Balance checks - absence of discontinuity of firm characteristics - RD plots

This figure is the RD plot generated using the variables described in table C.3. The x axis represents vote shares in favor of “once every three years”. The y axis represents the density of each vote share. Each dot represents the density averaged over a 0.25% bin. The interval around it is the 95% confidence interval of the mean. The curve is the fitted 4th order polynomial on either side of the cutoff. The lack of significant discontinuity demonstrated by the fact that the fitted line is still within the confidence interval of the bin estimate of the opposite side.

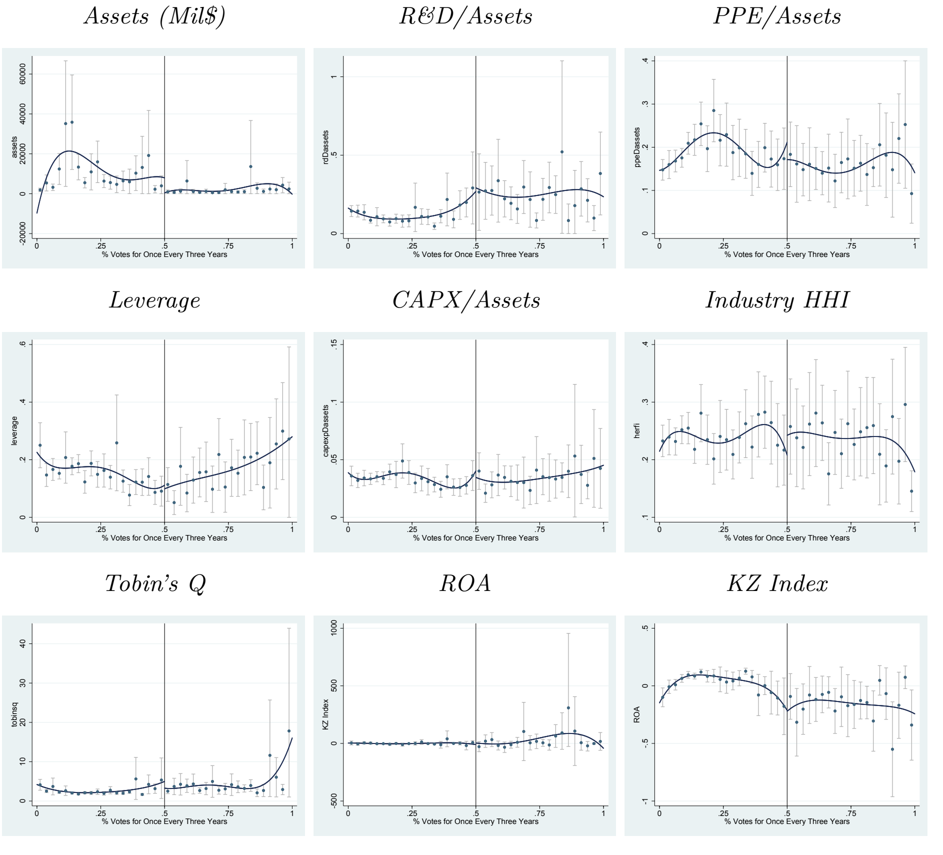


Table C.4. Balance checks - absence of discontinuity of *ex ante* innovation measures

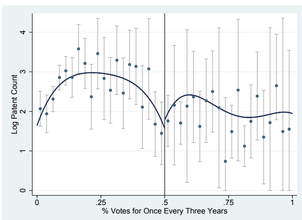
This table supports the validity of the RDD design by showing that *ex ante* (year 2010) values of innovation measures do not have discontinuities around the 50% vote share. I use local polynomial with local correctional term with an optimal bandwidth to estimate the discontinuous jump in the ex-ante outcome, following Imbens and Kalyanaraman (2012). Standard errors are presented in brackets. */**/*** indicates significant at the 10%, 5% and 1% levels, respectively.

	Log Pat. Count	Log Cit. Count	Exploi. Index	Cite Divers.
β	0.640	0.799	-0.095	0.009
	[0.412]	[0.567]	[0.071]	[0.040]
Obs	1073	936	946	946

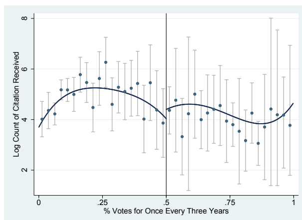
Figure C.9. Balance checks - absence of discontinuity of *ex ante* innovation measures - RD plots

This figure is the RD plot generated using the variables described in table C.4. The x axis represents vote shares in favor of "once every three years". The y axis represents the density of each vote share. Each dot represents the density averaged over a 0.25% bin. The interval around it is the 95% confidence interval of the mean. The curve is the fitted 4th order polynomial on either side of the cutoff. The lack of significant discontinuity demonstrated by the fact that the fitted line is still within the confidence interval of the bin estimate of the opposite side.[2]

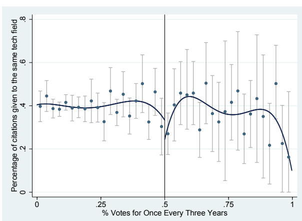
Log Pat. Count



Log Cit. Count



Exploi. Index



Exploi. Index

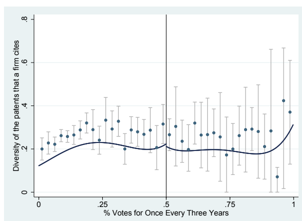


Table C.5. Dynamics of log patent count

This table is generated following the same RD protocol as Table C.3. The measure tested is the log patent count in the first, second and third year after the SOP frequency vote and SOP vote.

	Log Patent Count		
	first year	second year	third year
β	0.550	0.701	1.333**
	[0.447]	[0.591]	[0.528]
Obs	1043	1007	880

Figure C.10. Dynamics of log patent count

This figure is the RD plot generated using the variables described in table C.5. The x axis represents vote shares in favor of "once every three years". The y-axis represents the number of patents that are filed in the first, second and third year after the SOP frequency vote and SOP vote.

Log Patent Count

First Year *Second Year* *Third Year*

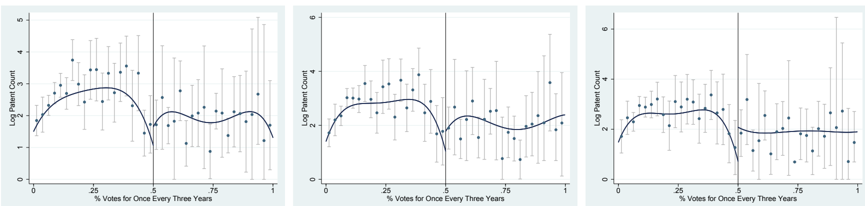


Table C.6. Dynamics of citation count

This table is generated following the same RD measure as table C.3. The measure used is the citation count for patents that are filed in the first, second and third year after the SOP frequency vote and SOP vote.

