Essays on competition in banking

van Boxtel, A.A.

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Essays on Competition in Banking

Proefschrift ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op maandag 16 maart 2015 om 16:15 uur door Anton Arie van Boxtel, geboren op 29 juni 1984 te Heemskerk.

Promotor: prof. dr. J.J.A.G. Driessen
Copromotor: dr. F. Castiglionesi

Overige commissieleden:
dr. A. Attar
prof. dr. J. Boone
prof. dr. W.B. Wagner

Preface

This thesis has been written over the past four years, during my time first as a master's student and then as a Ph.D. student in finance at Tilburg University. It consists of three papers on financial intermediation. Though each chapter is intended as an independent, stand-alone paper, there is one overarching theme: the effects of competition on financial intermediation. The classic reasoning behind competition is that it offers choice to consumers. This should drive down prices and improve quality for consumers. However, this simplified reasoning fails to take into account *externalities* between the competing parties.

The separate chapters of this thesis study the effects of these externalities in three contexts: Chapter 1 is a heavily modified version of my research master's thesis at Tilburg University. It studies the competition between banks as they try to hire talented workers, and how this competition leads to wage structures that induce excessive risk taking. Chapter 2 is co-authored with the co-promotor of this thesis, Dr Fabio Castiglionesi, and my other supervisor, Dr Fabio Feriozzi, who is currently working at IE Business School in Madrid. This chapter covers how the possibility for firms to privately contract with multiple investors leads to excessive liquidity provision. Finally, Chapter 3 covers how, in an economy in which a firm contracts investment and liquidity insurance with multiple investors, intermediaries are necessary for investment to be possible.

Before going to the heart of the thesis, there will first be an introduction for a general audience, both in Dutch and in English. Then, the abstracts for all three chapters follow in the "Abstracts" section of the introduction. These are intended for an audience of academics with a background in finance or economics. Academic introductions to each chapter will be at the beginning of the respective chapters. The bibliography is combined for all chapters. There are three appendices with proofs and some additional technical details. Except for some direct quotations, the entire body of the thesis is written in English.

Acknowledgements

Of course this thesis would not have existed if it were not for the support from so many different people over the years. As always, this list is far from complete and I apologize in advance for the omissions that are bound to occur.

I need to start by thanking my supervisors and co-authors, Fabio Castiglionesi and Fabio Feriozzi. Despite the fact that they have their nationalities, some of their academic background, and of course their first names in common, I could not have

wished for a more diverse and mutually complementary duo of supervisors.

Fabio Castiglionesi taught me economics. He always stimulated me to think about the economic intuitions in my (and our) work and to always keep the all-important bigger picture in mind. Thinking back to my arrival in Tilburg, I realize how much I learned about financial economics since then. Most of what I learned, I owe to him. I really enjoyed our frank and open discussions about my work, as well as on our joint work, and Fabio's indispensable advice on, and intermediation for, my academic career.

Fabio Feriozzi's attention to detail and insistence on formal correctness have greatly benefited our joint work, as well as my individual work. Very often, when I was running ahead of myself, Fabio reined me in and made sure I set my technical details straight before running on. Beside that, it is great to work with such a kind and friendly person as Fabio.

The Finance Department at Tilburg University has been a very stimulating and interesting environment to work in. The great seminar series and brown bag seminars have been inspiring and have also given me the opportunity to put my own work to the test. I would like to thank Luc Renneboog for reeling me into the department's great Ph.D. programme, and Juan Carlos Rodriguez for so smoothly organizing all the details so I could enter the programme.

I have found the department a very welcoming place to work in, with fun and interesting people. On a social level, interactions both with my fellow graduate student and with faculty were great. The research in this thesis also benefited substantially from discussions with many (former) department members and some of the senior Ph.D. students. Discussions with Fabio Braggion, Marco Da Rin, Vasso Ioannidou, Olivier de Jonghe, Peter Cziraki, Vincent van Kervel, Thomas Mosk, and Erik von Schedvin come to mind.

My cohort in the Ph.D. programme is a group of great, intelligent, and very fun people. Besides some discussions about research and the mutual support we had from each other when we were doing course work, this group has mostly been important in providing the necessary distractions from research. Despite the limited number of attractions that a city like Tilburg has to offer, my time there has been made much more enjoyable by Bernard van Doornik, Paola Morales, Çisil Sarisoy, Larissa Schäfer and, as a bit of an outsider, Patrick Tuijp. Thank you guys for all the great times.

My time in Toulouse has been a very important formative experience. The opportunities I had to interact with so many great economic theorists, concentrated in a single place, have been beyond useful. For this I primarily have to thank my host Andrea Attar. I am very grateful to him for making my stay in Toulouse possible, as well as for all the discussions we had that helped me shape the way I think about my research, especially in the area of non-exclusive competition.

In Toulouse I had the chance to discuss research with many different outstanding researchers that all have contributed greatly to my research, too many to mention here. Nonetheless, I would like to mention how honoured I feel that Jean Tirole managed to accord some of his precious time to me, for some extremely useful discussions on my work. Furthermore, my time in Toulouse has been made all the more pleasant by the presence of the very fun and welcoming population of other

Ph.D. students there, too many to mention by name.

A great thanks also goes out to the many people at the Institute for Advanced Studies (IHS) and the Vienna Graduate School of Finance. The IHS has been more than kind to hire me after such a smooth and pleasant application process. I remember my interview mostly for the great interaction and for the excellent feedback I received on my presentation of what is now Chapter 3 of this thesis. The Institute has kindly allowed and enabled me to work on the papers that make up this thesis, and the IHS and VGSF have provided an intellectually stimulating environment with smart and interesting people around me.

Of course there have been many people who have contributed directly and indirectly to the completion of this thesis outside academia, or before I officially started the Ph.D. programme. All the teachers at my primary schools, at my secondary schools, at Leiden University, and during my master's programme at the École Polytechnique. Too many to mention have taught me, shaped my thinking, challenged me intellectually, or put me in my place when it was needed.

All the friends I met during my studies in Leiden and in Paris, during my stay in the Congo, and during my times in Tilburg, Toulouse and Amsterdam have been a great source of support. Having a social life that does not revolve around academia has had many benefits: most of all the much needed distraction it provides, and the fact that it forces one to look outside the narrow confines of one's own academic research. I especially want to mention the guys from my club "Fidel" and from my old student house "Des Gueux" at Vliet 15 in Leiden. They have been, and still are, a solid base to which I can always return. For a nerdy academic like me, it is great to have such an interesting and diverse group of friends.

The best acknowledgements are saved for last: I of course need to thank my wonderful, sweet, and beautiful girlfriend Anna. To be together with an artist is the greatest thing that can happen to an academic: even if all our combined efforts to understand the world a little better prove to be futile, Anna, I know that you are there, making that same incomprehensible world a little more beautiful. But most of all, min käresta, du gör mig så lycklig. Lyckan, som du ger mig varje dag, gav mig den energi och inspiration som jag behövde, när jag färdigställde den här avhandlingen. Tack så jättemycket!

But finally, I need to thank my family. It has been, and still is, a privilege to grow up among such a loving family. My family stimulated my intellectual curiosity from an early age and taught me how much fun learning and discovering can be. Yet, they also kept me firmly grounded: I learned from them that good scholarly achievements do not entitle you to any feeling of superiority, but rather are a challenge to use the talents you have been given to the fullest. Words cannot describe the importance of the support, both moral and material, that my family has provided me at every step along the very erratic course that led me to this Ph.D.

Maria, Ineke, dank jullie wel vooral voor het geduld dat jullie hebben gehad met jullie betweterige broertje en al zijn rare projecten. Pap, mam, dank jullie wel voor alles. Zoals ik zei, het is niet in woorden uit te drukken hoe belangrijk jullie steun voor mij is geweest. Ik kan zonder een spoor van twijfel zeggen dat zonder jullie dit proefschrift er niet was geweest.

Contents

Introduction

Inleiding voor een algemeen publiek — Nederlands

In het komische stuk "De knecht van twee meesters" uit 1745 verhaalt de Venetiaanse toneelschrijver Carlo Goldoni over de avonturen van de knecht Truffaldino. Deze Truffaldino heeft al een meester voor wie hij werkt, maar krijgt op een gegeven moment de mogelijkheid gepresenteerd om tegelijkertijd ook nog voor een andere meester te werken. Hij heeft altijd honger en watertandt dan ook bij de gedachte om een dubbel loon te ontvangen en nog meer eten te kunnen kopen om zijn geweldige honger te stillen. Hij overweegt stilletjes:

> "...zou het niet mooi zijn om ze allebei te bedienen, dubbel loon te ontvangen en twee keer zo veel te eten? Het zou mooi zijn, als ze er nooit achter kwamen. En als ze erachter komen, wat heb ik dan te verliezen? Niks. Als de ene me de laan uitstuurt, houd ik de andere over."[1]

Hij grijpt de mogelijkheid met beide handen aan en gaat voor beide meesters werken. Al snel komt hij erachter dat dit moeilijker is dan hij had gedacht: hij haalt de taken die hij voor zijn verschillende meesters uit moet voeren door elkaar, komt door zijn werk voor de een niet meer aan het werk voor de ander toe en probeert angstvallig voor elk van zijn beide meesters verborgen te houden dat hij ook nog voor de ander werkt. Tegelijkertijd probeert hij ook nog zijn eigen, niet ongeringe, eetlust te verzadigen. Het is niet moeilijk voor te stellen in wat voor doldwaze avonturen hij hierdoor verzeild raakt. Gelukkig loopt het allemaal goed af en vindt Truffaldino zijn ware liefde. Eind goed, al goed.

Tegenwoordig spreken we niet meer van "meesters" en "knechten", maar in de economische theorie vinden we relaties zoals die tussen Truffaldino en zijn meester nog steeds bijzonder interessant. We spreken van *principaal-agentrelaties*. In dit geval zouden we Truffaldino de *agent* noemen, en zijn meester de *pincipaal*. Het centrale kenmerk van dit type relaties is *informatie-asymmetrie*: de knecht kent zijn vaardigheden beter dan de meester die kent en de meester kan niet alles wat de knecht doet in de gaten houden. De tak van de economische wetenschap die zich bezighoudt met dergelijke situaties heet *contracttheorie*. De contracttheorie bestudeert hoe verschillende partijen ondanks informatie-asymmetrie toch met elkaar zaken kunnen doen door contracten te schrijven.

[1] In het origineel: *"...No la saria una bella cossa servirli tutti do, e guadagnar do salari, e magnar el doppio? La saria bella, se no i se ne accorzesse. E se i se ne accorze, cossa pèrdio? Gnente. Se uno me manda via, resto con quell'altro."*, vertaling is van mijn eigen hand.

In het bijzonder in de financiële economie speelt informatie-asymmetrie een belangrijke rol. De contracttheorie is dan ook een belangrijk middel in het bestuderen van financieel-economische vraagstukken. Verzekeraars weten bijvoorbeeld niet hoe risicovol een verzekerde is voordat ze een polis afsluiten, en kunnen na het afsluiten van de polis ook niet constant in de gaten houden of de verzekerde zich niet roekeloos gedraagt. Banken kennen niet van tevoren de succeskansen van een bedrijf waar ze in investeren of de precieze kredietwaardigheid van iemand die geld komt lenen. Aandeelhouders van bedrijven kunnen niet alle werkzaamheden van directeurs en managers in de gaten houden. Zo zijn er tal van voorbeelden. Met contracttheorie kunnen we verklaren waarom je een eigen risico op je verzekering hebt, waarom er onderpand op leningen zit en waarom directeurs in aandelen en opties worden uitbetaald, in plaats van alleen in geld.

In de financieel-economische wetenschappelijke literatuur zijn veel van deze problemen bestudeerd in de context van *exclusieve* verhoudingen tussen principaal en agent, en dat terwijl een 18e-eeuwse Italiaanse toneelschrijver al kon bedenken wat een problematische — en hilarische — situaties het op kan leveren als een agent met meerdere principalen tegelijk kan handelen. Zelfs in modellen met meerdere banken, verzekeraars of investeerders die met elkaar concurreren, werd vaak aangenomen dat de klant, het bedrijf, de manager of de verzekerde uiteindelijk maar bij één aanbieder een contract kan afnemen. Dit wordt *exclusieve concurrentie* genoemd.

Exclusieve concurrentie levert vaak de beste uitkomsten op voor de agent, die als klant kan kiezen tussen meerdere contracten die worden aangeboden. De agent kiest dan die aanbieding, die voor hem het beste is. Als die aanbieding niet de best mogelijke is, en de aanbieder er winst op maakt, zal een andere aanbieder altijd met een net beter contract komen. Deze logica gaat echter niet meer op als de agent, net als Truffaldino, stiekem met meerdere aanbieders tegelijk in zee kan gaan. Dit komt omdat het contract dat je met één principaal hebt, invloed heeft op je gedrag ten opzichte van de anderen. Dergelijke economische situaties, met *niet-exclusieve* concurrentie, zijn in de laatste decennia binnen de economische theorie steeds uitvoeriger bestudeerd. Het is ook zeker niet lastig voor te stellen dat het bestuderen van deze situaties veel aan praktische relevantie heeft gewonnen. Met de opkomst van nieuwe technologieën, met nieuwe mogelijkheden om internationaal zaken te doen en met mazen in de regelgeving, die slechts met moeite de ontwikkelingen in de wereld kan bijhouden, is het steeds makkelijker geworden om met meerdere partijen tegelijk te handelen.

Hoofdstukken 2 en 3 (hoofdstuk 1 bespreek ik verderop in deze inleiding) gaan over de effecten van niet-exclusieve concurrentie in concrete financieel-economische situaties. Hoofdstuk 2 houdt zich bezig met de *liquiditeit* van bedrijven. In de gangbare modellen van liquiditeit komt een bedrijf met zijn investeerders overeen hoeveel die investeerders in het opstarten van een langetermijnproject steken en hoeveel er achter de hand gehouden moet worden voor latere onvoorziene kosten. Dit geld kan door het bedrijf zelf achter de hand worden gehouden in de vorm van een voorraad contanten, of door de investeerders worden voorzien in de vorm van een kredietlijn. Bij het aanhouden van liquiditeit speelt echter een afweging een rol: bij teveel liquiditeit lopen de mogelijke kosten op, hetgeen voor de investeerders beperkt hoeveel zij in het project kunnen steken. Bij te weinig liquiditeit bestaat het risico dat het bedrijf zijn kosten niet kan betalen en voortijdig opgedoekt moet worden, en zijn langetermijnopbrengsten niet kan realiseren. Zodoende is het optimaal om een vaste, maar begrensde, hoeveelheid liquiditeit af te spreken. Dit zal in een model met exclusieve concurrentie dan ook de uitkomst zijn.

Stel nu dat een bedrijf tot een zeker bedrag aan liquiditeit heeft, zeg duizend euro. Als dat bedrijf er dan achter komt dat het duizend en een euro nodig heeft, en het de mogelijkheid heeft om ongemerkt een andere investeerder te benaderen, dan zal het bedrijf graag de duizend euro van zijn oorspronkelijke liquiditeit gebruiken, en nog een euro van een andere investeerder aantrekken. Natuurlijk geldt dit nog steeds als het bedrijf tweeduizend euro nodig heeft, of drieduizend. Het is dus in de praktijk moeilijk om de hoeveelheid liquiditeit te beperken. Het tweede hoofdstuk van dit proefschrift leidt dit af in een formeel model en brengt dit in verband met het feit dat in veel verschillende landen recent de cashvoorraden van bedrijven om mysterieuze redenen aanzienlijk gegroeid zijn.

Het derde hoofdstuk gaat verder met het model van het tweede. Net zoals in hoofdstuk 2 hebben bedrijven een grote hoeveelheid liquiditeit nodig van investeerders. Als investeerders zelf maar een beperkte hoeveelheid geld hebben, zijn meerdere investeerders tezamen nodig om de potentiele kosten van het bedrijf te financieren. Indien meerdere investeerders samen echter het zelfde bedrijf financieren, heeft elk van de investeerders er baat bij om het andere bedrijf zoveel mogelijk voor de kosten op te laten draaien. Hoofdstuk 3 beschrijft een model waarin dat altijd mogelijk is, waardoor de economie ook niet functioneert als meerdere investeerders elk een op een met een bedrijf handelen.

De enige manier waarop er in deze situatie toch geïnvesteerd kan worden, is als de verscheidene investeerders hun geld aan een soort *tussenpersoon* geven, die vervolgens het geld weer gebruikt om in het bedrijf te investeren en om het bedrijf wanneer nodig van liquiditeit te voorzien. Op deze manier probeert hoofdstuk 3 het bestaan van financiële tussenpersonen, zoals banken, te verklaren. Dit wordt in verband gebracht met bevindingen uit de economische geschiedenis, over in welke periodes van de geschiedenis en in welke gebieden banken een belangrijke rol spelen in economische ontwikkeling. Vooral Duitsland aan het einde van de negentiende eeuw is in deze context veel bestudeerd.

Natuurlijk is er ook een eerste hoofdstuk. Dit gaat over een enigszins ander onderwerp, maar heeft nog steeds te maken met de effecten van concurrentie in de financiële sector. Het gaat over een onderwerp dat in de media uitvoerig besproken is: het verband tussen de beloningsstructuur van bankiers en risico. Vaak wordt een beeld geschetst van een “graaicultuur” van hebzuchtige bankiers die met hun onverantwoorde en roekeloze gedrag de economie aan de rand van de afgrond hebben gebracht, slechts om hun eigen bonus veilig te stellen.

Zonder een uitspraak te willen doen over hoe terecht de publieke verontwaardiging is, probeert het eerste hoofdstuk van dit proefschrift in een theoretisch model het verband tussen bonussen en risico’s in het bankwezen te bestuderen. In het publieke debat wordt de kritiek vaak gericht op bankiers of banken zelf, of wordt er gesproken over “cultuur”. Dit proefschrift, daarentegen, analyseert hoe bonussen het gevolg zijn van concurrentie op de arbeidsmarkt voor bankiers, en probeert te zien hoe de kenmerken van deze arbeidsmarkt direct kunnen leiden tot overmatig risico.

In het eerste hoofdstuk wordt deze arbeidsmarkt geanalyseerd met een vereenvoudigd model. Potentiële bankiers verschillen in hoe goed zij zijn in het doen van risicovolle investeringen. “Goede” bankiers kunnen voor de bank door het nemen van enig risico daadwerkelijk betere winsten behalen, terwijl “slechte” bankiers dit niet kunnen. De goede bankiers zijn de bankiers die goed zijn in het selecteren van de juiste projecten om te financieren, die de betere investeringen kunnen uitkiezen of de beste deals kunnen vinden. De slechte bankiers kunnen nog steeds op gemiddeld niveau presteren als ze geen risico nemen.