	Log Count of Citations Received		
	first year	second year	third year
β	0.508	-0.067	0.934*
	[.791]	[0.832]	[0.573]
Obs	848	792	635

Figure C.11. Dynamics of citation count

This figure is the RD plot generated using the variables described in table C.6. The x axis represents vote shares in favor of “once every three years”. The y-axis represents the citation count for patents that are filed in the first, second and third year after the SOP frequency vote and SOP vote.

Log Count of Citations Received

First Year | *Second Year* | *Third Year*

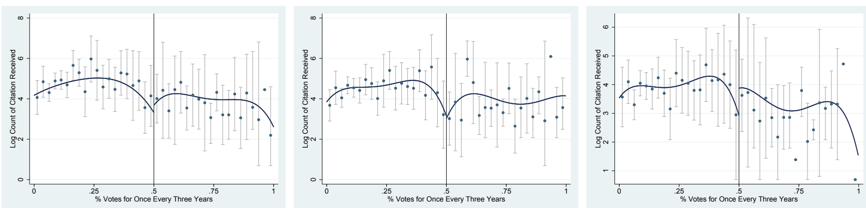


Table C.7. Exploration-exploitation dynamics: Exploitation Index

This figure is the RD plot generated using the protocol described in table C.3. The x axis represents vote shares in favor of “once every three years”. The measure tested is the exploitation index: ratio of the number of citations given to the same tech field by the patents of the firm of interest, divided by the total citations that these patents give. It measures how NOT explorative a firm is in a given year.

	Exploitation Index		
	first year	second year	third year
β	-0.298***	-.142	-0.002
	[0.066]	[0.144]	[0.128]
Obs	896	832	694

Figure C.12. Exploration-exploitation dynamics: Exploitation Index

This figure is the RD plot generated using the variables described in table C.7. The x axis represents vote shares in favor of “once every three years”. The measure tested is the exploitation index: ratio of the number of citations given to the same tech field by the patents of the firm of interest, divided by the total citations they give. It measures how NOT explorative a firm is for a given year.

Exploitation Index

First Year *Second Year* *Third Year*

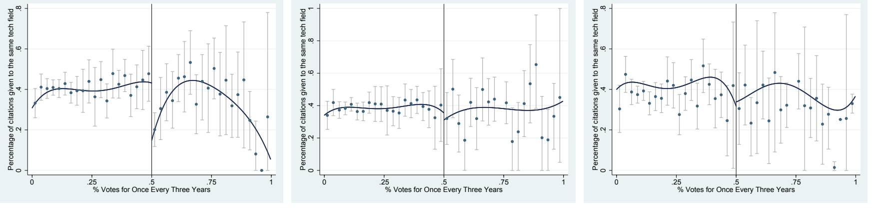


Table C.8. Exploration-exploitation dynamics: Diversity of the Citations Given by the firm of interest

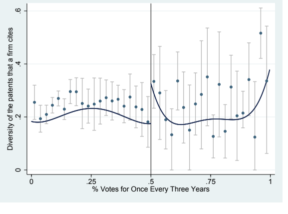
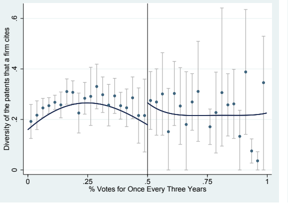
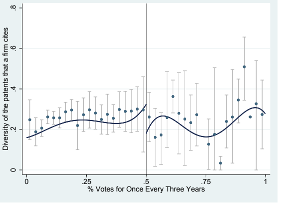
This figure is the RD plot generated using the protocol described in table C.3. The measure tested is the (one minus) Herfendahl index of the technological fields to which a firm's patents give citation. It measures how explorative a firm is in a given year. Then the same test is performed on the time series difference of this variable.

	Citation tech dispersion		
	first year	second year	third year
β	0.186***	0.116	0.018
	[0.062]	[0.090]	[0.077]
Obs	896	832	694
	Diff. of Citation tech dispersion		
	first year	second year	third year
β	0.200***	0.005	-0.003
	[0.054]	[0.063]	[0.070]
Obs	769	735	634

Figure C.13. Exploration-exploitation dynamics: Diversity of the Citations Given by the firm of interest

This figure is the RD plot generated using the variables described in table C.8. The measure tested is the (one minus) Herfendahl index of the technological fields to which a firm's patents give citation. It measures how explorative a firm is for a given year. Then the same test is performed on the time series difference of this variable.

Citation tech dispersion

First Year	*Second Year*	*Third Year*
		

Diff. of Citation tech dispersion

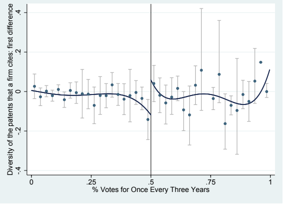
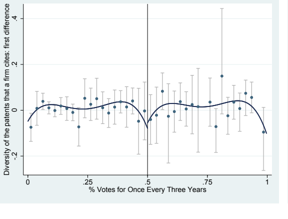
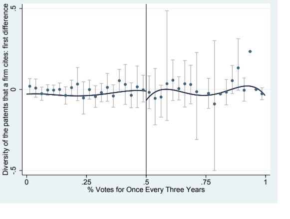
First Year	*Second Year*	*Third Year*
		

Table C.9. Robustness test: fixed bandwidth (5%) for the RDD

This table is the RD results generated during the procedure described in table C.3. For each of two measures of innovation output and the two measures of innovation strategy, instead of using a endogenously generated dynamic bandwidth like above, a fixed bandwidth of 5% was adopted.

	Patent count		
	first year	second year	third year
β	0.480	0.938	1.321**
	[0.471]	[0.699]	[0.632]
Obs	64	49	42
	Patent count		
	first year	second year	third year
β	0.995	1.055	0.848*
	[0.887]	[0.919]	[0.431]
Obs	49	40	35
	Exploitation quotient		
	first year	second year	third year
β	-0.431***	-0.508***	-0.109
	[0.149]	[0.165]	[0.180]
Obs	53	40	39
	Citation field dispersion		
	first year	second year	third year
β	0.209**	0.239**	-0.061
	[0.095]	[0.099]	[0.099]
Obs	53	40	39
	Diff. of Citation tech dispersion		
	first year	second year	third year
β	0.168***	0.082	-0.096*
	[0.061]	[0.062]	[0.057]
Obs	41	39	35

Table C.10. Robustness test: falsification test with cutoff set at 30%

This table is the RD results generated during the procedure described in table C.3. For each of two measures of innovation output and the two measures of innovation strategy, instead of setting the cutoff of the treatment and control at 50%, here we set the cutoff at 30%, a split that is not meaningful, to check if the result is spurious.