Als de bank een bankier in dienst wil nemen, dan weet ze niet of deze goed of slecht is. Door de juiste beloningsstructuur kan ze echter de goede bankiers selecteren. Door gemiddelde prestatie niet al te hoog te belonen, maar juist aan uitzonderlijk goede prestaties een bonus toe te kennen, schrikt de bank de minder goede bankiers af, maar trekt ze juist de betere aan, omdat die weten dat ze een grote kans hebben de bonus te krijgen.

De enige vraag die rest aangaande de beloningsstructuur is wat er gebeurt bij slechte prestaties. Een soort “straf” op slechte prestaties is de aangewezen manier om bankiers te weerhouden van het nemen van té veel risico. In de praktijk is de mate waarin een bank zijn werknemers kan bestraffen echter vaak beperkt: ontslag is meestal het ergste wat de bank kan doen, en zelfs dat is lastig. Deze beperking maakt het lastig om overmatig risico tegen te gaan.

Als goede bankiers een veel hogere gemiddelde beloning eisen dan minder goede bankiers, dan moet dat gebeuren door een hogere bonus. Dit maakt het noodzakelijk voor een bank om een beloningsstructuur aan te bieden die ertoe leidt dat bankiers overmatig risico nemen. Het feit dat goede bankiers een hogere beloning eisen kan een gevolg zijn van concurrentie op de arbeidsmarkt. Hoofdstuk 1 geeft een vereenvoudigd voorbeeld van een arbeidsmarkt waar dit gebeurt: in deze arbeidsmarkt concurreert één grote bank met meerdere kleine banken en is er één goede bankier tussen vele minder goede. De grote bank is bereid om veel meer te betalen dan de kleine banken voor deze goede bankier, aangezien zijn vaardigheden meer verschil uitmaken op de grotere investeringen van de grote bank. Dit komt overeen met hoe de betere bankiers voor grote investeringsbanken als Goldman Sachs en J.P. Morgan komen te werken.

Er is echter één probleem dat het lastig maakt voor de grote bank: de kleine banken zouden de goede bankier ook graag willen hebben. Om er voor te zorgen dat de goede bankier niet weggekaapt wordt, moet de grote bank dus een hoge bonus bieden, maar moet ze nog steeds een laag basissalaris bieden om ervoor te zorgen dat ze de minder goede bankiers weghoudt. Dit leidt tot een beloningsstructuur die op haar beurt weer bankiers aanzet tot het nemen van overmatig risico.

Hoe hoog dat risico is, hangt af van de structuur van de arbeidsmarkt: bij een zeer flexibele arbeidsmarkt kunnen kleine banken gemakkelijk goede bankiers wegkapen. Dit drijft de prijs van deze bankiers op, waardoor de banken extra hoge bonussen moeten geven, hetgeen weer leidt tot een hoog risico. Als de arbeidsmarkt minder flexibel is, is dit effect minder sterk en ligt het risico minder hoog. In beide gevallen kan het risico echter nog steeds overmatig zijn, als gevolg van de concurrentie op de arbeidsmarkt.

En dat brengt ons weer terug bij het overkoepelende thema van dit proefschrift: concurrentie in de financiële sector. Ik wil niet beargumenteren dat concurrentie in wezen slecht is. In het dagelijks leven komen we vaak genoeg voorbeelden tegen van hoe concurrentie leidt tot lagere prijzen en betere producten. De verscheidene hoofdstukken van dit proefschrift geven echter een aantal voorbeelden van negatieve effecten van vrije concurrentie in het bankwezen: het kan leiden tot overmatige liquiditeit, te weinig investering, overmatig risico of — soms ongewenste — concentratie in de financiële sector.

Introduction for a General Audience — English

In his comedy from 1745, "The Servant of Two Masters", the Venetian playwright Carlo Goldoni describes the adventures of the servant Truffaldino. Truffaldino, who already works for one master, is presented with the option of working for a second master at the same time. As he is always hungry, his mouth waters as he contemplates the possibility of receiving a double paycheck with which to buy food to still his tremendous appetite. He says to himself,

> "...wouldn't it be a beautiful thing to serve both of them, to gain two salaries, and to eat twice as much? It would be great, if they never realized. And if one of them realizes it, what do I lose? Nothing. If one of them sends me off, I'll be left with the other one."[2]

He proceeds to work for both masters, and quickly finds that this is harder than he thought: he starts confusing the tasks he has to do for his different masters, runs out of time to serve both, and tries anxiously to hide from each of his masters the fact that he is working for the other. Of course he still has to satisfy his own immense appetite, so one can of course imagine the series of comical misadventures this leads to. Luckily for Truffaldino, he finds love in the end, and all's well that ends well.

In economics, the kind of relationship between Truffaldino and his master is called a *principal-agent relationship*. In this situation, we would refer to Truffaldino as the *agent*, and to his master(s) as the *principal*, since we do not really talk about servants and masters anymore these days. These sorts of relationships are characterized by *information asymmetry*: a "servant" typically knows more about his own ability than his master does, and the master cannot monitor every single action the servant takes. The branch of economics that studies outcomes in these sorts of relationships is called *contract theory*. It studies how economic parties try to bridge these information asymmetries by writing contracts on observable outcomes.

Contract theory is a useful tool to describe situations in financial economics, where asymmetric information plays an important role: insurance companies cannot monitor the behaviour of their policy holders or know how risky they are, banks cannot fully observe the behaviour and the creditworthiness of firms or private borrowers that they finance, and shareholders cannot control what their executives are

[2] The original reads *"...No la saria una bella cossa servirli tutti do, e guadagnar do salari, e magnar el doppio? La saria bella, se no i se ne accorzesse. E se i se ne accorze, cossa pèrdio? Gnente. Se uno me manda via, resto con quell'altro."*, translation is my own.

doing. Thus, contract theory rationalizes why we have deductibles on our insurance or collateral on our mortgages, and why executives are paid in stocks and options, rather than just in cash.

However, even though an 18th-century Italian playwright already realized the problems that could arise from having multiple principals, many situations in financial economics have only been studied with the assumption that an agent would only have an *exclusive* relationship with a principal. In many cases there would be multiple principals (e.g., multiple banks, multiple insurance companies) competing in offering contracts to the agent, but the agent is restricted to choose only one of them. This we refer to as *exclusive competition*. Competition, it was reasoned, would always make sure that the agent best possible contract given the information asymmetries. The principal, on the other hand, would make no profit. If any principal would offer a contract that is not the best one for the agent, another principal would come in and offer a slightly better contract.

However, these papers tacitly assume that the agent can only choose the contract from one principal. If, like a modern-day Truffaldino, a firm can borrow from several banks, an executive can get paid by multiple firms, or a person can get insurance from multiple companies, the classic reasoning might not work anymore. This is because the contract an agent has with one principal has an influence on his behaviour towards the others. In the last decades, these situations of *non-exclusive competition* and *common agency* have become an active field of study in (financial) economics. One can even imagine that if economies are becoming freer and more international, and if technology or legal loopholes are making it easier for firms and private persons to secretly contract with different financiers, then studying these situations is becoming more and more relevant.

Chapters 2 and 3 (Chapter 1 will be discussed later on in this introduction) of this thesis are concerned with the effects of non-exclusive competition on very concrete situations in finance. Chapter 2 considers the effect on the *liquidity* of firms. In the conventional models of liquidity, investors and firms agree on how much they invest at the start-up of an investment project, and how much they keep in order to face potential, uncertain, costs at a later date. This liquidity is then supplied by allowing the firm to hold a cash buffer, or by investors providing a firm with a credit line up to a certain amount. There is a trade-off here: if investors supply too much liquidity, the potential costs of the project are higher, which limits the amount the investor can put into the project initially. If there is too little liquidity, the chance that the firm can face costs becomes smaller, limiting the potential revenue from the project. Thus, it is optimal to agree upon a certain limited amount of liquidity.

However, let's say that the firm can secretly contract with multiple investors. If it can only get liquidity up to, say, a thousand euros, and then faces a cost of a thousand and one euros, it would really like to have this additional euro. For this one euro, the firm can approach another investor, while still getting the thousand from its credit lines or cash buffer. This is still the case if the firm can get liquidity up to two thousand or three thousand euros, and so on. This means that, in principle, there is no limit to the liquidity that the firm can obtain from investors. The second chapter of this thesis studies this, and uses it to explain the puzzling fact of recent build-ups in cash holdings by firms.

The third chapter expands upon the second. In the second chapter I find that investors need to provide large, potentially unlimited amounts of liquidity to firms. The problem is that, if the investors themselves only have a limited amount of funds, these might not be sufficient to supply a firm with all the liquidity it needs. This would mean that multiple investors together need to team up to finance a firm. However, in that case, a new problem arises: each investor in a team of investors would rather see the others in the team provide liquidity to the firm before they do. Chapter 3 looks at situations in which investors would always have a possibility to make sure of that, through adjusting the price they charge the firm for liquidity. This leads to a market breakdown.

That means that one investor is not enough, but two are too many. The only way out of this conundrum is for the different investors to pass their money on to one single *intermediary*, who then invests in the firm. Thus, Chapter 3 explains why banks, and other intermediaries, are needed within a financial system. Furthermore, Chapter 3 ties this to some of the previous work done by economic historians, which asks why intermediated finance played such a major role in certain places and in certain episodes of history. Especially Germany in the late nineteenth century forms an interesting case study.

Of course, there is also a first chapter. This chapter is on a slightly different topic than the other two, though it also studies some of the negative effects of competition on the banking system. It discusses a topic that has been very actively discussed in the public sphere: bonuses and risk. The image of bankers taking massive risks, bringing our economy to the brink of collapse just to get their precious bonuses, has been painted all too often in the media. Much of the public outrage has been directly aimed at the bankers or banks themselves. Slogans about the "fat cats" on Wall Street having learned nothing, still carelessly taking excessive risks, abound.

Without taking a stand about how righteous this outrage is, the first chapter tries to theoretically understand the connection between risks and bonuses. And, as any paper in economics is supposed to, it tries to do so without hating the player, but rather by taking a good look at the game. The first question that needs to be answered is why banks offer their bankers a remuneration structure with bonuses, a second question is what constitutes "excessive" risk taking, and the third is whether and how the remuneration structure leads to excessive risk. The final, and probably most important, question is whether, if there are remuneration structures leading to excessive risk, we can do anything about it, for example by regulating the banking sector.

The first chapter analyzes a simple model in which banks try to hire bankers. Some bankers (the "good" bankers) are better than others, in the sense that if they take risk, they have a better chance of performing well. These might be the bankers that are really good at finding good deals, finding the right companies to finance, or finding the "next Apple" stock to invest in. The less skilled bankers (the "bad" ones) can still perform adequately and get an average performance.

If banks do not know the skill level of their prospective employees, they can still make sure they hire the good employees by rewarding good performance. They can make sure to hire only the good bankers, keeping out the bad ones, by not rewarding a mediocre performance very well. This can be achieved through a pay structure with a relatively low base salary, but high bonuses.

The question that is left about the remuneration structure is what happens in case of bad performance. In practice, there is a limit to how much a bank can "punish" its employees in case of bad performance. Firing an employee is often the worst they can do, and even that is often very complicated. However, the prospect of some sort of punishment is the way to keep bankers from taking excessive risk.

Chapter 1 models how, if good bankers command a much higher salary than bad ones, banks need to set a pay structure that induces excessive risk taking. This can be a result of competition on the labour market. In the simple model of the labour market, as presented in this thesis, there is one larger bank, with a large amount of assets to invest, competing on a labour market against many smaller banks. There is only one good banker out there. The large bank really wants to hire this banker, as investment skills will make more of a difference with the large investments of this bank than with those of the smaller banks. In the real world, this would mean that the better bankers end up working for Goldman Sachs or J.P. Morgan, simply because their skill makes more of a difference there.

Smaller banks, though, would still rather hire this one good banker than one of the bad ones, and thereby drive up the price of the good banker: the bigger bank needs to come up with a sufficiently lavish pay package to make sure the good banker does not choose to go to a smaller bank. The big bank, however, still needs to set its base salary low enough to make sure bad bankers do not accept employment. Thus, in order to attract the good banker, and only the good banker, the bank needs to set a pay package that includes a large bonus to reward good performance.

In the first chapter, it is shown that this bonus can be so high that it becomes attractive for the banker to choose an overly risky investment. Just how risky depends on the way the labour market is organized: if the labour market is very flexible, and bankers could potentially switch banks easily, then it becomes very costly to retain good workers, which drives up the potential bonuses and the resulting risk. If the labour market is relatively rigid, and bankers cannot easily switch between jobs, excessive risk can still arise as a consequence of labour market competition.

This brings us to the overarching theme of this thesis: competition in the financial sector. Of course this thesis does not in any way want to argue that competition is bad. In our daily life we often come across examples of how competition leads to better products at lower prices. However, this thesis argues that, especially in the financial sector, free competition can have some negative effects: it can lead to excessive liquidity holdings, underinvestment, consolidation and excessive risk.

Abstracts

Below are the abstracts of the three different chapters.

Abstract of Chapter 1

This chapter argues that excessive risk taking by financial institutions is a result of the need for these institutions to hire their traders from a labour market with dispersed talent. On the one hand, institutions want to hire only talented workers,

making sure untalented ones do not seek employment with them. In order to pick only the best bankers, the institution can offer a low base salary and high bonuses. Talented workers, knowing they have a large chance of obtaining the bonus, accept employment while less talented ones do not. On the other hand, if workers are protected by limited liability, these high-powered incentives can lead to excessive risk taking. This chapter first offers a simple model with workers of different abilities who have different outside options, and derives conditions under which excessive risk is taken. Then a labour market model is studied with banks of different sizes, in which the most talented worker ends up working for the biggest bank, where his talents are most productive. However, the competition from smaller banks endogenously raises the outside options for the good trader, giving rise to the need for high-powered incentives and scope of excessive risk taking. Then the effects of labour market mobility on the incentives to take risk are studied.

Abstract of Chapter 2

This chapter studies the effect of non-exclusive competition on liquidity provision in a generic financial intermediation setting. Consider the baseline model by Holmström and Tirole (1998) in which a firm in need of funds exclusively deals with a lender. The lender is willing to provide an up-front investment and a *finite* liquidity facility in exchange for part of the project's proceeds. The firm obtains a share of its project's payoff because of a moral hazard problem at the firm level. If the firm can privately contract with several lenders, there is a difficulty in limiting liquidity provision. Outside lenders can free ride upon the liquidity provided by an incumbent lender in exchange for the firm's original share. As a first result, this has the effect that the equilibrium from Holmström and Tirole (1998), with exclusive competition, is no longer sustained. As a second result, we show how an incumbent lender can ward off the outside lenders by offering unlimited liquidity support. The observed shift from exclusive to non-exclusive contracting environments could therefore help to explain the increase in liquidity holdings by firms.

Abstract of Chapter 3

This chapter argues that financial intermediaries serve to coordinate competition between investors. It starts out by modeling an economy without intermediaries: borrowers have access to a project that requires initial investment and faces a stochastic liquidity shock at an intermediate date, before realization of the project's proceeds. Investors can supply funds for starting the project and for insuring the project against liquidity shocks. Competition between investors is assumed to be *uncoordinated* and *non-exclusive.* The non-exclusive nature of the competition makes it impossible to limit the intermediate date liquidity supply to the borrower, as investors can try to extend the liquidity supply to a borrower by free-riding upon the liquidity supply of others. However, as the liquidity supply needs to be larger, multiple investors together are needed to finance each borrower, giving rise to a common agency problem: each investor wants to make sure the other investors are responsible for supplying liquidity more than he is. These two problems lead to an unraveling of the market if borrowers only contract directly with investors. Trade can be restored by intermediaries. The chapter discusses how either a social planner can restore trade by becoming an intermediary or how investors can offer to become intermediaries.

Chapter 1

Trader Compensation and Bank Risk: a Screening Approach

If you out for mega cheddar, you got to go high risk.

Ice T, *Don't Hate the Playa*

Bonuses for executives at banks, hedge funds and asset management companies have led to a great controversy as a result the financial crisis, becoming a major theme in the public debate. Within the public dialogue, bonuses are commonly connected to risk taking. The picture often painted in the public sphere is one of bankers taking excessive, value-destroying risks, attracted as they are by the prospects of high bonuses. From this picture a bewildering question arises: why would a rational bank set a pay structure that induces their traders to take excessive risks? The aim of this paper is to model the role that hidden information plays in the contractual relation between the bank and its traders, and to study the interplay between compensation and risk taking.

In attempting to tackle the issue of risk taking and compensation in the banking sector, one first needs to address the rationale behind a variable remuneration structure. In this paper, variable pay is not primarily used to induce an agent to provide some costly effort that potentially improves the probability distribution of the project this agent manages. Even though bankers can enhance their revenues through better client pitching, more research or closer monitoring, this paper takes unobservable skill differences to be the main rationale for a variable remuneration structure: banks try to screen possible traders by setting return-dependent wage schemes that deter less skillful traders from taking on a job at the bank. Meanwhile, more skillful ones accept these contracts, knowing full well that they have a better probability of earning a higher return, and thus a higher wage.

The reason why skill differences have a particularly salient effect in the financial sector is that the potential gains that a smart, talented worker can earn a bank or fund is only bounded by the funds the financial institution has available for trade, and not by production capacities of plants or demand for goods and services. Several papers in the macroeconomics and labour economics literature (Murphy, Shleifer, and Vishny, 1991; Bolton, Santos, and Scheinkman, 2011; Kneer, 2013) argue that, particularly in economies with highly deregulated banking sectors, banking offers a return on talent to skilled workers that the real economy does not. This paper does not aim to study this difference in returns to skill between the financial and real sectors, but rather takes the high rewards to skill as a given feature of the banking system. As the Squam Lake Working Group (Bernard, et al. 2010) put it,

> "(...), even among those with similar professional qualifications, there are tangible differences in the skills of financial employees, and even a small difference in skill can have an enormous impact on the profits of a financial firm."

The other feature that sets the financial sector apart from other sectors of the economy, and that plays an important role in this paper, is that in a financial firm, at any level of the organizational hierarchy, workers have a large amount of discretion over the risks they take, and to which they subsequently expose the financial firm for which they work. This discretion is often necessary as traders need to be able to react quickly to market movements, and corporate bankers must work to offer loans, underwritings or other services before the competition does. This goes all the way down to loan officers, who can decide on the loans they give to households and small businesses. Ex post, often only the profits and losses of a trade or the performance of a loan can be used in the compensation contract, whereas the discretionary risk that a worker takes remains difficult (or, in this case, impossible) to verify.