		Patent count	
	first year	second year	third year
β	0.710	1.642***	-0.247
	[0.683]	[0.599]	[0.545]
Obs	1043	1007	880
		Patent count	
	first year	second year	third year
β	-0.57	0.150	-0.581
	[0.824]	[0.674]	[0.788]
Obs	848	792	635
		Exploitation quotient	
	first year	second year	third year
β	-0.015	-0.038	-0.058
	[0.053]	[0.055]	[0.066]
Obs	896	832	694
		Citation field dispersion	
	first year	second year	third year
β	0.050	-0.082	-0.060
	[0.056]	[0.050]	[0.039]
Obs	896	832	694
		Diff. of Citation tech dispersion	
	first year	second year	third year
β	0.030	-0.048	0.022
	[0.047]	[0.059]	[0.034]
Obs	769	735	634

References

Acharya, Viral V., Krishnamurthy, Subramanian (2009). *Bankruptcy codes and innovation.*. Review of financial studies, 22(12):4949–4988.

Admati, Anat, Hellwig, Martin (2014). *The bankers' new clothes: What's wrong with banking and what to do about it.*. Princeton University Press.

Aghion, Philippe, Van Reenen, John, Zingales, Luigi. (2013). *Innovation and institutional ownership.*. The American Economic Review, 103(1):277–304.

Aït-Sahalia, Yacine, Saglam, Mehmet (2014). *High frequency traders: Taking advantage of speed*. NBER Working Paper.

Anand, Amber, Venkataraman, Kumar (2015). *Market conditions, fragility and the economics of market making.*. Journal of Financial Economics. Forthcoming.

Atanassov, Julian (2013). *Do hostile takeovers stifle innovation? evidence from antitakeover legislation and corporate patenting.*. The Journal of Finance, 68(3):1097–1131.

Bernstein, Shai (2015). *Does going public affect innovation?*. The Journal of Finance, 70(4):1365–1403.

Bertrand, Marianne, Mullainathan, Sendhil (2003). *Enjoying the quiet life? corporate governance and managerial preferences.*. Journal of political Economy, 111(5):1043–1075.

Biais, Bruno, Declerck, Fany, Moinas, Sophie (2015a). *Who supplies liquidity, how and when?*. Working Paper.

Biais, Bruno, Foucault, Thierry, Moinas, Sophie (2015b). *Equilibrium fast trading*. Journal of Financial Economics, 116(2):292–313.

Brogaard, Jonathan, Hendershott, Terrence, Riordan, Ryan (2014). *High-frequency trading and price discovery*. Review of Financial Studies, 27(8):2267–2306.

Budish, Eric, Cramton, Peter, Shim, John (2015). *The high-frequency trading arms race: Frequent batch auctions as a market design response*. The Quarterly Journal of Economics, 130(4):1621.

Calomiris, Charles W., Kahn, Charles M. (1991). *The role of demandable debt in structuring optimal banking arrangements.*. The American Economic Review, pages 497–513.

Carrion, Allen (2013). *Very fast money: High-frequency trading on the nasdaq*. Journal of Financial Markets, 16(4):680–711.

Cespa, Giovanni, Vives, Xavier (2015). *The beauty contest and short-term trading.*. The Journal of Finance, 70(5):2099–2154.

Chemmanur, Thomas J., Fulghieri, Paolo (2014). *Entrepreneurial finance and innovation: An introduction and agenda for future research.*. Review of Financial Studies, 27(1):1–19.

Chemmanur, Thomas J., Tian, Xuan (2018). *Do antitakeover provisions spur corporate innovation? a regression discontinuity analysis.*. pages 1–32. Journal of Financial and Quantitative Analysis.

Cheng, Ing-Haw, Milbradt, Konstantin (2012). *The hazards of debt: Rollover freezes, incentives, and bailouts.*. The Review of Financial Studies, 25(4):1070–1110.

Cochrane, John (2014). *Toward a run-free financial system.*. Across the Great Divide: New Perspectives on the Financial Crisis.

Constantinides, George M. (1986). *Capital market equilibrium with transaction costs.*. The Journal of Political Economy, pages 842–862.

Cournot, Antoine-Augustin (1838). *Recherches sur les principes mathématiques de la théorie des richesses par Augustin Cournot.*. chez L. Hachette.

Cuñat, Vicente, Gine, Mireia, Guadalupe, Maria (2012). *The vote is cast: The effect of corporate governance on shareholder value.*. The Journal of Finance, 67(5):1943–1977.

DIAMOND, DOUGLAS, DYBVIG, PHILIP (1983). *Bank runs, deposit insurance, and liquidity.*. The journal of political economy, pages 401–419.

DIAMOND, DOUGLAS, RAJAN, RAGHURAM (2000). *A theory of bank capital.*. The Journal of Finance, 55(6):2431–2465.

DIAMOND, DOUGLAS W., RAJAN, RAGHURAM G. (2001). *Banks and liquidity.*. The American Economic Review, 91(2):422–425.

EISENBACH, THOMAS (2017). *Rollover risk as market discipline: A two-sided inefficiency.*. Journal of Financial Economics, 126(2):252–269.

FOUCAULT, THIERRY, KADAN, OHAD, KANDEL, EUGENE (2013). *Liquidity cycles and make/take fees in electronic markets*. The Journal of Finance, 68(1):299–341.

GALSTON, BILL (2015). *Clinton gets it right on short-termism*. The Wall Street Journal.

GOLDSTEIN, ITAY, PAUZNER, ADY (2005). *Demand–deposit contracts and the probability of bank runs.*. Journal of Finance, pages 1293–1327.

GONZALEZ-URIBE, JUANITA, XU, MOQI (2017). *Corporate innovation cycles and ceo contracts.*.

GORTON, GARY, METRICK, ANDREW (2012). *Securitized banking and the run on repo.*. Journal of Financial economics, pages 425–451.

HALL, BRONWYN H., JAFFE, ADAM, TRAJTENBERG, MANUEL (2005). *Market value and patent citations.*. RAND Journal of economics, pages 16–38.

HALL, BRONWYN H., JAFFE, ADAM B., TRAJTENBERG, MANUEL (2001). *The nber patent citation data file: Lessons, insights and methodological tools.*. No. w8498. National Bureau of Economic Research.

HALL, BRONWYN H., LERNER, JOSH (2010). *The financing of r&d and innovation.*. Handbook of the Economics of Innovation, 1:609–639.

HOFFMANN, PETER (2014). *A dynamic limit order market with fast and slow traders*. Journal of Financial Economics, 113(1):156–169.

IMBENS, GUIDO, KALYANARAMAN, KARTHIK (2012). *Optimal bandwidth choice for the regression discontinuity estimator.*. The Review of economic studies.

Imbens, Guido W., Lemieux, Thomas (2008). *Regression discontinuity designs: A guide to practice.*. Journal of econometrics, 142(2):4615–635.

Jaffe, Adam, Manuel, Trajtenberg, Fogarty, Michael (2000). *Knowledge spillovers and patent citations: Evidence from a survey of inventors*. The American Economic Review, 90(2):215–218.

Jovanovic, Boyan, Menkveld, Albert J. (2015). *Middlemen in limit order markets.*. Working Paper.

Kerr, William R., Nanda, Ramana (2014). *Financing innovation.*. National Bureau of Economic Research, No. w20676.

Kerr, William R., Nanda, Ramana, Matthew, Rhodes-Kropf (2014). *Entrepreneurship as experimentation.*. The Journal of Economic Perspectives, 28(3):25–48.

Korajczyk, Robert A., Murphy, Dermot (2015). *High frequency market making to large institutional trades.*. Working Paper.

Lerner, Josh, Seru, Amit (2015). *The use and misuse of patent data: Issues for corporate finance and beyond.*. Booth/Harvard Business School Working Paper.

Malinova, Katya, Park, Andreas, Riordan, Ryan (2013). *Do retail traders suffer from high frequency traders?*. Working Paper.

Mann, William (2018). *Creditor rights and innovation: Evidence from patent collateral.*. Journal of Financial Economics, pages 25–47.

Manso, Gustavo (2011). *Motivating innovation.*. The Journal of Finance, 66(5):1823–1860.

McCrary, Justin (2008). *Manipulation of the running variable in the regression discontinuity design: A density test.*. Journal of Econometrics, 142(2):698–714.

Menkveld, Albert J., Yueshen, Bart Z. (2012). *Middlemen interaction and its effect on market quality*. Working Paper.

Merton, Robert (1969). *Lifetime portfolio selection under uncertainty: The continuous-time case.*. Lifetime portfolio selection under uncertainty: The continuous-time case., pages 247–257.

MORENO, DIEGO, TAKALO, TUOMAS (2016). *Optimal bank transparency.*. Journal of Money, Credit and Banking.

MORRIS, STEPHEN, SHIN, HYUN SONG (2006). *Endogenous public signals and coordination.*. Princeton, Unpublished.

NORDHAUS, WILLIAM D (1989). *Comment*. Brookings Papers on Economic Activity.

PHAM, HUYEN (2009). *Continuous-time stochastic control and optimization with financial applications*. Springer Science and Business Media.

ROBERTS, MICHAEL R., WHITED, TONI M. (2013). *Endogeneity in empirical corporate finance.*. Handbook of the Economics of Finance, 2:493–572.

ROCHET, JEAN-CHARLES, VIVES, XAVIER (2004). *Coordination failures and the lender of last resort: was bagehot right after all?*. Journal of the European Economic Association.

ROE, MARK (2015). *The imaginary problem of corporate short-termism*. The Wall Street Journal.

ROGER, L-C-G, ZANE, OMAR (2002). *A simple model of liquidity effects*. Advances in finance and stochastics, pages 161–176.

ROJCEK, JAKUB, ZIEGLER, ALEXANDRE (2016). *High-frequency trading in limit order markets: Equilibrium impact and regulation*. Working Paper.

SERU, AMIT (2014). *Firm boundaries matter: Evidence from conglomerates and r&d activity.*. Journal of Financial Economics, 111(2):381–405.

STEIN, JEREMY C. (1988). *Takeover threats and managerial myopia.*. The Journal of Political Economy.

TRAJTENBERG, MANUEL, HENDERSON, REBECCA, JAFFE, ADAM (1997). *University versus corporate patents: A window on the basicness of invention.*. Economics of Innovation and new technology, pages 19–50.

WAGNER, WOLF (2009). *Banking fragility and liquidity creation- options as a substitute for deposits*. Annals of Finance, 5(1):125–129.

YERMACK, DAVID (2010). *Shareholder voting and corporate governance.*. Annual Review of Financial Economics, 2(1):103–125.

YUESHEN, BART ZHOU (2014). *Queuing uncertainty in limit order market*. Working paper.

About the Author

Lingtian Kong was born in China. With a passion in both the natural and social sciences, he became interested in economics as a way to analyze societal phenomena by research tools shared with the natural sciences. He obtained his Master of Science degree in Finance and Economics at the London School of Economics in 2012.

In 2012, Lingtian started his doctoral studies at Rotterdam School of Management, Erasmus University under the supervision of Professor Marno Verbeek, Doctor Dion Bongaerts and Doctor Mark Van Achter. His PhD project focuses on theoretical and empirical analysis of the mal-function of the financial market. He took various courses offered at RSM as well as other institutions and had teaching experience with both graduate and undergraduate students. In the third year of his PhD program, Lingtian went on a research visit to the Haas School of Business at the University of California, Berkeley, hosted by Professor Terrence Hendershott. He presented his research work at several international conferences and research seminars.

Portfolio

Working papers

CEO evaluation frequencies and innovation. *(single authored)*

Trading speed competition: Can the arms race go too far? *(joint with Dion Bongaerts and Mark Van Achter)*

Asset risk and bank runs. *(joint with Dion Bongaerts)*

Conference presentations (* presented by co-authors)

2017: Financial Management Association Annual Meeting (Boston); International Finance and Banking Society Meeting (Ningbo); Annual Conference on Innovation Economics (Chicago).

2016: Stern Microstructure Conference* (the New York City); Tinbergen Institute (Ph.D. Lunch Seminar Series); Erasmus University (Ph.D. Seminars).

2015: European Finance Association Annual Meeting (Vienna)*; International Finance and Banking Society Meeting (Hangzhou); CEMS workshop on Market Liquidity (Brussels); Erasmus University (Ph.D. Seminars).

2014: Erasmus University (Ph.D. Seminars).

Research visits & Ph.D. courses

Research visit at UC Berkeley Haas School of Business, Berkeley (2015)

Seminar Corporate Finance, Erasmus University
Seminar Asset Pricing, Erasmus University
Behavioral Decision Theory, Erasmus University
Empirical Corporate Finance, Erasmus University
Empirical Asset Pricing, Erasmus University
Market Microstructure, Tinbergen Institute

Teaching activities

Master Thesis Supervision (2013-2017)
Bachelor Thesis Supervision (2013-2014, 2016)

Scholarships & grant

RSM Trustfund Research Visit Grant (2015)
AFA Student Travel Grant (2015)
Standard Bank Scholarship (2011-2012)

The ERIM PhD Series

The ERIM PhD Series contains PhD dissertations in the field of Research in Management defended at Erasmus University Rotterdam and supervised by senior researchers affiliated to the Erasmus Research Institute of Management (ERIM). All dissertations in the ERIM PhD Series are available in full text through the ERIM Electronic Series Portal: http://repub.eur.nl/pub. ERIM is the joint research institute of the Rotterdam School of Management (RSM) and the Erasmus School of Economics at the Erasmus University Rotterdam (EUR).