In this paper, skill differences give rise to wage differentiation. Good traders have an investment opportunity set that includes risky assets with a higher return than the risk-free investment. There is an optimal investment and with respect to this, there are both inefficiently prudent and inefficiently risky investments. Throughout most of the paper it is assumed that bad traders only have access to the risk-free investment. A bank needs to hire a trader to manage its trading budget. In order to make sure only a good trader accepts employment, it offers a contract with a rather low reward for a merely average return (which also the bad trader could get by investing in the safe investment), but a high reward for a high return (which only the good trader has a high enough chance of achieving). However, if the trader is protected by limited liability, the bank cannot keep him from taking excessive risk.

This dynamic is first studied in a context in which good traders exogenously have higher reservation utilities than bad ones. If traders are protected by limited liability, and good traders command a very high wage premium with respect to bad ones, payment for the good traders must be so convex, as a function of performance, that good traders are induced to take excessive risk.

The paper then goes on to model a stylized labour market in which banks of different sizes offer contracts to the traders. As good traders are more productive at larger banks, larger banks will end up hiring good traders. However, in order to make sure smaller banks do not make more attractive offers to good traders, larger banks need to pay a skill premium for good traders, potentially leading to excessive risk. The level of this skill premium is dependent upon the nature of the labour market: if the labour market is flexible, smaller banks can steal good traders away from larger banks, after learning the trader's type from the fact that he works for the large bank. Thus the premium large banks must pay to keep the good trader is very high in this case. In a less flexible labour market, banks only compete for talent ex ante, so that the good trader's rent is lower. It turns out risk in the more flexible labour market is higher.

The rest of this chapter is set up as follows: after a review of the literature in Section 1.1, Section 1.2 offers a simple model in which good and bad traders exogenously have different reservation utilities. Section 1.3 introduces a stylized labour market, with varying degrees of flexibility. Section 1.4 concludes.

1.1 Literature

On the theoretical side, the literature on the topic of remuneration in financial institutions has been surprisingly scant for a long time, but recently a number of papers have appeared that study (executive) compensation at banks and its interplay with risk taking. Thanassoulis (2012) derives how competition on the banking job market creates a negative externality, increasing the default risks of banks. In another paper by the same author (Thanassoulis, 2011), a model with both moral hazard and adverse selection is presented in which lesser-ability traders have the possibility to shift risks across time. Again because of the externality caused by competition, he finds that it is sometimes the constrained optimal solution to allow lesser-ability traders to shift risks.

In another recent paper, Bolton, Mehran, and Shapiro (2010) address the contracting problem between depositors, debtholders, bank shareholders and executives. They use a pure moral hazard model in which contracts for executives can only be based on market prices of the bank's debt and equity. They find that without regulatory intervention bank executives tend to shift risks to the detriment of debtholders and depositors. They also address how the CDS spread can be used as part of the compensation contract in order to mitigate risk-shifting incentives.

In terms of modeling, this paper is very close to the baseline model in Diamond (1998). In that model, agents' projects have three possible levels of payoffs. Effort gives the agent access to a range of distributions over these payoffs, and the contract needs to both give the agent the incentive to provide effort and align the agent's interests with those of the principal so that the agent chooses the project most profitable for the principal. They find that optimal payment is "almost" linear in the sense that the optimal payment schedule converges to a linear one when the cost of effort tends to zero.

There has been a body of empirical literature studying (executive) compensation and risk taking, mostly focusing on CEOs. Cheng, Hong, and Scheinkman (2010) address the cross-sectional heterogeneity in both executive pay and risk and find that, controlling for firm size and sub-industry, riskier banks tend to have higher executive compensation. They argue that in order to convince executives to work for inherently riskier firms, overall compensation should be higher. DeYoung, Peng, and Yan (2013) use a panel of commercial banks and find that, when executives have compensation packages with a large sensitivity to stock price *volatility*, their banks tend to be riskier according to several measures. They also find that, controlling for this effect, a higher sensitivity to the *level* of the stock price reduces risk taking.

In another recent paper, Fahlenbrach and Stulz (2011) perform a cross-sectional study on CEO pay packages and share performance during the recent 2007-2009 financial crisis. They find that banks in which CEO compensation in 2006 was more sensitive to share price performed worse during the financial crisis. This would support the hypothesis that pay that is strongly related to performance induces more risk taking. They do find, however, that sensitivity of pay packages with respect to share price volatility did not have any influence on share performance during the crisis.

Beside these very recent contributions, there is a large literature on executive compensation and risk, both at banks and other firms. Agrawal and Mandelker (1987) find that, for general firms, option-based executive compensation induces risk taking. Hubbard and Palia (1995) find that pay is more performance-related in less regulated sections of the banking industry. Houston and James (1995) report that the payment structure in the banking industry is significantly different from that in other industries. Their finding is that executive pay at banks is more conservative. Chen, Steiner, and Whyte (2006) find that banks with relatively more option-based compensation tend to be riskier. Cuñat and Guadalupe (2009) report that, following deregulation of the banking industry, variable pay increased at banks.

There exists a related literature on delegated portfolio management. In the classical delegated portfolio management set-up, the agent also has the post-contractual discretion to choose his exposure to a risk factor. This literature starts out with Bhattacharya and Pfleiderer (1985), who present a model in which an investor tries to screen potential money managers that differ in their forecasting accuracy of a normally distributed variable that influences portfolio returns. The model by Bhattacharya and Pfleiderer (1985) is one in which information is the most important good and the accuracy of that information is the determinant of quality. The authors also explicitly assume normal distributions. Their model has very similar assumptions and results to those in the present paper. The main difference between their paper and the present is that they do not explicitly address the issue of risk taking.

An explicit characterization of skill differences is common throughout the literature. A notable exception is the model by Foster and Young (2010), who use a distinction between skilled and unskilled managers that is akin to the one presented in the current work: in their model, a skilled and an unskilled manager can both easily replicate a benchmark. In their paper, the bad agent can also mimic the strategies of a skilled manager, but at a greater downside risk, making track records an imperfect way of measuring skill.

A strongly related paper is the one by Palomino and Prat (2003). They model the case of a money manager choosing assets on behalf of an investor. The manager has the discretion to choose both a level of costly effort and a level of risk, as such presenting the full moral hazard case of the setup in the present paper. They find that the incentive and participation constraints, combined with limited liability, rule out affine contracts. They also find that a binary bonus contract is among the optimal contracts. Lastly, they do find that deviations from optimal risk taking are possible, both in the direction of excessive risk and excessive prudence.

Other papers do not necessarily study the optimal contract, but rather take the contract as given and study its implications. Allen and Gorton (1993) explain the phenomenon of "churning", or trading without any specific reason or insight, by the fact that traders with no particular knowledge enter the market in order to obtain performance fees. Das and Sundaram (2002) study the pros and cons of incentive contracts and symmetric contracts in a signalling model. Hodder and Jackwerth (2007) numerically study the risk-taking incentives provided by typical hedge fund compensation contracts, taking into account the possible differences between the evaluation period of returns and the trading horizons of the fund manager. A comprehensive review of the delegated portfolio management literature can be found in the paper by Stracca (2006).

There have been some papers in the market microstructure literature studying the asset pricing implications of misalignment between the interests of a financial manager and the owner of the assets. As previously mentioned, the paper by Allen and Gorton (1993) that studies the effects of information asymmetry. Froot, Scharfstein, and Stein (1992) model how asset pricing anomalies can be caused by short-term thinking. More recently, Dasgupta and Prat (2008) built a model in which traders' career concerns lead to conflicts of interest and study how these career concerns affect market microstructure.

There is also some debate about to what extent traders' skills or the effort they put into information acquisition can influence their returns. Indeed, several studies (Malkiel, 1995, 2003; Gruber, 1996) find that, on average, active mutual fund investors do not perform better than the passive market benchmark. The fact that mutual fund managers, who can devote all their time and effort to investment-related activities, do not outperform the market can be explained by either the hypothesis that effort does not make much of a difference or that the pay structure in the highly regulated mutual fund industry does not induce effort. In order to examine the effect of skill, it is much more relevant to study persistent cross-sectional heterogeneity in fund performance. Berk and Green (2004) find only mixed evidence for persistence in relative performance, but also model how the absence of persistence does not necessarily imply the nonexistence of differential ability across managers. In a large study conducted among Finnish retail investors, Grinblatt, Keloharju, and Linnainmaa (2012) find that investors with a higher IQ exhibit more rational, "sensible" trading behaviour and gain higher returns.

In the theoretical literature as well as in the public debate, there has been mention of *deferred pay*. The public debate focuses on how deferred pay prevents traders from buying bubble assets, as is also put forward by the Squam Lake Working Group (Bernard, et al. 2010). The main theoretical contribution in the field focuses on another rationale behind deferring bonuses: Jarque and Prescott (2010) model the situation in which a banker's actions have an influence on both first- and second-period cash flows. If the second-period cash flow is informative about the banker's effort, the bank has to use this information in the trader's remuneration, which can only be done by deferring pay. The model by Thanassoulis (2011) also addresses the use of deferred pay, and its shortcomings, in mitigating the risk-shifting incentives of lesser-ability agents. Whatever the rationale, the idea of deferred pay has caught on, both in the industry itself[1] and among policy makers[2].

This paper differs from most of the literature relating moral hazard and (executive) compensation in that the main hidden action the agent can engage in post-contractually is choosing the risk he exposes the principal to. Furthermore, the adverse selection aspect of the problem necessitates a more complicated structure, in which the typical "bonus contracts" studied in the delegated portfolio management literature (see, e.g., Palomino and Prat, 2003) are not always feasible. The main difference to the delegated portfolio management literature that does study adverse selection is that, in the principal-agent set-up specific to banks, screening is more feasible as the bank is more likely to move first in the contract offering, whereas contracting in a delegated portfolio management setting is more likely to give rise to a signalling problem (as in Das and Sundaram, 2002). Nonetheless, the seminal contribution by Bhattacharya and Pfleiderer (1985) justifies screening by treating a principal like a large consortium of investors, who can set the terms of the contract.

Bénabou and Tirole (2013) also analyze compensation from a screening point of view, but instead of skill focus on the banker's intrinsic motivation to work as the central quality. They find that, if bankers can divide their time between a socially valuable task and a more monetarily rewarding task, competition on the labour market can skew the offered contracts in such a way that the bankers shift their attention away from the valuable task. The paper closest to the present is the one by Bijlsma, Boone, and Zwart (2012), which models a labour market for traders of varying skill, yielding excessive risk in competitive markets. However, the authors model the labour market in a Hotelling type of model, in which bankers' distances to their prospective employer matters, rather than the potential productivity of bankers (as in the present paper, proxied by the size of banks).

1.2 Model: Exogenous Reservation Utilities

In the first model, to get the basic intuition of the paper, I assume that good traders exogenously have higher reservation utilities than bad ones. Though this could be seen as a reduced form of the stylized banking labour market, which is presented later in the paper, it is also not a strange assumption to make from the onset. Aspiring bankers with higher skill levels could have better career options in other sectors, or be more productive when self-employed. This section derives conditions on the model parameters under which excessive risk is taken.

Players There is one monopolist principal, the *bank* (she) who needs to hire an agent, the *trader* (he), to invest in a financial asset. The bank needs precisely one trader to be able to invest. Hiring more traders will not give any more profits, as the bank only has a limited budget available for trading. There is a countably infinite set of traders, each with a privately known type $\vartheta \in \Theta = \{B, G\}$. I refer to these

[1] As is demonstrated by Crédit Suisse's *Partner Asset Facility*, cf. for example the article by Richard Beales (2008) in the New York Times

[2] Such as French president Nicolas Sarkozy, cf. the article by Perrine Créquy (2009) in Le Figaro

two types as *bad* and *good*. Only a finite subset of these traders is good, reflecting the assumption that trading skill is an exceedingly rare quality.

The assumptions on the number of traders needed by the bank are markedly different from those in some other papers (Bannier, Feess, and Packham, 2012; Bijlsma, Boone, and Zwart, 2012; Bénabou and Tirole, 2013). In these papers, banks hire both good and bad bankers and have both self-select into a type of contract and a corresponding level of risk. This makes sense for banks that engage, for example, in loan origination, where every banker potentially makes an additional profit for the bank that outweighs his wage costs. However, this paper tries to capture the typical situation at proprietary trading desks of investment banks and hedge funds, where small teams are in charge of these institutions' entire trading budget for a given asset class. These teams do not hire more traders, despite a large pool of applicants aspiring to be part of them. Adding traders to these teams will not be beneficial to the bank, as the potential revenues of these teams are constrained first and foremost by the budget they have available for trading.

The bank offers every trader a menu of compensation contracts until one of them accepts it. The trader, upon observing the offers, chooses either to accept one of the contracts or not to accept any of them. If the trader rejects, he obtains his reservation utility, $\underline{u}_\vartheta$. I assume that $\underline{u}_B < \underline{u}_G$. For now, this difference is taken as an exogenous feature of the model: talented traders might have better career prospects in other sectors, or be more productive when self-employed. However, this difference can also be regarded as a reduced form of competition on the labour market for traders; this difference can stem either from heterogeneity in bank size (Thanassoulis, 2011) or quality (Bannier, Feess, and Packham, 2012), or from banks' limited access to dispersed managerial talent (Bijlsma, Boone, and Zwart, 2012). In the next subsection, I will give an extensive form game in which the size differences between banks endogenously lead to differing reservation utilities.

The bank is risk neutral and does not face any cash constraints. All traders have the same utility function over wealth with Bernoulli kernel $u(\cdot)$. I assume that $u(\cdot)$ is strictly increasing and weakly concave. None of the players discount future cash flows. I denote by $w(\cdot)$ the inverse of $u(\cdot)$.

Investment opportunities After accepting the contract, the trader chooses an investment. Following Diamond (1998), the model specifies a set of three different possible outcomes $\{L, M, H\}$, with $H > M > L = 0$. An investment X is a random variable taking values in this outcome set and can, as such, be represented by a triple $(\mathbf{P}(X = L), \mathbf{P}(X = M), \mathbf{P}(X = H))$. Bad traders can only invest in risk-free assets (I will generalize this later on), whereas good traders have an investment opportunity set that can be indexed by a single parameter σ, taking values in a compact nonnegative real interval Σ, with $0 \in \Sigma$. This investment opportunity set contains the assets

$$\left\{X_\sigma^G := (1 - \sigma - g(\sigma), g(\sigma), \sigma) : \sigma \in \Sigma\right\}.$$

The parameter σ can be interpreted as a risk parameter. Note, however, that σ does not denote the standard deviation of the outcomes. The function g is non-increasing and $g(0) = 1$, so that both types of traders have access to a risk-free asset. Furthermore $g''(\cdot) < 0$ and there is a σ inside Σ for which the expected pay-off $g(\sigma)M + \sigma H$ is maximized. Note that for this value σ^* the first order condition

$$g'(\sigma^*) + \frac{H}{M} = 0 \tag{1.1}$$

holds. This also easily allows the characterization of inefficiently risky assets ($\sigma > \sigma^*$) and inefficiently prudent ones ($\sigma < \sigma^*$). It can be useful to think of the outcomes "low", "middle" and "high" as being with respect to a benchmark. In that case the risk-free asset represents an investment replicating a benchmark index or a market portfolio, whereas a higher σ investment represents an active trading strategy.

Contracts Contracts can be based only upon the outcome of the trader's investment and not on the choice of σ. The non-contractibility of σ, and the unobservability of the trader's type, means that contracts need to be designed to serve two distinct purposes: to align the risk preferences of the principal with those of the agent, and to screen the types of traders. Any contract consists of three wage levels w_L, w_M and w_H. Throughout the subsequent analysis, I will characterize the contracts in terms of the corresponding utility levels u_L, u_M and u_H. The bank can offer any menu of contracts, as long as the null contract is part of this menu. However, the outcomes do not change substantially if the bank is only allowed to offer one contract besides the no-trade option.

Before analyzing the screening and risk incentive constraints, I make one additional assumption: before the outcome of the investment is realized, the trader can engage in wasteful trades that lower this outcome. This entails that the compensation contract always needs to be non-decreasing in the outcome of the trade, as otherwise the agent has an incentive to engage in wasteful behaviour. This monotonicity requirement translates into the two constraints $u_H \geq u_M$ and $u_M \geq u_L$.

Constraints As previously stated, the contract between the principal and the agent affects the agent's incentive to reveal his type, but also affects the risk he chooses. This imposes a number of constraints, which will be formulated on u_L, u_M and u_H. As the amount of risk σ cannot be observed and is freely chosen by the agent, the principal faces the agent's risk incentive constraint, namely that the level of risk $\tilde{\sigma}$ that the agent ultimately chooses satisfies his best response correspondence:

$$\tilde{\sigma} \in \underset{\sigma}{\operatorname{argmax}} \left\{ \sigma \left(u_H - u_L\right) + g(\sigma) \left(u_M - u_L\right) + u_L \right\}, \tag{1.2}$$

which translates to the following first order condition:

$$g'(\tilde{\sigma}) + \frac{u_H - u_L}{u_M - u_L} = 0. \tag{1.3}$$

In order to motivate the good trader to accept the contract, his participation constraint must be satisfied, which can now be expressed in terms of $\tilde{\sigma}$:

$$\tilde{\sigma} \left(u_H - u_L\right) + g\left(\tilde{\sigma}\right) \left(u_M - u_L\right) + u_L \geq \underline{u}_G. \tag{1.4}$$

In order to make sure bad traders do not accept employment, the reward they get from accepting employment and investing in the risk-free asset must be below their reservation utility, giving the non-participation constraint

$$u_M \leq \underline{u}_B. \tag{1.5}$$

The bad trader will then choose not to accept employment, but rather enjoy his outside option.

The trader's risk incentive constraint allows the condition for excessive risk taking to be written directly in terms of the convexity of the compensation contract. As $g'(\cdot)$ is decreasing, $\tilde{\sigma}$ is greater (smaller) than σ^* if and only if $g'(\tilde{\sigma})$ is smaller (greater) than $g'(\sigma^*)$, meaning that inefficiently risky assets are chosen if and only if

$$\frac{u_H - u_L}{u_M - u_L} > \frac{H}{M},$$

and inefficiently prudent ones if and only if

$$\frac{u_H - u_L}{u_M - u_L} < \frac{H}{M}.$$

Limited liability If the trader is protected by limited liability, this translates into a minimum level of utility β that must be provided, adding the constraint $u_L \geq \beta$. The value β represents the lowest possible utility the principal must make sure the agent obtains. One could think about this as the utility corresponding to a minimum wage or to no wage at all. Alternatively, one could think of this as the utility corresponding to the maximum punishment the bank can exert on the trader by firing him, damaging his reputation and career prospects, or even pressing legal charges, if that is feasible.