DISSERTATIONS LAST FIVE YEARS

Akemu, O., *Corporate Responses to Social Issues: Essays in Social Entrepreneurship and Corporate Social Responsibility*, Promotors: Prof. G.M. Whiteman & Dr S.P. Kennedy, EPS-2017-392-ORG, https://repub.eur.nl/pub/95768

Alexiou, A. *Management of Emerging Technologies and the Learning Organization: Lessons from the Cloud and Serious Games Technology*, Promotors: Prof. S.J. Magala, Prof. M.C. Schippers and Dr I. Oshri, EPS-2016-404-ORG, http://repub.eur.nl/pub/93818

Alserda, G.A.G., *Choices in Pension Management*, Promotors: Prof. S.G. van der Lecq & Dr O.W. Steenbeek, EPS-2017-432-F&A, https://repub.eur.nl/pub/103496

Arampatzi, E., *Subjective Well-Being in Times of Crises: Evidence on the Wider Impact of Economic Crises and Turmoil on Subjective Well-Being*, Promotors: Prof. H.R. Commandeur, Prof. F. van Oort & Dr. M.J. Burger, EPS-2018-459-S&E, https://repub.eur.nl/pub/111830

Avci, E., *Surveillance of Complex Auction Markets: a Market Policy Analytics Approach*, Promotors: Prof. W. Ketter, Prof. H.W.G.M. van Heck & Prof. D.W. Bunn, EPS-2018-426-LIS, https://repub.eur.nl/pub/106286

Benschop, N, *Biases in Project Escalation: Names, frames & construal levels*, Promotors: Prof. K.I.M. Rhode, Prof. H.R. Commandeur, Prof. M. Keil & Dr A.L.P. Nuijten, EPS-2015-375-S&E, http://repub.eur.nl/pub/79408

Beusichem, H.C. van, *Firms and Financial Markets: Empirical Studies on the Informational Value of Dividends, Governance and Financial Reporting*, Promotors: Prof. A. de Jong & Dr G. Westerhuis, EPS-2016-378-F&A, http://repub.eur.nl/pub/93079

Bliek, R. de, *Empirical Studies on the Economic Impact of Trust*, Promotor: Prof. J. Veenman & Prof. Ph.H.B.F. Franses, EPS-2015-324-ORG, http://repub.eur.nl/pub/78159

Bouman, P., *Passengers, Crowding and Complexity: Models for Passenger Oriented Public Transport*, Prof. L.G. Kroon, Prof. A. Sch?bel & Prof. P.H.M. Vervest, EPS-2017-420-LIS, https://repub.eur.nl/

Brazys, J., *Aggregated Marcoeconomic News and Price Discovery*, Promotor: Prof. W.F.C. Verschoor, EPS-2015-351-F&A, http://repub.eur.nl/pub/78243

Bunderen, L. van, *Tug-of-War: Why and when teams get embroiled in power struggles*, Promotors: Prof. D.L. van Knippenberg & Dr. L. Greer, EPS-2018-446-ORG, https://repub.eur.nl/pub/105346

Burg, G.J.J. van den, *Algorithms for Multiclass Classification and Regularized Regression*, Promotors: Prof. P.J.F. Groenen & Dr. A. Alfons, EPS-2018-442-MKT, https://repub.eur.nl/pub/103929

Chammas, G., *Portfolio concentration*, Promotor: Prof. J. Spronk, EPS-2017-410-F&E,https://repub.eur.nl/pub/94975

Cranenburgh, K.C. van, *Money or Ethics: Multinational corporations and religious organisations operating in an era of corporate responsibility*, Prof. L.C.P.M. Meijs, Prof. R.J.M. van Tulder & Dr D. Arenas, EPS-2016-385-ORG, http://repub.eur.nl/pub/93104

Consiglio, I., *Others: Essays on Interpersonal and Consumer Behavior*, Promotor: Prof. S.M.J. van Osselaer, EPS-2016-366-MKT, http://repub.eur.nl/pub/79820

Darnihamedani, P. *Individual Characteristics, Contextual Factors and Entrepreneurial Behavior*, Promotors: Prof. A.R. Thurik & S.J.A. Hessels, EPS-2016-360-S&E, http://repub.eur.nl/pub/93280

Dennerlein, T. *Empowering Leadership and Employees' Achievement Motivations: the Role of Self-Efficacy and Goal Orientations in the Empowering Leadership Process*, Promotors: Prof. D.L. van Knippenberg & Dr J. Dietz, EPS-2017-414-ORG, https://repub.eur.nl/pub/98438

Deng, W., *Social Capital and Diversification of Cooperatives*, Promotor: Prof. G.W.J. Hendrikse, EPS-2015-341-ORG, http://repub.eur.nl/pub/77449

Depecik, B.E., *Revitalizing brands and brand: Essays on Brand and Brand Portfolio Management Strategies*, Promotors: Prof. G.H. van Bruggen, Dr Y.M. van Everdingen and Dr M.B. Ataman, EPS-2016-406-MKT, http://repub.eur.nl/pub/93507

Duijzer, L.E., *Mathematical Optimization in Vaccine Allocation*, Promotors: Prof. R. Dekker & Dr W.L. van Jaarsveld, EPS-2017-430-LIS, https://repub.eur.nl/pub/101487

Duyvesteyn, J.G. *Empirical Studies on Sovereign Fixed Income Markets*, Promotors: Prof. P. Verwijmeren & Prof. M.P.E. Martens, EPS-2015-361-F&A, https://repub.eur.nl/pub/79033

Elemes, A., *Studies on Determinants and Consequences of Financial Reporting Quality*, Promotor: Prof. E. Peek, EPS-2015-354-F&A, https://repub.eur.nl/pub/79037

Ellen, S. ter, *Measurement, Dynamics, and Implications of Heterogeneous Beliefs in Financial Markets*, Promotor: Prof. W.F.C. Verschoor, EPS-2015-343-F&A, http://repub.eur.nl/pub/78191

Erlemann, C., *Gender and Leadership Aspiration: The Impact of the Organizational Environment*, Promotor: Prof. D.L. van Knippenberg, EPS-2016-376-ORG, http://repub.eur.nl/pub/79409

Eskenazi, P.I., *The Accountable Animal*, Promotor: Prof. F.G.H. Hartmann, EPS-2015-355-F&A, http://repub.eur.nl/pub/78300

Evangelidis, I., *Preference Construction under Prominence*, Promotor: Prof. S.M.J. van Osselaer, EPS-2015-340-MKT, http://repub.eur.nl/pub/78202 Faber, N., *Structuring Warehouse Management*, Promotors: Prof. M.B.M. de Koster & Prof. A. Smidts, EPS-2015-336-LIS, http://repub.eur.nl/pub/78603

Feng, Y., *The Effectiveness of Corporate Governance Mechanisms and Leadership Structure: Impacts on strategic change and firm performance*, Promotors: Prof. F.A.J. van den Bosch, Prof. H.W. Volberda & Dr J.S. Sidhu, EPS-2017-389-S&E, https://repub.eur.nl/pub/98470

Fernald, K., *The Waves of Biotechnological Innovation in Medicine: Interfirm Cooperation Effects and a Venture Capital Perspective*, Promotors: Prof. E. Claassen, Prof. H.P.G. Pennings & Prof. H.R. Commandeur, EPS-2015-371-S&E, http://hdl.handle.net/1765/79120

Fisch, C.O., *Patents and trademarks: Motivations, antecedents, and value in industrialized and emerging markets*, Promotors: Prof. J.H. Block, Prof. H.P.G. Pennings & Prof. A.R. Thurik, EPS-2016-397-S&E, http://repub.eur.nl/pub/94036

Fliers, P.T., *Essays on Financing and Performance: The role of firms, banks and board*, Promotors: Prof. A. de Jong & Prof. P.G.J. Roosenboom, EPS-2016-388-F&A, http://repub.eur.nl/pub/93019

Frick, T.W., *The Implications of Advertising Personalization for Firms, Consumer, and Ad Platfroms*, Promotors: Prof. T. Li & Prof. H.W.G.M. van Heck, EPS-2018-452-LIS, https://repub.eur.nl/pub/110314

Fytraki, A.T., *Behavioral Effects in Consumer Evaluations of Recommendation Systems*, Promotors: Prof. B.G.C. Dellaert & Prof. T. Li, EPS-2018-427-MKT, https://repub.eur.nl/pub/110457

Gaast, J.P. van der, *Stochastic Models for Order Picking Systems*, Promotors: Prof. M.B.M de Koster & Prof. I.J.B.F. Adan, EPS-2016-398-LIS, http://repub.eur.nl/pub/93222

Giurge, L., *A Test of Time; A temporal and dynamic approach to power and ethics*, Promotors: Prof. M.H. van Dijke & Prof. D. De Cremer, EPS-2017-412-ORG, https://repub.eur.nl/

Gobena, L., *Towards Integrating Antecedents of Voluntary Tax Compliance*, Promotors: Prof. M.H. van Dijke & Dr P. Verboon, EPS-2017-436-ORG, https://repub.eur.nl/pub/103276

Groot, W.A., *Assessing Asset Pricing Anomalies*, Promotors: Prof. M.J.C.M. Verbeek & Prof. J.H. van Binsbergen, EPS-2017-437-F&A, https://repub.eur.nl/pub/103490

Harms, J. A., *Essays on the Behavioral Economics of Social Preferences and Bounded Rationality*, Prof. H.R. Commandeur & Dr K.E.H. Maas, EPS-2018-457-S&E, https://repub.eur.nl/pub/108831

Hekimoglu, M., *Spare Parts Management of Aging Capital Products*, Promotor: Prof. R. Dekker, EPS-2015-368-LIS, http://repub.eur.nl/pub/79092

Hengelaar, G.A., *The Proactive Incumbent: Holy grail or hidden gem? Investigating whether the Dutch electricity sector can overcome the incumbent's curse and lead the sustainability transition*, Promotors: Prof. R.J. M. van Tulder & Dr K. Dittrich, EPS-2018-438-ORG, https://repub.eur.nl/pub/102953

Hogenboom, A.C., *Sentiment Analysis of Text Guided by Semantics and Structure*, Promotors: Prof. U. Kaymak & Prof. F.M.G. de Jong, EPS-2015-369-LIS, http://repub.eur.nl/pub/79034

Hollen, R.M.A., *Exploratory Studies into Strategies to Enhance Innovation-Driven International Competitiveness in a Port Context: Toward Ambidextrous Ports*, Promotors: Prof. F.A.J. Van Den Bosch & Prof. H.W. Volberda, EPS-2015-372-S&E, http://repub.eur.nl/pub/78881

Hurk, E. van der, *Passengers, Information, and Disruptions*, Promotors: Prof. L.G. Kroon & Prof. P.H.M. Vervest, EPS-2015-345-LIS, http://repub.eur.nl/pub/78275

Jacobs, B.J.D., *Marketing Analytics for High-Dimensional Assortments*, Promotors: Prof. A.C.D. Donkers & Prof. D. Fok, EPS-2017-445-MKT, https://repub.eur.nl/pub/103497

Kahlen, M. T., *Virtual Power Plants of Electric Vehicles in Sustainable Smart Electricity Markets*, Promotors: Prof. W. Ketter & Prof. A. Gupta, EPS-2017-431-LIS, https://repub.eur.nl/pub/100844

Kampen, S. van, *The Cross-sectional and Time-series Dynamics of Corporate Finance: Empirical evidence from financially constrained firms*, Promotors: Prof. L. Norden & Prof. P.G.J. Roosenboom, EPS-2018-440-F&A, https://repub.eur.nl/pub/105245

Karali, E., *Investigating Routines and Dynamic Capabilities for Change and Innovation*, Promotors: Prof. H.W. Volberda, Prof. H.R. Commandeur and Dr J.S. Sidhu, EPS-2018-454-S&E, https://repub.eur.nl/pub/106274

Keko. E, *Essays on Innovation Generation in Incumbent Firms*, Promotors: Prof. S. Stremersch & Dr N.M.A. Camacho, EPS-2017-419-MKT, https://repub.eur.nl/pub/100841

Kerkkamp, R.B.O., *Optimisation Models for Supply Chain Coordination under Information Asymmetry*, Promotors: Prof. A.P.M. Wagelmans & Dr. W. van den Heuvel, EPS-2018-462-LIS

Khattab, J., *Make Minorities Great Again: a contribution to workplace equity by identifying and addressing constraints and privileges*, Promotors: Prof. D.L. van Knippenberg & Dr A. Nederveen Pieterse, EPS-2017-421-ORG, https://repub.eur.nl/pub/99311

Kim, T. Y., *Data-driven Warehouse Management in Global Supply Chains*, Promotors: Prof. R. Dekker & Dr C. Heij, EPS-2018-449-LIS, https://repub.eur.nl/pub/109103

Klitsie, E.J., *Strategic Renewal in Institutional Contexts: The paradox of embedded agency*, Promotors: Prof. H.W. Volberda & Dr. S. Ansari, EPS-2018-444-S&E, https://repub.eur.nl/pub/106275

Krämer, R., *A license to mine? Community organizing against multinational corporations*, Promotors: Prof. R.J.M. van Tulder & Prof. G.M. Whiteman, EPS-2016-383-ORG, http://repub.eur.nl/pub/94072

Kysucky, V., *Access to Finance in a Cros-Country Context*, Promotor: Prof. L. Norden, EPS-2015-350-F&A, http://repub.eur.nl/pub/78225

Lee, C.I.S.G., *Big Data in Management Research: Exploring New Avenues*, Promotors: Prof. S.J. Magala & Dr W.A. Felps, EPS-2016-365-ORG, http://repub.eur.nl/pub/79818

Legault-Tremblay, P.O., *Corporate Governance During Market Transition: Heterogeneous responses to Institution Tensions in China*, Promotor: Prof. B. Krug, EPS-2015-362-ORG, http://repub.eur.nl/pub/78649

Lenoir, A.S. *Are You Talking to Me? Addressing Consumers in a Globalised World*, Promotors: Prof. S. Puntoni & Prof. S.M.J. van Osselaer, EPS-2015-363-MKT, http://repub.eur.nl/pub/79036