1.2.1 First Best

In case the trader types can be observed, the bank can simply reject bad traders and just offer a contract to good traders. If the bank pays the trader a flat wage, the trader is indifferent between all different types of investments and we can assume that he chooses the optimal asset $\sigma = \sigma^*$. In this case the moral hazard problem is irrelevant. The good trader earns his reservation utility and the bank earns the expected return on the optimal asset, minus the wage $w\left(u^G\right)$ she needs to pay the trader.

1.2.2 Without Limited Liability

As stated previously, when the types of the traders are unobservable, the bank needs to set wages in such a way that only a good trader accepts employment. In order to do so, the bank needs to set a low payment in case of a medium return on the investment. In order to attract the good trader, she can set a high reward for a high return on the investment, knowing that this only makes the contract more attractive to the good trader. If the trader is not protected by limited liability, the bank can punish a bad performance as severely as she wants, and thus has a means to discipline the trader into not taking excessive risk. Thus, the bank optimizes his expected profits

$$\tilde{\sigma}\left[H - v\left(u_H\right)\right] + g\left(\tilde{\sigma}\right)\left[M - v\left(u_M\right)\right] - \left(1 - g\left(\tilde{\sigma}\right) - \tilde{\sigma}\right) v\left(u_L\right),$$

subject to the good trader's best response correspondence

$$g'\left(\tilde{\sigma}\right) + \frac{u_H - u_L}{u_M - u_L} = 0.$$

The following constraints are needed to ensure the participation of the good trader, and the non-participation of the bad trader:

$$\begin{aligned} \tilde{\sigma}\left(u_H - u_L\right) + g\left(\tilde{\sigma}\right)\left(u_H - u_L\right) + u_L &\geq \underline{u}_G \\ u_M &\leq \underline{u}_B. \end{aligned}$$

Finally, the problem must satisfy the monotonicity constraints $u_H \geq u_M \geq u_L$.

Solving the bank's problem, the following result obtains.

Proposition 1.1. *If the trader is risk neutral, then*

- *the good trader obtains his reservation utility and the bad one does not accept employment,*
- *the compensation contract is linear in the sense that*

$$\frac{w_H - w_L}{w_M - w_L} = \frac{u_H - u_L}{u_M - u_L} = \frac{H}{M},$$

- *the chosen level of risk is equal to the optimal level of risk:* $\tilde{\sigma} = \sigma^*$.

If the trader is risk averse, then

- *the good trader obtains his reservation utility and the bad one does not accept employment,*
- *the chosen level of risk is lower than the optimal level of risk:* $\tilde{\sigma} < \sigma^*$.

The first part of this proposition is quite straightforward: since the trader is risk neutral, the bank does not need to insure the trader, and as she can freely set u_L, she can perfectly align the trader's incentives with her own. The second result is essentially a risk-sharing result: because of the screening, there needs to be a variable wage structure, but then the trader faces risk. The optimal solution weighs the trader's risk premium, which is a cost to the bank, against the profits that can be made from taking more risk.

1.2.3 With Limited Liability

If the trader is protected by limited liability, the bank cannot freely punish bad performance, making it harder to discipline the trader into not taking too much risk. As will be shown later on, in case good traders demand a much higher wage with respect to the limited liability level than bad traders do, the screening problem can lead to such a convex wage schedule that the bank can only be sure to hire the good trader if he lets him take excessive risk. In this case, the principal's optimization problem remains virtually the same, only with the added constraint that $u_L > \beta$. If $\beta \geq \overline{u}_B$ it becomes impossible to screen out bad traders. I will leave this possibility aside and focus on the cases in which $\beta < \overline{u}_B$.

One can derive sufficient conditions under which risk taking is excessive without solving the bank's problem: the participation and non-participation constraints, together with the limited liability constraint, impose a minimum convexity on the trader's rewards. From the good trader's participation constraint, we have

$$u_H - u_L \geq \frac{1}{\tilde{\sigma}} \left(\underline{u}_G - u_L - g\left(\tilde{\sigma}\right) \left(u_M - u_L\right)\right),$$

so that we can bound the trader's best response correspondence in the following way:

$$-g'\left(\tilde{\sigma}\right) = \frac{u_H - u_L}{u_M - u_L} \geq \frac{1}{\tilde{\sigma}} \left(\frac{\underline{u}_G - u_L}{u_M - u_L} - g\left(\tilde{\sigma}\right)\right). \tag{1.6}$$

As $\underline{u}_G > \underline{u}_B \geq u_M \geq u_L \geq \beta$, the above expression implies that

$$g\left(\tilde{\sigma}\right) - \tilde{\sigma} g'\left(\tilde{\sigma}\right) \geq \frac{\underline{u}_G - \beta}{\underline{u}_B - \beta}. \tag{1.7}$$

This allows me to prove the following sufficient condition for excessive risk taking by the trader.

Proposition 1.2. *If*

$$\frac{\underline{u}_G - \beta}{\underline{u}_B - \beta} > \frac{\mathbf{E}X^*_\sigma}{M}, \tag{1.8}$$

then $\tilde{\sigma} > \sigma^*$.

Proof. Note that

$$g(\sigma^*) - \sigma^* g'(\sigma^*) = g(\sigma^*) + \sigma^* \frac{H}{M} = \frac{1}{M} \mathbf{E}X^G_{\sigma^*}.$$

This means that if condition (1.8) holds,

$$g(\tilde{\sigma}) - \tilde{\sigma} g'(\tilde{\sigma}) \geq g(\sigma^*) - \sigma^* g'(\sigma^*),$$

which, because of the concavity of $g(\cdot)$, implies that $\tilde{\sigma} > \sigma^*$. QED

The intuition behind this proposition is that if the difference between the limited liability constraint and the bad trader's outside option is rather low, the difference between the compensation in case of a medium investment return and the compensation for a bad performance of the investment will be very small. If at the same time the good trader commands a very high wage, the bank needs to set a very high reward for a good performance to still make it attractive for him to accept employment. These two effects together will lead to such a convex wage schedule that the hired trader ends up taking excessive risk. Note that it is also implicitly assumed that it is better to hire the good trader and have him invest in the risky $\tilde{\sigma}$ asset, rather than to hire the bad trader and have him invest in the safe asset, i.e., that

$$\tilde{\sigma}\left(H - (w(u_H) - w(u_L))\right) + g(\tilde{\sigma})\left(M - (w(u_M) - w(u_L))\right) - w(u_L) \geq M - \underline{u}_B.$$

When analyzing the bank's optimization problem, one also finds that, in case of excessive risk taking, the bound on the convexity is sharp, as is stated in the following proposition.

Proposition 1.3. *If* $\frac{\underline{u}_G - \beta}{\underline{u}_B - \beta} \geq \frac{\mathbf{E} X^*_\sigma}{M}$, *the good trader takes a risk* $\tilde{\sigma}$ *such that*

$$g(\tilde{\sigma}) - \tilde{\sigma} g\prime(\tilde{\sigma}) = \frac{\underline{u}_G - \beta}{\underline{u}_B - \beta}.$$

The intuition behind the above lemma is that, as $\tilde{\sigma}$ is necessarily in the inefficiently risky region, raising risk will both make it more expensive for the bank to pay the risk-averse trader and make the expected return on the asset lower. Thus, it is optimal to have $\tilde{\sigma}$ as low as possible.

The two sides in condition (1.8) have interesting interpretations. The left hand side is the ratio of wage premia: the numerator specifies how much the good trader needs or wants to earn above the absolute minimum, while the denominator represents the same quantity for the bad trader. The right hand side of condition (1.8) is the expected gross revenue of the optimal asset, normalized by M. As M is also the expected revenue of the risk-free asset, the RHS of (1.8) can also be interpreted as the discounted expected return on the optimal risky asset. Alternatively, it represents the ratio between the expected revenues of the two traders' respective optimal assets.

1.2.4 Enlarging the Bad Trader's Investment Opportunity Set

Enriching the model slightly, I now assume that the bad trader has a larger set of investments available than just the risk-free one, i.e., both types have a set of available assets that can be indexed by a single parameter σ, taking values in a compact nonnegative real interval Σ_ϑ, with $\{0\} \subset \Sigma_B \subset \Sigma_G$. Again, the good trader has access to the set of assets

$$\left\{X^G_\sigma := (1 - \sigma - g(\sigma), g(\sigma), \sigma) : \sigma \in \Sigma_G\right\},$$

and the bad trader to

$$\left\{X^B_\sigma := (1 - \sigma - b(\sigma), b(\sigma), \sigma) : \sigma \in \Sigma_B\right\}.$$

Both functions $b(\cdot)$ and $c(\cdot)$ are decreasing and convex, with $b(0) = g(0) = 1$ and $b(\sigma) < g(\sigma)$ for all $\sigma \in \Sigma_B \setminus \{0\}$. This means that, for any positive level of risk, the good trader has strictly better investments available than the bad trader.

How difficult it becomes to select only the good trader depends on the bad trader's investment opportunity set: in order to screen out the bad trader, the bank now faces the non-participation constraint that, for all $\sigma \in \Sigma_B$,

$$\sigma\,(u_H - u_L) + b(\sigma)\,(u_M - u_L) + u_L \leq \underline{u}_B, \tag{1.9}$$

or, put differently,

$$\max_{\sigma \in \Sigma_B} \{\sigma\,(u_H - u_L) + b(\sigma)\,(u_M - u_L) + u_L\} \leq \underline{u}_B. \tag{1.10}$$

This maximization has a corner solution if $b'(0) + \frac{u_H - u_L}{u_M - u_L} \leq 0$, in which case it boils down to the old $u_M \leq \underline{u}_B$, meaning that a solution $(\tilde{\sigma}, u_H, u_M, u_L)$ to the previously-studied problem (with $\Sigma_B = \{0\}$) will still be the solution to the problem with the richer investment opportunity set as long as $b'(0) \leq g'(\tilde{\sigma})$.

If, however, $b'(0) \geq g'(\tilde{\sigma})$, the old solution no longer holds: the bad trader would have an incentive to accept the contract and invest in a risky asset with $\sigma > 0$. In that case, the bank's optimization problem is still to maximize

$$\sigma\,(H - (w(u_H) - w(u_L))) + b(\sigma)\,(H - (w(u_H) - w(u_L))) - w(u_L),$$

where the good trader's participation constraint

$$\sigma\,(u_H - u_L) + g(\sigma)\,(u_M - u_L) + u_L \geq \underline{u}_G$$

and the good trader's incentive constraint

$$g'(\sigma) + \frac{u_H - u_L}{u_M - u_L} = 0$$

still hold. The bad trader's non-participation constraint then becomes

$$\sigma_B\,(u_H - u_L) + b(\sigma_B)\,(u_M - u_L) + u_L \leq \underline{u}_B$$

where σ_B is such that

$$b'(\sigma_B) + \frac{u_H - u_L}{u_M - u_L} = 0.$$

Solving this problem is beyond the scope of this paper, but note that it imposes a stronger constraint on the convexity of the wage schedule. Indeed, deducting the bad trader's non-participation constraint from the good trader's participation constraint gives

$$(\tilde{\sigma} - \sigma_B)\frac{u_H - u_L}{u_M - u_L} + g(\tilde{\sigma}) - b(\sigma_B) \geq \frac{\underline{u}_G - \underline{u}_B}{u_M - u_L}.$$

Noting that the right hand side of the equation is greater than $\frac{\underline{u}_G - \underline{u}_B}{\underline{u}_B - \beta}$, and filling in the expressions from the incentive compatibility constraints gives

$$g(\tilde{\sigma}) - \tilde{\sigma} g'(\tilde{\sigma}) \geq \frac{\underline{u}_G - \underline{u}_B}{\underline{u}_B - \beta} + b(\sigma_B) - \sigma_B b'(\sigma_B).$$

As $b(\cdot)$ is convex, $b(\sigma_B) - \sigma_B b'(\sigma_B) > b(0) = 1$, so that the right hand side of this inequality is strictly greater than $\frac{u_G - \beta}{\underline{u}_B - \beta}$. This implies that the lower bound on the risk in this case is sharper than in the model with $\Sigma_B = \{0\}$: as it becomes harder to distinguish between the bad and the good trader, it becomes more likely that very high-powered incentives are needed to be sure that the good trader is hired. On the other hand, as should be noted, making the bad trader better will both raise the expected revenue the bad trader can earn for the bank and make it more expensive to hire the good trader. Thus, if the quality of traders is nearly indistinguishable, it becomes more attractive not to screen and just hire the "bad" trader.

1.3 Model: the Banking Labour Market

In order to analyze how the differences in reservation utilities arises endogenously, I model a very stylized labour market in which there are "small" and "large" banks. Again, every bank only needs one trader, but big and small banks differ in the size of the budget that this trader will be managing. In this context, an investment bank with considerable proprietary trading activity will be "bigger" than a commercial bank with a similar balance sheet size. The good trader's talent at making profitable investments will be more lucrative if the trader has a larger budget to invest. As a large bank gets more out of hiring the good trader than a small one, large banks will be willing to pay more. I will be focusing primarily on equilibria in which large banks hire good traders, and small banks hire bad traders. However, as small banks are also willing to pay more for good traders than for bad ones, good traders command a wage premium from the large banks. This wage premium can lead to excessive risk taking in much the same way as before.

The contracts, and the resulting levels of risk, are dependent upon the precise rules of the labour market. In order to study this, I provide two very stylized types of labour markets : a less flexible one and a more flexible one. In the less flexible labour market, all banks offer their contracts to the different traders; a trader simply observes all contracts on offer from the different banks, and chooses the one that he prefers, or chooses to exercise his outside option. In a more flexible labour market, a trader can, after being hired by one bank, still choose to go and work for a different bank. This means that, in an equilibrium in which large banks hire good traders, being hired by a large bank conveys a trader's quality. This makes it more attractive for small banks to hire good traders, as they no longer have to deal with screening out the bad traders when they make a predatory offer on a good trader. This further drives up the good trader's rent, imposing stronger constraints on the contract that the large bank has to offer the good trader. I will go on to show that this leads to higher risk taking.

Between the two versions of the labour market model, the only difference lies in the timing of the contract stage. In both versions, all traders have a reservation utility $\underline{u}$. There are $N+1$ banks, $i = 0, 1, \ldots, N$, with respective sizes $J := I_0 > I_1 = I_2 = \ldots = I_N =: I$. I assume that $I \geq \underline{u}$, so that it is profitable for the smaller banks to hire a bad trader to invest in the risk-free asset. There is a large number $T > N$ of bad traders and just one good trader. On the preferences and the investment opportunity sets of the traders I make some simplifying assumptions: the bad trader has only the risk-free asset available and the good trader has the set of assets

$$\left\{X_\sigma^G = (\sigma, g(\sigma), 1-\sigma-g(\sigma)) : \sigma \in \Sigma^G\right\}$$

available, where $g(\sigma) = 1 - \frac{\alpha}{2}\sigma^2$, for some positive α. I also assume $M = 1$ and $L = 0$. This means that $\sigma^* = \frac{H}{\alpha}$ and

$$\mathbf{E}X_G^{\sigma^*} = 1 + \frac{1}{2}\frac{H^2}{\alpha}.$$

Both traders are risk neutral, so that $u(\cdot)$ and $w(\cdot)$ are the identity.

1.3.1 The Less Flexible Labour Market

In order to model a less flexible labour market, I assume that a trader, once he has accepted a contract, will start working for that bank until his investment returns are realized and the game ends. At the contracting stage, all banks offer contracts (u_H, u_M, u_L) to all the traders, upon which these traders decide to accept or reject. If multiple traders accept the contract of a bank, a trader will be assigned at random. Note first of all that the fact that T is greater than N guarantees that all banks will employ a trader in equilibrium, as any bank that ends up without a trader would have an incentive to deviate by hiring an unemployed bad trader.

Now, assume that in this model equilibria exist in which the large bank hires the good trader and all the other banks hire the bad ones. In that case, there is a constraint for the large bank to offer the good trader such a high wage that the smaller banks prefer to hire a bad trader and have him invest in the risk-free asset. Thus, the large bank must offer an expected utility $\bar{u}$ to the good trader that makes it unattractive for smaller banks to lure away the good trader.

In order to see what this utility is, I first take any given outside option $\tilde{u}$ for the bad trader, and then see what a small bank's profit $V(\tilde{u})$ would be from offering a compensation package that offers a minimum utility of $\tilde{u}$. That is to say, if another bank were to offer a compensation package worth $\tilde{u}$ to the good trader, $V(\tilde{u})$ represents the maximum profit a small bank could get from topping that offer and hiring the good trader. Thus $V(\tilde{u})$ is defined as the solution to the following problem.

$$\max_{\sigma, u_H, u_M, u_L} \left\{ \left(\sigma H - L + \left(1 - \frac{\alpha}{2}\sigma^2\right)\right) I - \sigma\left(u_H - u_L\right) - \left(1 - \frac{\alpha}{2}\sigma^2\right)\left(u_M - u_L\right) - u_L \right\}$$

subject to

$$\begin{aligned}
& u_M \leq \underline{u} \\
& u_L \geq \beta \\
& \alpha\sigma = \frac{u_H - u_L}{u_M - u_L} \\
& \sigma\left(u_H - u_L\right) + \left(1 - \frac{\alpha}{2}\sigma^2\right)\left(u_M - u_L\right) + u_L \geq \tilde{u}.
\end{aligned}$$

As a small bank would still need to screen out the bad trader, in this optimization problem she is bound by the same constraints as before. The first constraint is the non-participation constraint that keeps bad traders from taking the contract; the second is a limited liability constraint. The third constraint is the incentive compatibility constraint that the trader chooses the optimal level of risk given the compensation package, and the fourth is an individual rationality constraint.

The individual rationality constraint is always binding. With respect to the other constraints, the problem allows corner solution or internal solutions, dependent upon the parameters of the problem: as before, the limited liability and non-participation constraints then impose a minimum convexity in the payment schedule for the trader. If these become binding, this means that the bank is restricted in the risk level it can let the trader take. If these constraints do not bind, the bank can freely give the trader the incentives to take the optimal investment decision and implement $\sigma = \sigma^*$. Solving this problem gives the following expression for the value function $V(\cdot)$:

$$V(\tilde{u}) = \begin{cases} \left(1 + \frac{H^2}{2\alpha}\right) I - \tilde{u} & \text{if } \frac{\tilde{u}-\beta}{\underline{u}-\beta} \leq 1 + \frac{H^2}{2\alpha} \\ \left(1 + H\sqrt{\frac{2}{\alpha}\frac{\tilde{u}-\underline{u}}{\underline{u}-\beta}} - \frac{\tilde{u}-\underline{u}}{\underline{u}-\beta}\right) I - \tilde{u} & \text{if } \frac{\tilde{u}-\beta}{\underline{u}-\beta} > 1 + \frac{H^2}{2\alpha}. \end{cases}$$

Now if the large bank wants to be sure that the small banks cannot hire the good trader, she must offer him an expected utility so high that the small banks would rather hire a bad trader and invest in a risk-free asset than try to top the good bank's offer. This means that the large bank needs to offer an expected utility $\bar{u}$ such that the small bank's expected profit from topping the offer, $V(\bar{u})$, is not greater than what she would obtain from hiring the bad trader, $V(\bar{u}) \leq I - \underline{u}$. The best $\bar{u}$ is the smallest one satisfying this criterion. As $V(\cdot)$ is decreasing and continuous, $\bar{u}$ solves the equation $V(\bar{u}) = I - \underline{u}$. This equation then gives the following expression for the rent that the good trader can earn:

$$\bar{u} - \underline{u} = \begin{cases} \frac{1}{2}\frac{H^2}{2\alpha} I & \text{if } \underline{u} - \beta \geq I \\ (\underline{u} - \beta)\left(\frac{2I}{\underline{u}-\beta+I}\right)^2 \frac{H^2}{2\alpha} & \text{if } \underline{u} - \beta < I. \end{cases} \tag{1.11}$$

Assuming $\beta > 0$, one has that $\underline{u} \geq I$ and the smaller banks would not be able to pay the bad trader to invest in the safe asset.

Now we can address the optimization problem for the large bank. In order not to lose the trader to a small bank, she needs to provide the good trader with at least $\overline{u}$. The large bank still needs to screen out bad traders, so that when she determines the wage schedule to optimize her own profits, she is restricted by the following constraints:

$$\begin{aligned} & u_M \leq \underline{u} \\ & u_L \geq \beta \\ & \alpha\sigma = \frac{u_H - u_L}{u_M - u_L} \\ & \sigma\left(u_H - u_L\right) - \left(1 - \frac{\alpha}{2}\sigma^2\right)\left(u_M - u_L\right) + u_L \geq \tilde{u}. \end{aligned}$$

In an exercise similar to the one in Section 1.2.3, one finds that the good trader takes excessive risk if

$$\frac{\overline{u} - \underline{u}}{\underline{u} - \beta} > \frac{H^2}{2\alpha}.$$

From expression (1.11), it follows that

$$\frac{\overline{u}-\underline{u}}{\underline{u}-\beta}=\left(\frac{2I}{\underline{u}-\beta+I}\right)^2\frac{H^2}{2\alpha},$$

which, because $I > \bar{u}-\beta$, is necessarily greater than $\frac{H^2}{2\alpha}$. This implies that the large bank can only be sure to hire the good trader if she offers a compensation package that induces excessive risk.

The next question is whether this strategy is profitable for the large bank, and especially whether it pays more than hiring a bad trader and investing in the risk-free asset. As the following lemma states, this is always the case.

Lemma 1.4. *If $\beta \geq 0$ and $I \geq \underline{u}$, the large bank hires the good trader and the small banks hire bad traders. The good trader subsequently takes an excessive risk of*

$$\sigma=\frac{2I}{\underline{u}-\beta+I}\frac{H}{\alpha}.$$

The intuition behind this lemma is that, if the small bank is indifferent towards hiring the trader, giving him $\bar{u}$ as well as letting him take excessive risk on the one hand, and hiring the small trader and having him invest in the safe asset on the other (which is the definition of $\overline{u}$), the large bank will be even more willing to hire the good trader and have him invest in a profitable, but excessively risky asset. This means that if the opportunities to punish the trader in case of a bad outcome are relatively limited, then excessive risk taking is an inherent feature of this model.

1.3.2 The Flexible Labour Market

I model a more flexible labour market by letting the banks offer their contracts sequentially to the different traders. As the hiring decision by the large bank is the most interesting in this context, I assume the banks offer their contracts in numerical order. If the trader accepts the contract, his employment at the large bank and his contract terms become verifiable information, but the trader keeps the real option to take the offer from another bank and renege on earlier offers.

I am still interested in equilibria in which the large bank gets the good trader. However, in these equilibria, upon knowing which trader has been hired by the good bank, other banks will rationally infer that this trader must be the good one. These banks can then offer this good trader a targeted contract, without needing to screen out bad traders. This entails that it becomes cheaper for the small banks to offer an undercutting contract, thus driving up the good trader's rent.

In this case, we can repeat the exercise we did before. However, now knowing the good trader's type, the small bank no longer faces the non-participation constraint. This means that, for any expected utility $\tilde{u}$ that the large bank offers, the small bank can obtain an expected revenue of $V(\tilde{u})$, which is defined as the solution to the following problem:

$$\max_{\sigma,u_H,U_M,u_L}\left\{\left(\sigma H+\left(1-\frac{\alpha}{2}\sigma^2\right)\right)I-\sigma\left(u_H-u_L\right)-\left(1-\frac{\alpha}{2}\sigma^2\right)\left(u_M-u_L\right)-u_L\right\}$$

subject to

$$\begin{aligned}
&u_L \geq \beta \\
&\alpha\sigma = \frac{u_H - u_L}{u_M - u_L} \\
&\sigma\left(u_H - u_L\right) + \left(1 - \frac{\alpha}{2}\sigma^2\right)\left(u_M - u_L\right) + u_L \geq \tilde{u}.
\end{aligned}$$

This problem is relaxed with respect to the one studied in the previous subsection: the constraint that screens out the bad traders is no longer present. In a sequential equilibrium in which the large bank gets the good trader, the other banks can infer with certainty which trader is the good one. Noticing that the participation constraint is binding, and that, in making the contract, there is freedom to choose u_H, u_M, and u_L in such a way that $\sigma = \sigma^*$, one finds that

$$V(\tilde{u}) = \left(1 + \frac{H^2}{2\alpha}\right) I - \tilde{u}.$$

Thus, in order to make sure that the smaller banks do not want to give a predatory offer to the good trader after he accepts employment, the large bank must offer the trader an expected utility of at least

$$\bar{u} := \min\left\{\tilde{u} : V(\tilde{u}) \leq I - \underline{u}\right\}.$$

This gives that

$$\bar{u} = \underline{u} + \frac{H^2}{2\alpha} I.$$

The good trader's rent is thus exactly the incremental profit he could earn the small bank, with respect to the bad trader.

In case the large bank wants to hire the good trader, she now faces the constraints of

$$\begin{aligned}
&u_L \geq \beta \\
&\alpha\sigma = \frac{u_H - u_L}{u_M - u_L} \\
&\sigma\left(u_H - u_L\right) + \left(1 - \frac{\alpha}{2}\sigma^2\right)\left(u_M - u_L\right) + u_L \geq \tilde{u}.
\end{aligned}$$

Again, all these constraints must bind at the optimum, giving $u_M = \underline{u}$, $u_L = \beta$ and u_H such that

$$\frac{1}{2}\alpha\sigma^2 = \frac{\overline{u} - \underline{u}}{\underline{u} - \beta} = \frac{H^2}{2\alpha}\frac{I}{\underline{u} - \beta}.$$

Again, the trader necessarily takes excessive risk. In this case, the bank will make an expected profit of

$$\left(1 + \frac{H^2}{2\alpha}\sqrt{\frac{I}{\underline{u} - \beta}}\left(2 - \sqrt{\frac{I}{\underline{u} - \beta}}\right)\right) J - \overline{u}.$$

In order to see if these equilibria are possible, it is necessary to compare the expression with the expected profit that the bank makes from hiring the bad trader. A quick look at this expression reveals that if $I > 4(\underline{u} - \beta)$, the above expression is smaller than $J - \overline{u}$, and the large bank is better off hiring a bad trader and obtaining a profit of $J - \underline{u}$. So, assume now that $\frac{1}{4}I \leq \underline{u} - \beta \leq I$. In that case, the large bank still needs to ensure that

$$\left(1 + \frac{H^2}{2\alpha}\sqrt{\frac{I}{\underline{u} - \beta}}\left(2 - \sqrt{\frac{I}{\underline{u} - \beta}}\right)\right) J - \overline{u} \geq J - \underline{u},$$

which translates to

$$\frac{H^2}{2\alpha}\sqrt{\frac{I}{\underline{u} - \beta}}\left(2 - \sqrt{\frac{I}{\underline{u} - \beta}}\right) J \geq \overline{u} - \underline{u}.$$

As $\overline{u} - \underline{u} = \frac{H^2}{2\alpha}$, the above expression simplifies to

$$J \geq \frac{\frac{\underline{u} - \beta}{I}}{2\sqrt{\frac{\underline{u} - \beta}{I}} - 1} I. \tag{1.12}$$

This gives the following lemma.

Lemma 1.5. *Assume* $\frac{1}{4}I \leq \underline{u} - \beta \leq I$. *If condition* (1.12) *is satisfied, a subgame perfect equilibrium exists in which the large bank hires the good trader, the small banks hire bad traders, and the good trader takes an excessive risk of*

$$\sigma = \frac{H}{\alpha}\sqrt{\frac{I}{\underline{u} - \beta}}$$

1.3.3 Comparative Analysis

In both types of labour markets, this paper finds excessive risk taking. Focusing on the equilibria in which the large bank hires the good trader, the risk taken in the less flexible labour market equals

$$\sigma^{LF} = \frac{2I}{\underline{u} - \beta + I}\frac{H}{\alpha} = \frac{2I}{\underline{u} - \beta + I}\sigma^*,$$

and the risk taken in the more flexible labour market equals

$$\sigma^F = \frac{H}{\alpha}\sqrt{\frac{I}{\underline{u} - \beta}} = \sigma^*\sqrt{\frac{I}{\underline{u} - \beta}}.$$

Note that both are increasing in $\frac{I}{\underline{u} - \beta}$: if I increases, i.e., when small banks are relatively big, they are willing to pay more for good traders. This drives up good traders' rent and thus the risk that the large bank must allow the good trader to take. Similarly, when $\underline{u} - \beta$ is very small, the reservation utility of bad traders is relatively low, and the limited liability constraint is relatively high, so that it becomes harder to screen out bad traders and punish bad performance at the same time.

Since $\underline{u} - \beta < I$ by assumption, it can be seen that $\sigma^F > \sigma^{LF} > \sigma^*$: excessive risk is taken in both types of labour markets, but more so in the flexible one than in the non-flexible one. In the more flexible labour market, it becomes more attractive for the small bank to make a predatory offer to the good trader, thus driving up the good trader's rent. This observation carries the strongest regulatory consequence: making the labour market more flexible will drive up the rents of good traders, and thus lead to higher risk. A regulator intent on reducing risk in the financial sector could thus introduce measures to make the labour market *less* flexible. All of the equilibria above are constrained efficient, so that taxes or bonus caps cannot make the outcome more efficient; however, a regulator could intervene by introducing tighter regulation in the financial labour market.

1.4 Conclusion

In this paper, I find that incomplete information on trader skill can lead to excessive risk taking: good traders demand a wage premium, and the only way banks can separate them from bad traders is by offering a high upside bonus. This leads to overly high-powered incentives, and thus to excessive risk taking. The wage premium that leads to excessive risk taking can be modeled as an exogenous feature, arising from better outside options for good traders, but can also arise as an endogenous result of competition on a labour market. The nature of the labour market has an impact on the level of risk that traders end up taking: in more flexible labour markets, the threat of traders leaving for other firms raises the price of good traders, so that they need to be given more high-powered incentives than in less flexible labour markets. It can thus be argued that regulators have the option of making labour markets less flexible in order to reduce excessive risk taking by large financial institutions.

A meaningful welfare analysis is not possible in this model, as only banks and their workers are present in the model. This model therefore overlooks two types of externalities. First of all, the bank's risk taking can impose a negative externality on bondholders and depositors, as well as, for the more systemically important banks, on the financial system. The other type of externality is on the real sector: if banks hire very talented workers, these workers will not be productive in the real sector. A more complete analysis of welfare should take this into account. This model, if depositors or financial contagion are introduced, could be used to study whether intervening in bankers' pay provides an efficient policy tool to mitigate the first type of externality.

It deserves further study whether this model, especially that of the labour market, can still be analyzed in a tractable manner if preferences are generalized. It would be especially interesting to see to what extent the excessive risk taking result would persist with risk-averse traders.

In this model, investment technologies and prices have been treated as exogenous. An interesting question would be to see how the pay-off structure of bankers influences asset prices. The framework in this paper could be adapted to involve an asset market, in order to see whether certain stylized facts of financial markets can be explained by bankers' incentives.

Chapter 2

Credit Market Competition and Liquidity Provision

CO-AUTHORS: FABIO CASTIGLIONESI AND FABIO FERIOZZI[1]

> They clashed together when they met and then at that point each turned about and rolled his weight back again, shouting: 'Why hoard?' and 'Why Squander?'
>
> Dante Alighieri, *La Divina Commedia, Inferno, Canto VII*[2]

Cash holdings by firms have seen a dramatic increase over the past decades (Bates, Kahle, and Stulz, 2009; Duchin, 2010). At a practical level, the propensity of firms to hold on to their liquid resources is puzzling since there are no clear reasons as to why firms would find it convenient to sit on an increasing pile of low-return cash. In the theoretical literature, the incentives to hold (or provide) liquidity have been studied mainly within the context of exclusive contracting relationships. That is, the borrower can only trade with one lender, either a monopolist or one from a competitive pool of lenders. In this framework, an exogenous change in competition (from a monopolist lender to a competitive one, or vice versa) does not affect firm liquidity holdings.

This paper tries to approach the issue of liquidity provision to firms in a novel manner. It considers the role of non-exclusive relationships in a generic model of investment and liquidity à la Holmström and Tirole (1998). If contracting between firms and lenders is exclusive, it is optimal for a lender to guarantee a limited provision of liquidity to the firm to provide for cost shocks that might arise before the firm's project matures. However, if it is possible for a firm to privately contract with several investors at the same time, other investors could expand this liquidity

[1]Fabio Castiglionesi is currently an associate professor at the Finance Department at Tilburg University in Tilburg, The Netherlands. Fabio Feriozzi currently works as an assistant professor of finance at IE Business School in Madrid, Spain.

[2]Translation D. Sinclair. In the original, these lines read
"Percoteansi 'ncontro; e poscia pur
si rivolgea ciascun, voltando a retro,
gridando: "Perche tieni?" e "Perche burli?""

provision by free riding upon the provision given by incumbent lenders. This means that it becomes harder, if not impossible, to limit the liquidity provision to the firm.

In Holmström and Tirole (1998), our main benchmark, a firm has a profitable project, but will face a stochastic liquidity shock before the project matures. This liquidity shock takes the form of a cost that needs to be paid in order for the project to be carried out and its payoff realized. The firm will contract with investors to obtain the resources needed for the initial investment, but will also secure a liquidity guarantee. Under exclusive contracting, the size of this liquidity provision is the one that optimally weighs the potential cost that needs to be paid against the benefit of being able to carry the project to maturity.

Contracting the liquidity provision ex ante is optimal because the firm has only a limited capacity of refinancing due to a moral hazard problem that causes a wedge between pledgeable income and total income. The constrained first-best allocation can be implemented by contracting a finite maximum amount of liquidity that the firm can withdraw, and a flat repayment to the lender. The latter is needed to guarantee the firm a large enough share of its project's final proceeds in order to make sure the firm exerts effort.

When competition between lenders is assumed to be non-exclusive, i.e., when lenders are not able to verify the relations between the firm and other lenders, a problem with the finite liquidity provision arises. As the liquidity needs of the firm are private information (as in Holmström and Tirole, 2011), the exclusive competition allocation provides an incentive for outside lenders to deviate. An outside lender can offer some additional, and optional, liquidity provision to the firm. If the firm accepts, and finds itself needing the additional liquidity, it can privately understate its liquidity needs to the incumbents and use the additional liquidity from the outside lender. If the additional liquidity provision is needed, the deviating lender can receive a payment from the firm's original incentive share.

This deviation means that the original Holmström and Tirole (1998) allocation can no longer be sustained. As a next step, we try to characterize the equilibria that do arise when competition is not exclusive. In particular, assuming outside lenders have the possibility to fully dilute the original lender, we show that an equilibrium can be found in which an incumbent lender provides unlimited liquidity support to the firm. This makes it impossible for the outside lenders to free ride, and it allows the incumbent to regain exclusivity. To motivate our analysis, we briefly document how since the late 1960s it has become increasingly common for firms to have several banking relationships (see, among others Braggion and Ongena, 2011). As we try to argue in this paper, this historical shift towards non-exclusivity could have consequences for the economy as a whole. As a particular example, consider the United Kingdom. After the deregulation in the British banking market, there has been a rise of multiple banking relationships from the late 1960s through the early 1980s (Braggion and Ongena, 2011). This rise is accompanied by an increase in firms' aggregate cash holdings, as shown in Figure 2.1.

The results in this paper point towards a channel through which a shift from exclusivity towards non-exclusivity could induce a higher liquidity provision (and therefore possibly higher liquidity holdings). It might be necessary for lenders to provide a firm with more liquidity in order to ward off competition from outside

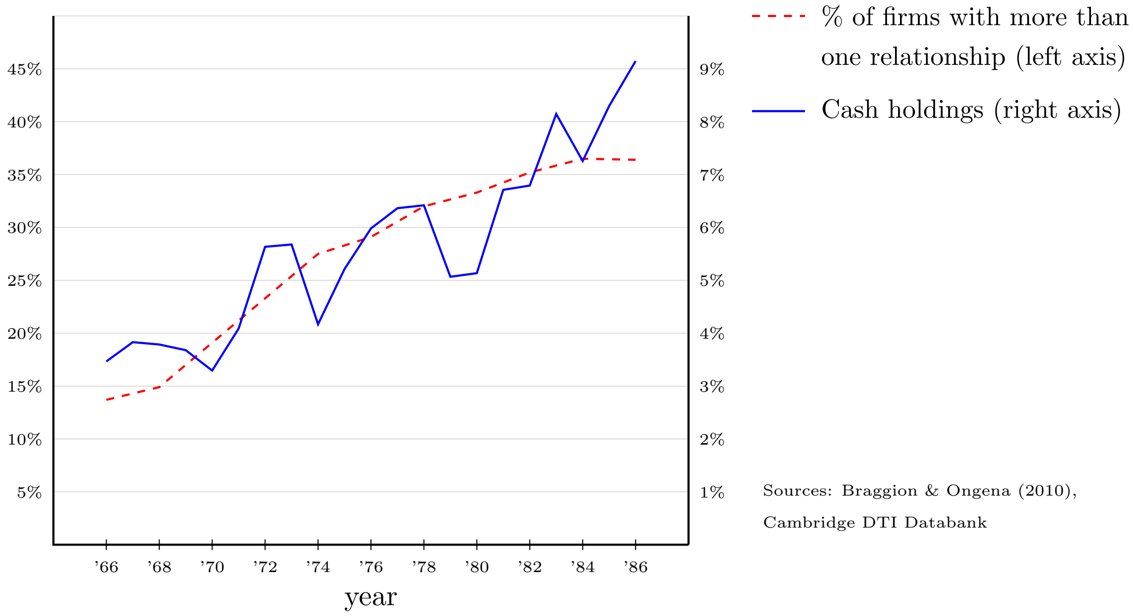

Figure 2.1: Graph of cash holdings and relationship multiplicity in the UK

lenders. Clearly, more evidence is needed, but this is the basic observation that motivates our model.