Li, D., *Supply Chain Contracting for After-sales Service and Product Support*, Promotor: Prof. M.B.M. de Koster, EPS-2015-347-LIS, http://repub.eur.nl/pub/78526

Liu, N., *Behavioral Biases in Interpersonal Contexts*, Supervisors: Prof. A. Baillon & Prof. H. Bleichrodt, EPS-2017-408-MKT, https://repub.eur.nl/pub/95487

Ma, Y., *The Use of Advanced Transportation Monitoring Data for Official Statistics*, Promotors: Prof. L.G. Kroon & Dr J. van Dalen, EPS-2016-391-LIS, http://repub.eur.nl/pub/80174

Maira, E., *Consumers and Producers*, Promotors: Prof. S. Puntoni & Prof. C. Fuchs, EPS-2018-439-MKT, https://repub.eur.nl/pub/104387

Mell, J.N., *Connecting Minds: On The Role of Metaknowledge in Knowledge Coordination*, Promotor: Prof. D.L. van Knippenberg, EPS-2015-359-ORG, http://hdl.handle.net/1765/78951

Meulen,van der, D., *The Distance Dilemma: the effect of flexible working practices on performance in the digital workplace*, Promotors: Prof. H.W.G.M. van Heck & Prof. P.J. van Baalen, EPS-2016-403-LIS, http://repub.eur.nl/pub/94033

Micheli, M.R., *Business Model Innovation: A Journey across Managers' Attention and Inter-Organizational Networks*, Promotor: Prof. J.J.P. Jansen, EPS-2015-344-S&E,http://repub.eur.nl/pub/78241

Moniz, A, *Textual Analysis of Intangible Information*, Promotors: Prof. C.B.M. van Riel, Prof. F.M.G de Jong & Dr G.A.J.M. Berens, EPS-2016-393-ORG, http://repub.eur.nl/pub/93001

Mulder, J. *Network design and robust scheduling in liner shipping*, Promotors: Prof. R. Dekker & Dr W.L. van Jaarsveld, EPS-2016-384-LIS, http://repub.eur.nl/pub/80258

Neerijnen, P., *The Adaptive Organization: the socio-cognitive antecedents of ambidexterity and individual exploration*, Promotors: Prof. J.J.P. Jansen, P.P.M.A.R. Heugens & Dr T.J.M. Mom, EPS-2016-358-S&E, http://repub.eur.nl/pub/93274

Okbay, A., *Essays on Genetics and the Social Sciences*, Promotors: Prof. A.R. Thurik, Prof. Ph.D. Koellinger & Prof. P.J.F. Groenen, EPS-2017-413-S&E, https://repub.eur.nl/pub/95489

Oord, J.A. van, *Essays on Momentum Strategies in Finance*, Promotor: Prof. H.K. van Dijk, EPS-2016-380-F&A, http://repub.eur.nl/pub/80036

Peng, X., *Innovation, Member Sorting, and Evaluation of Agricultural Cooperatives*, Promotor: Prof. G.W.J. Hendriks, EPS-2017-409-ORG, https://repub.eur.nl/pub/94976

Pennings, C.L.P., *Advancements in Demand Forecasting: Methods and Behavior*, Promotors: Prof. L.G. Kroon, Prof. H.W.G.M. van Heck & Dr J. van Dalen, EPS-2016-400-LIS, http://repub.eur.nl/pub/94039

Petruchenya, A., *Essays on Cooperatives: Emergence, Retained Earnings, and Market Shares*, Promotors: Prof. G.W.J. Hendriks & Dr Y. Zhang, EPS-2018-447-ORG, https://repub.eur.nl/pub/105243

Plessis, C. du, *Influencers: The Role of Social Influence in Marketing*, Promotors: Prof. S. Puntoni & Prof. S.T.L.R. Sweldens, EPS-2017-425-MKT, https://repub.eur.nl/pub/103265

Pocock, M., *Status Inequalities in Business Exchange Relations in Luxury Markets*, Promotors: Prof. C.B.M. van Riel & Dr G.A.J.M. Berens, EPS-2017-346-ORG, https://repub.eur.nl/pub/98647

Pozharliev, R., *Social Neuromarketing: The role of social context in measuring advertising effectiveness*, Promotors: Prof. W.J.M.I. Verbeke & Prof. J.W. van Strien, EPS-2017-402-MKT, https://repub.eur.nl/pub/95528

Protzner, S. *Mind the gap between demand and supply: A behavioral perspective on demand forecasting*, Promotors: Prof. S.L. van de Velde & Dr L. Rook, EPS-2015-364-LIS, http://repub.eur.nl/pub/79355

Pruijssers, J.K., *An Organizational Perspective on Auditor Conduct*, Promotors: Prof. J. van Oosterhout & Prof. P.P.M.A.R. Heugens, EPS-2015-342-S&E, http://repub.eur.nl/pub/78192

Riessen, B. van, *Optimal Transportation Plans and Portfolios for Synchromodal Container Networks*, Promotors: Prof. R. Dekker & Prof. R.R. Negenborn, EPS-2018-448-LIS, https://repub.eur.nl/pub/105248

Rietdijk, W.J.R. *The Use of Cognitive Factors for Explaining Entrepreneurship*, Promotors: Prof. A.R. Thurik & Prof. I.H.A. Franken, EPS-2015-356-S&E, http://repub.eur.nl/pub/79817

Rösch, D. *Market Efficiency and Liquidity*, Promotor: Prof. M.A. van Dijk, EPS-2015-353-F&A, http://repub.eur.nl/pub/79121

Roza, L., *Employee Engagement in Corporate Social Responsibility: A collection of essays*, Promotor: Prof. L.C.P.M. Meijs, EPS-2016-396-ORG, http://repub.eur.nl/pub/93254

Schie, R. J. G. van, *Planning for Retirement: Save More or Retire Later?* Promotors: Prof. B. G. C. Dellaert & Prof. A.C.D. Donkers, EOS-2017-415-MKT, https://repub.eur.nl/pub/100846

Schoonees, P. *Methods for Modelling Response Styles*, Promotor: Prof. P.J.F. Groenen, EPS-2015-348-MKT, http://repub.eur.nl/pub/79327

Schouten, K.I.M. *Semantics-driven Aspect-based Sentiment Analysis*, Promotors: Prof. F.M.G. de Jong, Prof. R. Dekker & Dr. F. Frasincar, EPS-2018-453-LIS, https://repub.eur.nl/pub/112161

Schouten, M.E., *The Ups and Downs of Hierarchy: the causes and consequences of hierarchy struggles and positional loss*, Promotors; Prof. D.L. van Knippenberg & Dr L.L. Greer, EPS-2016-386-ORG, http://repub.eur.nl/pub/80059

Smit, J. *Unlocking Business Model Innovation: A look through the keyhole at the inner workings of Business Model Innovation*, Promotor: Prof. H.G. Barkema, EPS-2016-399-S&E, http://repub.eur.nl/pub/93211