The rest of this paper is as follows. Section 2.1 reviews the literature. Then, section 2.2 presents the model. Section 2.3 describes the allocation in autarky, that is, when there are no outside investors. Section 2.4 solves the model for the benchmark allocation under exclusive competition (Section 2.4.1) and monopoly (Section 2.4.2). Section 2.5 presents the analysis under the assumption of non-exclusive competition, where we show how the benchmark allocation under exclusive competition is not sustained in equilibrium (Section 2.5.1) and we characterize the alternative equilibrium allocation (section 2.5.2). Section 2.6 concludes. Proofs for this chapter are in Appendix A. Appendix B shows how the main results still hold if the model considers a game with a finite, but large enough, number of players.

2.1 Literature

A growing literature describes the inefficiencies caused by non-exclusivity in the presence of adverse selection (Beaudry and Poitevin, 1995; Attar, Mariotti, and Salanié, 2011, 2014) or moral hazard (Bizer and DeMarzo, 1992; Kahn and Mookherjee, 1998; Parlour and Rajan, 2001; Bisin and Guaitoli, 2004; Attar and Chassagnon, 2009; Attar, Casamatta, Chassagnon, and Décamps, 2010; Castiglionesi and Wagner, 2013). The literature so far has mainly focused on firm solvency issues, while our focus is on firm liquidity provision.

A common finding in the literature is that non-exclusivity leads to less lending and trading, and in particular non-exclusivity brings under-insurance (Kahn and Mookherjee, 1998) (Bisin and Guaitoli, 2004). This paper instead presents an example in which non-exclusivity induces over-insurance: lenders are willing to provide unlimited liquidity support to prevent free riding by other lenders. Lenders can free ride on one another's liquidity provision, thus extending the liquidity supply to the firm. An unlimited liquidity support thus serves to regain exclusivity. As far as we are aware, this is the first example of over-insurance under non-exclusivity.

In much of the literature, liquidity or cash holdings are explained using practical, and firm-specific, reasons, giving cross-sectional predictions about differences in liquidity holdings. Our paper instead explains changes in liquidity holdings through the overall nature of competition, which is something that typically happens *outside* the firm: this could give a possible explanation of the fact that build-ups in liquidity often seem to happen across the board, rather than at the level of individual firms.

2.2 Model

The model is a non-exclusive competition equivalent to the model of liquidity demand in a generic lender-borrower relationship (subject to moral hazard) as introduced by Holmström and Tirole (1998), and extensively treated in Tirole (2006, Section 5.3) and in Holmström and Tirole (2011, Chapter 2). The model describes a three-period production economy. There is a single good used for both consumption and investment, which will be referred to as "money". There are two kinds of players: firms and lenders. All players are risk neutral and do not discount future consumption.

Firm and projects. There is a firm, protected by limited liability, with a finite endowment A at $t = 0$. The firm has access to a project, with constant returns to scale, that can be either successful or unsuccessful. The success of the project is revealed at $t = 2$. For an investment of I units of money at $t = 0$, the project yields RI at $t = 2$ when successful and 0 when unsuccessful.

The firm can influence the success probability of the project by privately choosing a level of *effort* $e \in \{L, H\}$. If the firm chooses low effort $e = L$, it obtains a private benefit B and the project will have a success probability equal to $p_L > 0$. If the firm chooses high effort $e = H$, there will be no private benefit, but the project will have a higher success probability equal to p_H. We have $0 < p_L < p_H \leq 1$. We denote the difference in the success probabilities by $\Delta p = p_H - p_L$.

The present value of the project (conditional upon continuation) when the firm exerts high effort is defined as

$$\rho_1 := p_H R. \tag{2.1}$$

This present value gives the total expected surplus that can be generated by the project if it is continued after $t = 1$. In order to induce the firm to work, its share of the final project's return has to be higher than the potential benefits of shirking, i.e., given an up-front investment I and a repayment D to the lenders, we need

$$p_H(RI - D) \geq p_L(RI - D) + BI. \tag{2.2}$$

This constraint allows us to define the *expected pledgeable income* (conditional upon continuation) as the maximum income per unit of investment that can be pledged to lenders without violating incentive compatibility:

$$\rho_0 := p_H \left(R - \frac{B}{\Delta p} \right). \tag{2.3}$$

This pledgeable income is the maximum amount, in expected value and conditional upon continuation, that an investor can expect to receive from the firm. This amount will be useful when analyzing the possibility of credit rationing. In terms of this pledgeable income, the incentive compatibility constraint can be rewritten as

$$p_H D \leq \rho_0 I. \tag{2.4}$$

The liquidity shock. At $t = 1$, before the effort choice is made, there is a reinvestment stage at which the firm observes an exogenous liquidity shock ϑ. This liquidity shock is a nonnegative random variable drawn from a continuous, atomless and strictly increasing distribution $F(\cdot)$ with density $f(\cdot)$. A liquidity shock ϑ, when the investment size equals I, means that a total cash injection of ϑI is needed in order to continue the project. If this cash cannot be raised, the firm is liquidated, yielding zero. If this cash can be raised, the project will continue. We assume that there is no possibility of continuation at a smaller scale.

Lenders. Lenders have an unlimited amount of money at $t = 0$ and $t = 1$ and can provide money to the firm to finance investment in the project. Throughout the main text, in order to develop the most important intuitions, we treat the lending sector as subject to free entry: contracts are offered simultaneously, but we assume that in any equilibrium, there is at least one inactive lender. Appendix B formally describes a game with a finite number of lenders.

Contracts. The contracting stage occurs at $t = 0$. In this stage each lender can offer a contract to the firm. The contract C_i is an array $(J_i, L_i, D_i(\cdot))$, that specifies three elements:

- the up-front amount J_i to finance the *initial* investment in the project;
- the liquidity policy function $L_i(\cdot)$ that determines the liquidity provision at $t = 1$. Upon learning the shock ϑ, the firm sends a message $m_i(\vartheta)$ about the liquidity shock to each lender $i \in S$ and the function $L_i(\cdot)$ maps each message into an amount of liquidity that lender i provides to the firm;
- the repayment function $D_i(\cdot)$ that is a function of the message m_i that the firm sends to lender i.

The firm can send different messages to each of the lenders and each message m_i can not be seen by anyone but the firm and lender i. The strategy of the firm at $t = 1$ is thus a vector of messages in the space $\mathcal{M}^S$. It is assumed that the space of messages $\mathcal{M}$ is of the same dimension as the space of potential liquidity shocks (i.e., $\mathcal{M}$ is one-dimensional). As the lender can only respond to the message with the amount of liquidity provision and the repayment amount, any more dimensions added to the message space would be redundant. One can think of the message as a report of ϑ or just a demand for an amount of liquidity, i.e. $L_i(m_i) = m_i$.

The set of lenders with whom the firm trades is called the set of *active* lenders and denoted by S. The free-entry assumption translates into the assumption that no matter how large S, there is always an inactive lender who can offer a contract on top of those that have already been traded. The choice of S determines, among others, the firm's total aggregate investment size $I = A + \sum_{i \in S} J_i$. It is assumed that the total amount I, i.e. all the money borrowed from the lenders, has to be invested in the firm's project.

We assume that the firm does not have any consumption or investment opportunities at $t = 1$, so that any liquidity withdrawn at $t = 1$ in excess of the firm's liquidity need is wasted. This justifies the assumption made in the Holmström and Tirole (1998); Holmström and Tirole (2011) and Tirole (2006) models that firms do not withdraw more liquidity than needed, even if incremental cash at $t = 1$ is sold at a very low marginal price. In case of a continuous distribution of shocks, and a non-decreasing price of liquidity, this entails that the firm does not withdraw more liquidity than needed:

$$\sum_{i \in S} L_i(m_i) \in \{0, \vartheta I\} . \tag{2.5}$$

For continuous shock distributions one can assume, to the same effect, that, although the liquidity needs cannot be observed, the post-shock holdings of the firm can be. In that case, lenders can punish the firm for having a liquidity position that is too high. The post-shock cash holdings equal $\sum_{i \in S} L_i(m_i) - \vartheta I$ in case $\max_{(m_i)_i} \left\{\sum_{i \in S} L_i(m_i)\right\} \geq \vartheta I$ and $\sum_{i \in S} L_i(m_i)$ in case $\max_{(m_i)_i} \left\{\sum_{i \in S} L_i(m_i)\right\} < \vartheta I$. Therefore the lenders can prevent the firm from withdrawing more than it needs, so that the firm faces the additional constraint (2.5). In other words, the firm can either withdraw exactly the amount it needs, if the liquidity facilities are sufficient, or not withdraw any liquidity.

As argued in Holmström and Tirole (1998); Holmström and Tirole (2011), and Tirole (2006, Section 5.3), the specific form that these contracts can take may vary. A straightforward interpretation is a combination of start-up financing and a credit line. However, these contracts can also be interpreted as a combination of initial financing and an amount of cash that is provided up front. This up-front cash can be accompanied by covenants and provisions on the level of cash holdings.[3] Of course, a combination of credit lines and cash provided up front is also possible.

It is not possible for the lenders to write any contingency upon the contractual relationship the firm has with other lenders. This makes it more complicated to define concepts like seniority, or to differentiate between debt, equity, and hybrid claims. Throughout most of the analysis it is assumed that lenders write plain debt contracts. The firm's cut of the project's proceeds takes the form of an equity claim.

Timing. The firm's effort choice comes after the reinvestment stage. The timeline of the model can be seen in Figure 2.2.

[3]The standstill agreement between Kirk Kekorian and Chrysler in February 1996 shows that cash holdings can be limited contractually and court-enforced.

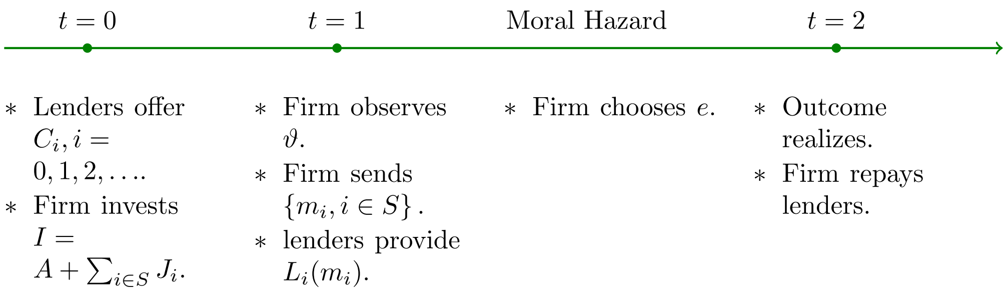

Figure 2.2: Timeline for the basic model

Parameter assumptions. First of all, we assume that there are gains from trade: the project, given that the firm exerts high effort, has a positive ex ante net present value (NPV) , i.e., there exists a $\widetilde{\vartheta}$ such that

$$F(\widetilde{\vartheta})\rho_1 > 1 + \int_0^{\widetilde{\vartheta}} \vartheta f(\vartheta) \mathrm{d}\vartheta. \tag{2.6}$$

For a liquidity level $\tilde{\vartheta}$, the left hand side represents the expected total surplus generated per unit invested, whereas the right hand side represents the total cost of the project. For all $\widetilde{\vartheta}$ we have

$$F\left(\widetilde{\vartheta}\right)\rho_0 < 1 + \int_0^{\widetilde{\vartheta}} \vartheta f(\vartheta) \mathrm{d}\vartheta. \tag{2.7}$$

This means that, for any level of liquidity, the total cost of the project is always greater than the total expected amount that can be pledged to outside investors. This means that there is credit rationing. Otherwise, the firm would be able to borrow an infinite amount of funds for its project. Another assumption is that, if the firm exerts low effort, the total NPV of the project will never be positive. That is, for all $\widetilde{\vartheta}$,

$$F(\widetilde{\vartheta})\,(p_L R + B) < 1 + \int_0^{\widetilde{\vartheta}} \vartheta f(\vartheta) \mathrm{d}\vartheta. \tag{2.8}$$

It will be useful to define the *expected unit cost of effective investment* $c(\cdot)$ as

$$c\left(\widetilde{\vartheta}\right) := \frac{1 + \int_0^{\widetilde{\vartheta}} \vartheta f(\vartheta) \mathrm{d}\vartheta}{F(\widetilde{\vartheta})}.$$

This function weighs the costs and benefits of providing more liquidity. The numerator gives the expected cost (per unit of investment) if liquidity is provided up to $\tilde{\vartheta}$. The denominator gives the benefit of providing liquidity: the probability of being able to finance the liquidity shock and having the project yield a return. As a general property, for any continuous distribution $F(\cdot)$, the function $c(\cdot)$ has a unique local and global minimum at ϑ^*, where $\vartheta^* = c(\vartheta^*)$ and has two asymptotes:

$\lim_{\tilde{\vartheta}\to 0} c\left(\tilde{\vartheta}\right) = \infty$ and $\lim_{\vartheta\to\infty} c\left(\tilde{\vartheta}\right) = 1 + \mathbf{E}\vartheta < \infty$. Using these properties of $c(\cdot)$ allows us to re-state assumptions (2.6) and (2.7) in the following concise form:

$$\rho_0 < c(\vartheta^*) < \rho_1.$$

A typical graph of the function $c(\cdot)$ is shown in Figure 3.1.

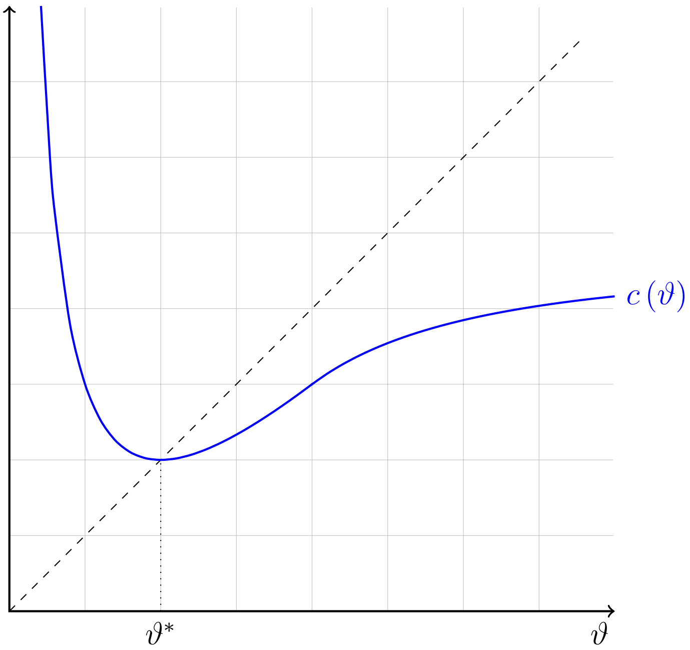

Figure 2.3: Typical shape of the expected unit cost of investment function

Aggregate allocation. In order to analyze the aggregate quantities, we indicate with $\overline{\vartheta}$ the maximum liquidity shock that the firm can finance at $t = 1$. Note that it is possible for $\overline{\vartheta}$ to equal infinity. The interpretation of $\overline{\vartheta}$ is the total liquidity position of the firm. This is the combination of available cash and the total of available credit lines. For each $\vartheta \in [0, \overline{\vartheta}]$, we define the corresponding total debt load as

$$D(\vartheta) := \min_{\{m_i : i\in S\}} \left\{ \sum_{i\in S} D_i(m_i) : \sum_{i\in S} L_i(m_i) = \vartheta I \right\}.$$

The highest level of liquidity that can still be secured in exchange for an incentive-compatible repayment is denoted by

$$\underline{\vartheta} := \max\{\vartheta : p_H D(\vartheta) \leq \rho_0 I\}.$$

If $p_H D(\vartheta) \leq \rho_0 I$ everywhere, we define $\underline{\vartheta} = \overline{\vartheta}$.

2.3 Autarky

We begin the analysis under the assumption that the credit market is absent. Studying the autarky case serves to determine the firm's reservation utility. In autarky, the firm has two choices: either consume his endowment at $t = 0$, or invest part of his endowment, keeping the rest in cash in order to insure the project. In the first case, the total consumption equals the endowment A. In the second case, if the firm invests a fraction γ of its own endowment A to start up the project, storing $(1-\gamma)A$ in cash to finance the possible liquidity shock in the intermediate period, the firm is insured as long as $\vartheta < \frac{1-\gamma}{\gamma}$. This gives the firm a total consumption of

$$\gamma F\left(\frac{1-\gamma}{\gamma}\right)\left[\rho_1 - \int_0^{\frac{1-\gamma}{\gamma}} \vartheta f(\vartheta)\,\mathrm{d}\vartheta\right] A + (1-\gamma)\,A.$$

By choosing to invest a fraction γ of its endowment and keeping a fraction $1-\gamma$ in cash, the firm is self-insured against liquidity shocks up to $\vartheta_A := \frac{1-\gamma}{\gamma}$. The firm's decision whether or not to invest is dependent upon the profits it can obtain by investing, as is stated in the following proposition:

Proposition 2.1. *Under autarky, the firm invests a fraction $\gamma A = \frac{A}{1+\vartheta_A}$ of its endowment, where ϑ_A satisfies the first order condition*

$$F(\vartheta_A)\,\rho_1 - \left(1 + \int_0^{\vartheta_A} \vartheta f(\vartheta)\,\mathrm{d}\vartheta\right) = (1+\vartheta_A)\, f(\vartheta_A)\,(\rho_1 - \vartheta_A)\,, \tag{2.9}$$

giving the firm a total consumption of

$$U_A := \left(\gamma\left(F(\vartheta_A)\,\rho_1 - \int_0^{\vartheta_A} \vartheta f(\vartheta)\,\mathrm{d}\vartheta\right) + (1-\gamma)\right) A.$$

Regarding the part that is invested in the project, there is a level of γ for which it is possible to obtain a revenue greater than one. This means that the firm will always choose to invest at least part of its endowment in the project. On the other hand, the firm must also keep at least some part of its endowment as a cash buffer, as otherwise the project can never be realized. This means that choosing $\gamma = 1$ or $\gamma = 0$ is never optimal and the firm's problem always has an interior solution. The firm's autarky consumption U_A serves as a reservation utility: in any equilibrium allocation, the firm must receive at least U_A, as otherwise it will be better off not trading, but instead investing on its own.

2.4 Benchmark Allocations

We now characterize the two "exclusive" benchmark allocations. The first case can be thought of as the "classic" competitive market and corresponds exactly to the baseline solution in Holmström and Tirole (1998). The second consists of a monopoly, in which the lending sector consists of only a single lender, who can make a take-it-or-leave-it offer to the firm.

2.4.1 Exclusive Competition

In this section, we study the case of exclusive competition. Any lender can ex ante enforce his contract to be exclusive, precluding the borrower from trading with other lenders. Alternatively one can assume that there exists some exogenous factor that constrains the borrower to pick only one contract from the contracts offered. In this case, whenever a lender offers a profitable contract, another lender can undercut this contract. This entails that, in equilibrium, lenders earn zero profits and the profits for the firm are maximized. Thus, this case corresponds to the "classic" notion of competition, and to the baseline case studied in Holmström and Tirole (1998). The following proposition characterizes the exclusive competition allocation. Temporarily allowing for the results that the firm's repayment is constant and incentive compatible, the firm's total gross surplus is characterized by the following proposition:

Proposition 2.2. *The allocation under exclusive competition is characterized by an investment size*

$$I^C = \frac{A}{F(\vartheta^*)(\vartheta^* - \rho_0)},$$

a flat incentive-compatible repayment such that $p_H D^C(\vartheta) = \rho_0 I^C$ *for all* ϑ*, and a maximum liquidity provision* $\vartheta^C = \vartheta^*$*.*

The complete proof is provided in Appendix A. The explanation of the result is as follows. Because of competition, the total surplus of the project goes to the firm and must be maximized. First of all, as the surplus generated under low effort is lower than that generated under high effort, there is no reason to have any state of the world in which the incentive compatibility constraint is violated and in which the firm provides low effort. Furthermore, as a higher repayment raises the financing capacity of the lender, and the surplus increases with investment size, the incentive compatibility constraint must be binding. To see why ϑ^C must equal ϑ^*, consider the the optimal allocation, characterized by an up-front investment $I^C = A + J^C$, a constant repayment D^C and a maximum liquidity provision ϑ^C

$$p_H F\left(\vartheta^C\right)\left(RI^C - D^C\right),$$

which has to be maximized subject to the lender's break-even constraint

$$\left(1 + \int_0^{\vartheta^C} \vartheta f(\vartheta)\,\mathrm{d}\vartheta\right) I^C - A \leq F\left(\vartheta^C\right) p_H D^C$$

and the incentive compatibility constraint for the firm

$$D^C \leq \left(R - \frac{B}{\Delta p}\right) I^C.$$

With both constraints binding, one can reduce the optimization to one over the the total level of liquidity. As can be seen from the resulting expression for total surplus

$$\frac{\rho_1 - c\left(\vartheta^C\right)}{c\left(\vartheta^C\right) - \rho_0} A,$$

raising the liquidity level has two effects: on the one hand, it raises the probability of being able to finance the liquidity shock and thus augments the expected return per unit invested. On the other hand, it raises the total expected cost per unit of investment, thus tightening the financing capacity of the lender and the total project size that can be financed. The optimal level of liquidity is thus the one that ideally weighs the costs of providing insurance against the benefits. Indeed this happens when $c\left(\vartheta^C\right)$ is minimized, i.e., when $\vartheta^C = \vartheta^*$. Because ϑ^* minimizes $c(\cdot)$, by assumption (2.6), there are gains from trade. This means that the firm is always better off with an outside investor: under autarky the firm always has at least part of its capital remain unproductive. This means that the individual rationality constraint of the firm is always satisfied under the exclusive competition allocation.

2.4.2 Monopoly

When the lending sector is monopolistic, there is only one lender, who makes a take-it-or-leave-it offer to the firm. In this case, the lender will maximize its profits, subject to the firm's incentive compatibility and individual rationality constraints. The result is very similar to the one under exclusive competition:

Proposition 2.3. *The allocation under monopoly is characterized by an investment size*

$$I^M = \frac{U_A}{F(\vartheta^*)(\rho_1 - \rho_0)},$$

a flat incentive-compatible repayment $D^M = \left(R - \frac{B}{\Delta p}\right) I^M$*, and a maximum liquidity provision* $\vartheta^M = \vartheta^*$.

The intuition is as follows. By a similar argument as in the previous paragraph, the repayment is constant and incentive compatible everywhere. The monopoly contract can thus be characterized by an up-front investment $I^M = A + J^M$, a repayment D^M, and a maximum liquidity provision ϑ^M. The monopolistic lender maximizes the expected profits

$$F(\vartheta^M)p_H D - \left(\left(1 + \int_0^{\vartheta^M} \vartheta f(\vartheta)\mathrm{d}\vartheta\right) I - A\right), \tag{2.10}$$

subject to the firm's individual rationality constraint

$$p_H F(\vartheta^M)(RI^M - D^M) \geq U_A, \tag{2.11}$$

and the firm's incentive compatibility constraint $D^M \leq \left(R - \frac{B}{\Delta p}\right) I^M$. As was shown in the previous paragraph, there is always an allocation that is individually rational for both the lender and the firm, so that the lender can obtain nonnegative profits, while satisfying both the individual rationality and the incentive compatibility constraint of the firm. With similar arguments as in the previous paragraph, the incentive compatibility constraint is always binding. So is the firm's individual rationality constraint, as the lender does not want to leave any surplus to the firm on top of its reservation utility. Plugging these two constraints into expression (2.10), the following expression for the lender's profits obtains:

$$A - \frac{c(\vartheta^M) - \rho_0}{\rho_1 - \rho_0} U_A.$$

This expression is decreasing in $c\left(\vartheta^M\right)$ and therefore the optimal level of liquidity provision under monopoly minimizes the function $c(\cdot)$, so that $\vartheta^M = \vartheta^*$. That is, the monopolistic lender efficiently weighs the costs and benefits of providing liquidity to the firm. Comparing the allocation under exclusive competition to the monopoly allocation, it can be noted that both the repayment as a fraction of total revenues and the relative liquidity level are the same. However, comparing the investment sizes, one can see that the investment size under monopoly is strictly smaller than the one under exclusive competition: $I^M < I^C$. The intuition behind this is as follows: the lender can obtain at most the pledgeable income from the project and has to supply the total cost of the project (minus the firm's fixed endowment). Then, by the credit rationing assumption (2.7), the costs per unit invested are higher than the revenue that the lender can obtain per unit invested. This means that the lender's profits are decreasing in investment size. However, in any allocation, the firm receives an amount at least equal to its potential private benefits, meaning that the firm's utility is increasing in total investment size. Thus, under monopoly, when the lender's profits are maximized, the project size must be as small as possible, and under exclusive competition, where the firm's profit are maximized, the investment size must be as large as possible.

2.5 Non-Exclusive Competition

We now leave the situation in which competition is simply assumed in reduced form, as in Section 2.4.1. Rather we choose to model competition explicitly as a non-exclusive phenomenon: firms can privately choose to contract with multiple lenders. First, we show that, once competition is assumed to be non-exclusive, the benchmark allocation as derived under exclusive competition is no longer sustained in equilibrium. Second, we characterize an alternative equilibrium in such an environment.

2.5.1 Impossibility of the Benchmark Allocations

In both benchmark cases, the optimal contract consists of a certain amount of initial investment I, a finite liquidity provision policy ϑ^*, and a flat and incentive-compatible repayment D. However, there are problems with this allocation in a non-exclusive environment, as stated by the following proposition:

Proposition 2.4. *In equilbria satisfying some refinement criteria as specified in the appendix*[4]*, the exclusive competition allocation with*

$$I = \frac{A}{F(\vartheta^*)(\vartheta^* - \rho_0)},$$

with $\bar{\vartheta} = \vartheta^*$ *and with* $p_H D(\vartheta) = \rho_0$ *for all* ϑ*, cannot be sustained.*

[4]This refinement amounts to a form of trembling-hand perfection over the contract acceptance decision by the firm.

The formal proof is in Appendix A. The intuition of the result is as follows. In the exclusive competition allocation, the firm obtains a positive share of the revenue. An entrant can provide the firm with additional liquidity (i.e., on on top of $\overline{\vartheta}$) in exchange for the firm's share of the revenue.[5] Upon accepting this additional liquidity, the firm can now also continue if $\overline{\vartheta} < \vartheta \leq \overline{\vartheta} + \Delta\vartheta$, by raising $\overline{\vartheta}$ from the active lenders and up to $\Delta\vartheta$ from the entrant. The entrant finds it profitable to make this offer up to an amount $\Delta\vartheta$ as long as the expected revenue from receiving the firm's share (given that the firm uses the facility and provides low effort) exceeds the expected amount of liquidity that the lender needs to provide. That is, as long as

$$E\left((\vartheta - \overline{\vartheta})I \mid \overline{\vartheta} < \vartheta < \overline{\vartheta} + \Delta\vartheta\right) < p_L \frac{B}{\Delta p}.$$

Notice that the deviation considered above is possible irrespective of the size of the investment. This implies that the deviation cannot be prevented by covenants on or monitoring of the investment size. In the model without liquidity shocks, some effects of non-exclusivity can be mitigated by appropriate covenants Attar, Casamatta, Chassagnon, and Décamps (2010). However, in the present economy, even if the investment size could be imposed ex ante, the result of Proposition 2.4 would still hold.

Proposition 2.4 shows that with non-exclusive competition, the possibility that the firm has the ability to understate its liquidity needs at $t = 1$ turns out to have an important implication. Under exclusive competition, truth telling was trivially satisfied because the firm had no incentive to underreport its type and no possibility to overreport it. Under the deviation in Proposition 2.4, the deviating lender can offer additional liquidity support by free riding on the support by the active lenders. The firm can still access the original liquidity facilities by *underreporting* its type to the active lenders in the states of the world where this is needed. In these states of the world the deviating lender reduces the total surplus from the project, but he externalizes the losses to the other lenders.

2.5.2 Equilibrium

This section deals with establishing the existence of an equilibrium in the non-exclusive setting, and deriving its characteristics. By an equilibrium in this free-entry model we mean any set of traded contracts under which there is always an inactive lender, and no inactive lender has an incentive to offer another contract. In order to analyze the equilibrium, we introduce two further assumptions:

Assumption 1: Observability of the project size. In a model without a liquidity shock, the size of the project reveals ex post to the lenders whether or not the firm has contracted with other lenders. We assume that after $t = 0$, the total size of the project can be observed and used as a contracting variable. As is documented in Attar, Casamatta, Chassagnon, and Décamps (2010), observability of the project size can mitigate some of the inefficiencies caused by non-exclusivity.

[5] Legal restrictions can exist on the possibility of a manager to sell his stake in the firm. Yet, even if the manager cannot sell his stake, he can dilute it by issuing more debt.

In their paper, the lender has only limited means to punish the firm in case the firm has attracted additional funds from other lenders. The presence of the liquidity shock gives lenders an additional disciplining device as lenders are now able to cut the firm's liquidity supply in case the firm has attracted too much initial funding.

Note, however, that ϑ and the various L_i still cannot be used in contracts. The idea is that the project size I corresponds to the actual size of the firm's assets and is thus easily verifiable. As a balance sheet item, it is easy to verify in court. However, it is possible for the firm to hide additional cash flows, for example by obtaining supplies in kind, by using off-shore transactions, or by using derivative structures.

In order to deal with this contractual contingency on I, a slight abuse of notation needs to be introduced. The contracts now specify, besides the same J_i, the functions $L_i(\cdot,\cdot)$ and $D_i(\cdot,\cdot)$ that depend on both the message sent by the firm and the total investment level. On the aggregate level, we denote by $D(\vartheta, I)$ the lowest level of debt that guarantees a liquidity injection of ϑI, and by $\overline{\vartheta}(I)$ the maximum liquidity shock that can be withdrawn, *for a given level* of investment I. We still indicate the equilibrium quantities with $D(\vartheta)$ and $\overline{\vartheta}$.

Assumption 2: Strategic default. In order to build the equilibrium it is important to specify whether or not the firm can strategically default, i.e. assume a greater total debt load $D(\cdot)$ than the total potential revenue of the project RI. We assume that this kind of strategic default is possible. If the total debt load can exceed the revenue from the project, the question is how this total revenue is divided among the creditors in case the firm is bankrupt. We assume the pro rata distribution of the debt claims among the various creditors, weighted by the size of the claim. No creditor can impose seniority.

We denote by $\widehat{D}_i(\cdot)$ the amount of repayment lender i gets in equilibrium as a function of the realized liquidity shock, and we have

$$\widehat{D}_i(\vartheta) = \begin{cases} D_i(m_i(\vartheta)) & \text{when} \quad D(\vartheta) \leq RI \\ \frac{D_i(m_i(\vartheta))}{D(\vartheta)} RI & \text{when} \quad D(\vartheta) > RI. \end{cases}$$

Given that strategic default is possible, the firm is free to choose as high a debt level as possible. This means that, in any continuation equilibrium, the only constraint on the liquidity shock that can be financed is the total possible supply of liquidity. That is, we have $\overline{\vartheta} = \max \left\{ \frac{\sum_{i \in S} L_i}{I} \right\}$.

The assumption of pro rata sharing in case of bankruptcy implies a possibility for dilution that exacerbates the result of Proposition 2.4. Independently of the repayment, an entrant can always dilute active lenders, making a finite liquidity provision impossible. We have the following proposition:

Proposition 2.5. *Under Assumption 2, any allocation with a finite $\overline{\vartheta}$ cannot be sustained in equilibrium.*

The proof of the proposition is given in Appendix A. The intuition is as follows. For any finite level of liquidity, there is a corresponding total debt load. A deviating lender can enter the market and offer some more liquidity in exchange for a repayment that only has to be paid if the firm uses the additional facility, as in the proof of Proposition 2.4. The difference here is that, no matter what the current total debt level $D(\overline{\vartheta})$, the entrant can set a repayment that is high enough to dilute the existing creditors. Under the assumption that every creditor is treated equally, the active lenders cannot prevent this deviation.

Because of Proposition 2.5, the liquidity supply to the firm must be unlimited in any equilibrium. This means, however, that if an entrant increases the project size without offering additional liquidity insurance, the firm could still get full insurance. However, by Assumption 1, incumbent lenders could prevent such behaviour from entrants by conditioning the contract on the observed investment size.

Proposition 2.6. *Assume $\vartheta^* - p_L R \geq \frac{\Delta p}{p_L}$. If one lender, say lender 0, offers the contract $(J_0, L_0(\cdot,\cdot), D_0(\cdot,\cdot))$ with*

$$J_0 = \frac{A}{1 + E(\vartheta) - \rho_0},$$

$$L_0(m_0, I) = \begin{cases} m_0 I & \text{if } \; I = A + J_0 \\ 0 & \text{if } \; I \neq A + J_0, \end{cases}$$

and

$$D_0(m_0, I) = \begin{cases} (R - \frac{B}{\Delta p})I & \text{if } \; I = A + J_0 \\ RI & \text{if } \; I \neq A + J_0, \end{cases}$$

then there is no profitable deviation for the entrants.

The proof is in Appendix A. The result can be understood by first of all noting that the liquidity shock gives the lender an additional way in which it is able to punish the firm for seeking additional up-front investment. The threat by the lender of cutting off the liquidity supply makes it possible for the firm to commit to not seeking up-front investment from any of the other lenders. Similarly, because of the already unlimited liquidity supply, the firm has no reason to seek additional support from outsiders. This contract sees the lender break even, so there is no way for an entrant to undercut lender 0's offer.

We can now establish the condition under which there are gains from trade with an infinite liquidity provision. This will make it possible to characterize the existence of an equilibrium. We have the following.

Proposition 2.7. *Assume $\vartheta^* - p_L R \geq \frac{\Delta p}{p_L}$ and*

$$\frac{(\rho_1 - \rho_0)\, A}{(1 + \mathbf{E}\vartheta) - \rho_0} \geq U_A; \tag{2.12}$$

then, there exists an equilibrium allocation with trade and infinite liquidity provision.

The formal proof is given in Appendix A. The intuition behind this result is as follows: because of Proposition 2.5, we are focusing on equilibria with an unlimited liquidity supply. The inequality (2.12) provides the condition for such an equilibrium to be feasible: if it is satisfied, lenders can provide unlimited liquidity insurance and still break even, all the while providing a utility of at least U_A to the firm. This means that there can be an allocation that is individually rational for both the firm and the lender(s). The contract mentioned in Proposition 2.6 then constitutes a strategy for one lender to finance the firm and provide unlimited liquidity support.

We are now ready to synthesize all of the above results into one main proposition, that states the existence of an equilibrium, and the uniqueness of the equilibrium allocation:

Proposition 2.8. *Assume the conditions for Proposition 2.7 are met; then, there is an equilibrium under non-exclusive competition and any equilibrium is characterized by the aggregate allocation* $(I, \overline{\vartheta}, D)$ *with*

$$I = \frac{A}{1 + \mathbf{E}\vartheta - \rho_0},$$
$$\overline{\vartheta} = \infty,$$

and

$$D = \left(R - \frac{B}{\Delta p}\right) I.$$

Note that lenders break even in this allocation, and, given that liquidity support must be unlimited, the firm's utility is maximized. This proposition follows intuitively from the results above. As was already established in Proposition 2.7, there is an equilibrium with an infinite liquidity provision. Because of Proposition 2.5, any equilibrium must be characterized by an infinite liquidity provision. Furthermore, because of the possibility of entrants undercutting any allocation that leaves a profit to the lenders, the firm's utility must be maximized, subject to the familiar incentive compatibility and break-even constraints, but now with the added constraint that liquidity provision must be unlimited.

When comparing this result to the benchmark of exclusive competition, it can first be noted that $1 + \mathbf{E}\vartheta = \lim_{\vartheta \to \infty} c(\vartheta) > c(\vartheta^*) = \vartheta^*$, so that $1 + \mathbf{E}\vartheta - \rho_0 > \vartheta^* - \rho_0 > F(\vartheta^*)(\vartheta^* - \rho_0)$, which gives

$$I = \frac{A}{1 + \mathbf{E}\vartheta - \rho_0} < \frac{A}{F(\vartheta^*)(\vartheta^* - \rho_0)} = I^C.$$

As the total cost of financing per unit of investment size is higher under full insurance than it is under optimal insurance, lenders will be able to finance only a smaller investment. By the very definition of the competitive allocation, the surplus then generated is the maximal, given the investor's break-even constraint and the firm's incentive constraint. This means that the surplus generated in the equilibrium under non-exclusive competition, satisfying those same constraints, generates a lower surplus. Indeed the surplus generated under non-exclusive competition equals

$$(\rho_1 - \vartheta^*) I = \frac{\rho_1 - \vartheta^*}{1 + \mathbf{E}\vartheta - \rho_0} A,$$

and the surplus generated under exclusive competition equals

$$F(\vartheta^*)(\rho_1 - \vartheta^*) I^C = \frac{\rho_1 - \vartheta^*}{\vartheta^* - \rho_0} A,$$

which, as $1 + \mathbf{E}\vartheta > \vartheta^*$, is greater than the non-exclusive surplus.