Straeter, L.M., *Interpersonal Consumer Decision Making*, Promotors: Prof. S.M.J. van Osselaer & Dr I.E. de Hooge, EPS-2017-423-MKT, https://repub.eur.nl/pub/100819

Stuppy, A., *Essays on Product Quality*, Promotors: Prof. S.M.J. van Osselaer & Dr. N.L. Mead. EPS-2018-461-MKT, https://repub.eur.nl/pub/111375

Subaşi, B., *Demographic Dissimilarity, Information Access and Individual Performance*, Promotors: Prof. D.L. van Knippenberg & Dr W.P. van Ginkel, EPS-2017-422-ORG, https://repub.eur.nl/pub/103495

Szatmari, B., *We are (all) the champions: The effect of status in the implementation of innovations*, Promotors: Prof. J.C.M & Dr D. Deichmann, EPS-2016-401-LIS, http://repub.eur.nl/pub/94633

Tuijl, E. van, *Upgrading across Organisational and Geographical Configurations*, Promotor: Prof. L. van den Berg, EPS-2015-349-S&E, http://repub.eur.nl/pub/78224

Turturea, R., *Overcoming Resource Constraints: The Role of Creative Resourcing and Equity Crowdfunding in Financing Entrepreneurial Ventures*, Promotors: Prof. P.P.M.A.R Heugens, Prof. J.J.P. Jansen & Dr. I. Verheuil, EPS-2019-472-S&E, https://repub.eur.nl/pub/112859

Valogianni, K. *Sustainable Electric Vehicle Management using Coordinated Machine Learning*, Promotors: Prof. H.W.G.M. van Heck & Prof. W. Ketter, EPS-2016-387-LIS, http://repub.eur.nl/pub/93018

Vandic, D., *Intelligent Information Systems for Web Product Search*, Promotors: Prof. U. Kaymak & Dr Frasincar, EPS-2017-405-LIS, https://repub.eur.nl/pub/95490

Verbeek, R.W.M., *Essays on Empirical Asset Pricing*, Promotors: Prof. M.A. van Dijk & Dr M. Szymanowska, EPS-2017-441-F&A, https://repub.eur.nl/pub/102977

Vermeer, W., *Propagation in Networks:The impact of information processing at the actor level on system-wide propagation dynamics*, Promotor: Prof. P.H.M.Vervest, EPS-2015-373-LIS, http://repub.eur.nl/pub/79325

Versluis, I., *Prevention of the Portion Size Effect*, Promotors: Prof. Ph.H.B.F. Franses & Dr E.K. Papies, EPS-2016-382-MKT, http://repub.eur.nl/pub/79880

Vishwanathan, P., *Governing for Stakeholders: How Organizations May Create or Destroy Value for their Stakeholders*, Promotors: Prof. J. van Oosterhout & Prof. L.C.P.M. Meijs, EPS-2016-377-ORG, http://repub.eur.nl/pub/93016

Vlaming, R. de., *Linear Mixed Models in Statistical Genetics*, Prof. A.R. Thurik,Prof. P.J.F. Groenen & Prof. Ph.D. Koellinger, EPS-2017-416-S&E, https://repub.eur.nl/pub/100428

Vries, H. de, *Evidence-Based Optimization in Humanitarian Logistics*, Promotors: Prof. A.P.M. Wagelmans & Prof. J.J. van de Klundert, EPS-2017-435-LIS, https://repub.eur.nl/pub/102771

Vries, J. de, *Behavioral Operations in Logistics*, Promotors: Prof. M.B.M de Koster & Prof. D.A. Stam, EPS-2015-374-LIS, http://repub.eur.nl/pub/79705

Wagenaar, J.C., *Practice Oriented Algorithmic Disruption Management in Passenger Railways*, Prof. L.G. Kroon & Prof. A.P.M. Wagelmans, EPS-2016-390-LIS, http://repub.eur.nl/pub/93177

Wang, P., *Innovations, status, and networks*, Promotors: Prof. J.J.P. Jansen & Dr V.J.A. van de Vrande, EPS-2016-381-S&E, http://repub.eur.nl/pub/93176

Wang, R., *Corporate Environmentalism in China*, Promotors: Prof. P.P.M.A.R Heugens & Dr F. Wijen, EPS-2017-417-S&E, https://repub.eur.nl/pub/99987

Wang, T., *Essays in Banking and Corporate Finance*, Promotors: Prof. L. Norden& Prof. P.G.J. Roosenboom, EPS-2015-352-F&A, http://repub.eur.nl/pub/78301

Wasesa, M., *Agent-based inter-organizational systems in advanced logistics operations*, Promotors: Prof. H.W.G.M van Heck, Prof. R.A. Zuidwijk & Dr A. W. Stam, EPS-2017-LIS-424, https://repub.eur.nl/pub/100527

Wessels, C., *Flexible Working Practices: How Employees Can Reap the Benefits for Engagement and Performance*, Promotors: Prof. H.W.G.M. van Heck, Prof. P.J. van Baalen & Prof. M.C. Schippers, EPS-2017-418-LIS, https://repub.eur.nl/

Williams, A.N., *Make Our Planet Great Again: A Systems Perspective of Corporate Sustainability*, Promotors: Prof. G.M. Whiteman & Dr. S. Kennedy, EPS-2018-456-ORG, https://repub.eur.nl/pub/111032

Witte, C.T., *Bloody Business: Multinational investment in an increasingly conflict-afflicted world*, Promotors: Prof. H.P.G. Pennings, Prof. H.R. Commandeur & Dr M.J. Burger, EPS-2018-443-S&E, https://repub.eur.nl/pub/104027

Yuan, Y., *The Emergence of Team Creativity: a social network perspective*, Promotors: Prof. D. L. van Knippenberg & Dr D. A. Stam, EPS-2017-434-ORG, https://repub.eur.nl/pub/100847

Ypsilantis, P., *The Design, Planning and Execution of Sustainable Intermodal Port-hinterland Transport Networks*, Promotors: Prof. R.A. Zuidwijk & Prof. L.G. Kroon, EPS-2016-395-LIS, http://repub.eur.nl/pub/94375

Yuferova, D. *Price Discovery, Liquidity Provision, and Low-Latency Trading*, Promotors: Prof. M.A. van Dijk & Dr D.G.J. Bongaerts, EPS-2016-379-F&A, http://repub.eur.nl/pub/93017

Zhang, Q., *Financing and Regulatory Frictions in Mergers and Acquisitions*, Promotors: Prof. P.G.J. Roosenboom & Prof. A. de Jong, EPS-2018-428-F&A, https://repub.eur.nl/pub/103871

Zuber, F.B., *Looking at the Others: Studies on (un)ethical behavior and social relationships in organizations*, Promotor: Prof. S.P. Kaptein, EPS-2016-394-ORG, http://repub.eur.nl/pub/94388