Under monopoly, the total investment size was

$$\frac{U^A}{F(\vartheta^*)\,(\rho_1 - \rho_0)} =: I^M,$$

giving an expected total surplus under monopoly of

$$F(\vartheta^*)\,(\rho_1 - \vartheta^*)\,I^M = \frac{\rho_1 - \vartheta^*}{\rho_1 - \rho_0} U^A.$$

Now, if Assumption (2.12) holds,

$$I = \frac{A}{1 + \mathbf{E}\vartheta - \rho_0} > \frac{U^A}{(\rho_1 - \rho_0)}.$$

So the total surplus generated under non-exclusive competition, $(\rho_1 - \vartheta^*)\,I$, still exceeds the total surplus under monopoly. A comparison between I and I^M is not straightforward to make, but we can conclude that, in case a full-insurance equilibrium under non-exclusive competition exists, the total surplus generated will be between the monopoly surplus and the exclusive competition surplus.

2.6 Conclusion

In this paper we analyze a model of non-exclusive competition among lenders that provide investment and liquidity support to a firm. The paper highlights a new and interesting channel through which non-exclusivity can help to explain an increase in firm liquidity holdings. Indeed, the increasing multiplicity of bank-firm relationships seems to have been a forerunner of the build-up in cash holdings. It is important to notice that our paper takes a very generic view of liquidity: our aggregate variable representing liquidity can represent the up-front provision of liquid securities or an insurance in the form of credit lines. Clearly the former translates directly in higher cash holdings, while the latter does not necessarily affect them.

Nonetheless, our interpretation of the equilibrium with unlimited liquidity support is that it will, at least partly, be implemented through higher cash holdings. This is particularly true if lenders cannot commit to providing credit lines. Concerns about the difference between cash holdings and credit lines (as in Acharya, Davydenko, and Strebulaev, 2012) point to a second channel through which non-exclusivity leads to higher cash holdings. As a general result, the free-riding problem under non-exclusivity lowers the NPV per unit of investment, as the optimal liquidity policy is not implementable. This reduces the scope for trading between lenders and firms, forcing more firms to self-insure against liquidity shocks. This leads to a shift from credit lines towards cash holdings, increasing the liquidity that firms need to hold.

Chapter 3

Competition, Common Agency, and the Need for Financial Intermediation

The bank is something more than men, I tell you.

John Steinbeck, *The Grapes of Wrath*

Banks, and other financial intermediaries, play an important role in bringing capital from investors to entrepreneurs and households in any advanced economy around the world. The reason why banks are needed as a middle-man has been debated by economists. This paper aims to provide a novel rationale for the existence of bank-like financial intermediaries: banks arise as an institution to coordinate competition between investors.

In addressing the raison d'être of financial intermediation, three arguments stand out in the literature. Diamond and Dybvig (1983) document how it is optimal for investors to pool the choice of the maturity structure of their investments if they face uncertainty regarding their liquidity needs. Diamond (1984) argues how if several investors invest in the same company, every one of them will have an incentive to shirk away from monitoring, making it optimal for investors to delegate monitoring to a single institution. Finally, Diamond and Rajan (2001) argue that the possibility of bank runs disciplines banks into not extracting rents when rolling over an investment, thereby improving efficiency ex ante.

In this paper, intermediaries are necessary as an institution to coordinate the flow of capital from investors competing in a *non-exclusive* and *uncoordinated* fashion to borrowers. Competition is non-exclusive in the sense that investors can neither observe nor control the contracts that borrowers have with other investors. Competition is uncoordinated in the sense that if several investors finance the same borrower together, they cannot coordinate on the contracts that each one of them trades with that borrower.

This paper features a simplified Holmström and Tirole (1998) type model of investment: borrowers have access to a project that requires an initial fixed investment, pays off at a later date and is subject to a liquidity shock at an intermediate date. Under optimal contracting, investors provide money to the borrower for the initial investment and set a maximum level up to which the liquidity shock can be financed at the intermediate date. In exchange for their investment, they ask for a repayment from the project's proceeds at the final date.

This paper finds an explanation for the presence of intermediaries by describing an economy that cannot reach an equilibrium if investors and borrowers can only contract directly with one another. This is illustrated by first restricting investors in this economy to only deal with borrowers directly, and vice versa. Under this restriction, because of the non-exclusive and uncoordinated nature of the contracting, a two-sided free-riding problem arises. On the one hand, a similar problem as in Boxtel, Castiglionesi, and Feriozzi (2013) is present: for any finite credit line that incumbent investors provide, outside investors can free-ride upon this provision and offer an additional "emergency" liquidity provision. They can do so by diluting the incumbent investors in the states of the world where this additional liquidity provision is used. This problem ultimately makes it impossible to limit the intermediate date liquidity provision.

On the other hand, if liquidity support is unlimited, various investors have to provide liquidity together, as the shock might exceed the limited endowment of each individual investor. The potential expected repayment that the borrower can offer remains limited. This means each investor has an incentive to change the pricing of his liquidity provision in such a way that the other investors are responsible for providing liquidity, thus investing and not insuring, free-riding on the insurance others are providing.

These two free-riding problems make that there is no equilibrium under the restriction that banks and borrowers can only contract directly with one another. In order to circumvent this double free-riding problem, an institution can coordinate investment by collecting money from investors and using it to invest in and provide liquidity insurance to borrowers, thus acting as an intermediary. This paper will discuss how such an institution can either be implemented by a social planner or arise endogenously trough contracting between investors, effectively making one or more investors intermediaries. This intermediary can use the pooled resources of several investors to fully insure the firm's liquidity shock, and is thus not susceptible to free-riding by investors increasing the liquidity supply.

One investor becomes the intermediary for others and thus the sole entity directly trading with a company. This dynamic is related to a common observation in the economic history literature: in the latter half of the nineteenth century, especially in Germany[1], and, to some extent, in the United States, financing by large (universal) banks has played a pivotal role in the development of areas like mining, railroads, utilities, and heavy industries. This paper models production technologies that require some initial investment and face potentially high, yet initially uncertain, costs or reinvestment needs at a later date. Thus, I aim to capture precisely the salient features of technologies that played a role in these areas: after investment, development costs are still uncertain and in case of a technical failure or unexpected cost overrun, it is next to impossible to scale down operations. In this paper, it is indeed assumed that if a borrower cannot face its intermediate date liquidity needs,

[1] In the literature review I will give more references to papers discussing the role of bank financing in industrial development

it needs to abandon the project, yielding a liquidation value of zero.

For already established banking markets, this paper could shed light on consolidation and concentration in the banking sector. Recent merger waves in the 1990s in the US (Calomiris, 1999; Calomiris and Karceski, 2000) and in the 1980s and 1990s in Europe (Karceski, Ongena, and Smith, 2005; Boot, 1999) have been studied, mostly from an efficiency perspective. This paper could shed light on how increased international and domestic competition, new technologies, and deregulation might have precipitated these concentration movements. If it became easier for banks to compete over financing the same firms, then by the mechanism described in this paper, banks would have needed to pool their investment together. This pooling would allow them to circumvent the non-exclusivity and common agency problems arising from many smaller banks competing with one another.

Similarly, a banking sector with multiple banks competing should develop "intermediaries for intermediaries". Thus, the model in this paper can be applied to understand the origin of investment banking, and, more recently, the rise of syndication and originate-to-distribute models of banking. In most syndicated deals, a group of banks appoint one bank as the *lead arranger*, who is also in charge of credit lines to the firm receiving the loan. This lead bank thus becomes an incidental intermediary for the others in a dynamic not dissimilar to the one described in this paper. In the originate-to-distribute model, one bank originates loans for particular borrowers, and re-sells them to other investors, including other banks.

From a theoretical viewpoint, the two problems leading to market failure under direct contracting between borrowers and investors fall into two distinct, but related categories: on the one hand, there is an *exclusivity* problem and on the other hand, there's a *common agency* problem. The exclusivity problem stems from the inability of the borrower to commit to only contract with a given number of investors without seeking additional trade with other investors. The common agency problem stems from the fact that multiple investors are needed to finance one borrower and there exist contractual externalities between different investors financing the same borrower.

The exclusivity problem is of the nature of those documented in Bizer and DeMarzo (1992) and Arnott and Stiglitz (1991): in a large class of allocations, there are potential gains from trade between the agent (the borrower) and a possible entrant, in which the entrant imposes a negative externality on the incumbents. Specifically following Boxtel, Castiglionesi, and Feriozzi (2013), in this paper there is no possibility to limit the liquidity supply to the borrower, as potential entrants can always extend the liquidity supply on top of what is already offered.

The need to provide the borrower with unlimited liquidity leads to the fact that in order to finance the borrower, multiple investors are necessary. This situation, with one agent and multiple principals is referred to as *common agency.* As has been shown both with moral hazard (Bernheim and Whinston, 1986) and with adverse selection (Martimort and Stole, 2002), the contractual externalities in these common agency situations are such that the resulting contracts differ substantially from those in classical single principal-single agent situations. In this paper, the common agency problem is similar to the one described in Castiglionesi and Wagner (2013): although all investors have an interest in seeing the borrower continue, each one also has an interest in the other ones providing the funds for the borrower to do so, so that each one will try to adjust its contract with the borrower in such a way that others provide more liquidity.

The rest of the paper is set up as follows: section 3.1 reviews the related literature. Section 3.2 presents the model and its basic assumptions. The main intuitions of the model are treated in an example in section 3.3. Section 3.4 then proves the no-trade results for a more general model. Section 3.5 describes how intermediaries can restore trade. Section 3.6 discusses the robustness of the model's most important implications to different modeling assumptions. Section 3.7 concludes and presents some ideas for further research. Formal proofs of the main results can be found in Appendix C.

3.1 Literature

This paper models financial intermediaries as a means to overcome two types of externalities between investors. To make a very blunt classification, these sorts of externalities fall within the type commonly studied in the *common agency* literature on the one hand, and the *non-exclusivity* literature on the other. Even though these two literatures are very similar and closely related at a theoretical level, the common agency literature tends to deal with situations in which an agent finds himself contracting with multiple principals, whereas the non-exclusivity literature deals with situations in which the agent *could* contract with several principals.

In the classic literature on competition with asymmetric information it is often, implicitly or explicitly, assumed that agents only deal with only one of many principals. The competitive mechanism of principals undercutting each others' offers would then lead to a solution that is optimal for the agent, only constrained by the relevant moral hazard or adverse selection problems. However, as Pauly (1974) notes, the possibility of agents privately contracting with several counterparties at the same time leads to equilibria that are inefficient also with respect to the constrained optima, as those constrained optima can leave room for a private trade between the agent and non-incumbent principals. These trades would then impose externalities on incumbent lenders.

In economies with moral hazard, potential gains from trade between the agent and outside principals are possible, as the outside principals can externalize the incentive effect of their contracts. In insurance economies (Arnott and Stiglitz, 1991; Bisin and Guaitoli, 2004; Attar and Chassagnon, 2009), trade can be restored by *latent* contracts. (Bizer and DeMarzo, 1992) model a consumption economy in which a borrower can sequentially approach multiple banks. Under the constrained optimal contract, the borrower has an incentive to approach other banks.

A different problem stemming from non-exclusivity in economies with moral hazard is documented in the papers by Parlour and Rajan (2001) and Attar, Casamatta, Chassagnon, and Décamps (2010): if an agent can choose contracts from multiple principals, there is a problem with the classic Bertrand (1883) workings of competition. If one principal offers a contract that leaves him some rent, other principals do not want to undercut, as the agent might be induced to accept both the first contract and the undercutting contract. This gives equilibria in which one principal can extract up to monopoly rents, as no other principal can undercut him if he does.

Another strand of the non-exclusivity literature focuses on adverse selection. Attar, Mariotti, and Salanié (2011) and Attar, Mariotti, and Salanié (2014) study how markets with adverse selection work in the presence of private contracts between agents and multiple principals. This adverse selection can also start playing a role at an intermediate date, after contracts have been agreed upon: Boxtel, Castiglionesi, and Feriozzi (2013) document how firms receiving liquidity support from different investors want to contract additional lines of liquidity support, thus giving that any limited level of liquidity provision leads to potential deviations from outside investors. This mechanism also plays an important role in this paper.

Very closely related, and on a technical level very similar, is the literature on common agency: situations in which an agent contracts bilaterally with multiple principals at the same time and the principals impose indirect externalities on each other through the incentive effects of their contracts. Bernheim and Whinston (1986) model an economy with moral hazard. As providing incentives through a contract is costly, each principal will want to other principals to provide the incentives, leading to a free-riding problem. Martimort and Stole (2002) and Peters (2001) look at common agency situations with adverse selection and find that the revelation principle found in single-principal mechanism design problems does not hold in multi-principal settings. A very practical common agency problem close to the one in this paper is discussed in Castiglionesi and Wagner (2013).

This paper aims to address the formation of large bank-like intermediaries as the most common instrument in an economy to get funds from investors to firms. The economic history literature has noted some times and places where banks started playing an exceptionally large role. Especially in Germany in the late 19th and early and 20th centuries, large universal banks dominated the financing of German firms. The special relation between German banks and industrial firms has been noted by contemporaries (Jeidels, 1905; Riesser, 1910). In a seminal analysis, Gerschenkron et al. (1962) states that

> ...the German banks, and along with them the Austrian and Italian banks, established the closest possible relations with industrial enterprises. A German bank, as the saying went, accompanied an industrial enterprise from the cradle to the grave, from establishment to liquidation throughout all the vicissitudes of its existence.[2]

He argues that it was the presence of large universal banks, or *Großbanken* that allowed the German economy to mobilize enough capital for the second industrial revolution and to "catch up" with the more industrialized economy of the United Kingdom.

This "Gerschenkron hypothesis", that universal banks with a great deal of corporate control were needed especially to mobilize capital in order to rapidly indus-

[2] This echoes an earlier quote by Jeidels (1905) of the same gist: "The banks attend an industrial undertaking from its birth to its death, from promotion to liquidation, they stand by its side whilst it passes through the financial processes of economic life, whether usual or unusual, helping it and at the same time profiting from it."

trialize the relatively backward German economy[3] has afterwards been extensively discussed, refined, and reinterpreted. Da Rin (1996) reinterprets the important role of universal banks in Germany as being focused on *information*, in monitoring, control and coordination of investment. Da Rin and Hellmann (2002) provide a formal theoretical model of the big-push dynamics that underlie the Gerschenkron hypothesis. Guinnane (2002) analyzes the development and role of all banks in Germany, not just the Großbanken, in the 19th and early 20th centuries from the point of view of *delegated monitoring* (cf. Diamond, 1984) and argues that German banks were particularly apt at performing this task.

Numerous studies have compared the German experience to the experience in other countries. During the American "Gilded Age" a number of large financiers, the best known of which is J.P. Morgan, have played a pivotal role in financing American industrialization. The financiers of the house of Morgan were often active on the boards of directors of the firms they financed. DeLong (1991) finds that firms with J.P. Morgan representatives on their boards were 20% more valuable than those without J.P. Morgan men on their boards. Ramirez (1995) finds that this difference is most likely attributable to liquidity issues. Having close ties to a bank makes it easier for firms to raise funds in times of high liquidity needs. An interesting case study is presented in Chandler (1954), focusing on the patterns of railroad financing in the US: even though railroads were often equity financed, firms relied on intermediaries to raise equity. Often larger equity-financed railroads ran into liquidity problems, as one would expect from the analysis in this paper.

For all the similarities between American investment banking and German universal banking, differences are noted by Calomiris (1993, 1995). He argues that even before the Clayton and Glass-Steagal acts of 1914 and 1933 respectively, the regulatory branching and activity restrictions on banks were such that the American system performed significantly worse in financing industrial development.

The other comparison that is often made is between the United States and Germany on the one hand, where bank financing played a relatively important role and Great Britain on the other, where, according to Gerschenkron et al. (1962), banks were "obsessive about liquidity and only lent on a short term, hands-off basis." (cf. Guinnane, 2002). Davis (1963) hypothesizes that industrial development in the U.K. in the late nineteenth century started lagging behind that in the U.S. and Germany, because the U.K. lacked the kind of large financial institutions that the U.S. and Germany had. This difference in growth has been documented extensively in Lewis (1978) and attributed by some scholars (such as DeLong, 1991) to the different financial systems present in the different countries.

A few papers compare financing by universal banks to other forms of financing *within* an economy. The aforementioned papers by DeLong (1991) and Ramirez (1995) do so for the United States around the turn of the twentieth century. Hoshi, Kashyap, and Scharfstein (1991) compare firms in post-WWII Japan that have close ties to so-called *keiretsus*, large financial conglomerates, to firms that lack these close ties. They find that the former group of firms has a smaller sensitivity to liquidity shocks, indicating that they are less liquidity constrained. Becht and Ramírez

[3]and that the even more backward Russian economy needed the even more draconian force of centralized state planning.

(2003) perform a similar exercise, looking at bank affiliation in Germany and find that mining and steel companies with ties to universal banks were significantly less liquidity constrained than non-affiliated firms.

As to the question why banks are needed in a financial system, two seminal arguments stand out in the theoretical literature. Diamond and Dybvig (1983) state that investors need banks in order to pool the uncertainty they have about future liquidity needs. Pooling investments can lead to a better mix of liquid short-term and profitable long-term investments than each investor individually optimizing his investment over different horizons. Diamond (1984) provides another explanation: investors need to perform the costly task of *monitoring* in order to mitigate information asymmetries with borrowers. However, if other investors monitor already, each investor is better off not expending the monitoring effort, free-riding on the monitoring provided by others. Ultimately this means that everyone is better off if one investor gets assigned to be a *delegated monitor,* intermediating for the other investors.

This paper is closer to Diamond (1984) in the sense that it models the necessity of banks from features of the borrowers' investment technology. However, this paper is substantially different from Diamond (1984) on a rather fundamental level. It is the inability of both agents and principals to commit to contracts that causes inefficiencies in this model. The technologies in this paper are exogenous, so a moral hazard problem as in Diamond (1984) does not exist. Investors have no special ability to overcome information asymmetries that exist between investors and borrowers, be it through some costly monitoring expenditure or through learning about the borrowers' technology. This also makes a difference in terms of interpretation: in Diamond (1984), banks perform the economically productive task of monitoring, and centralizing this task to one party gives economies of scale. In this paper, however, intermediaries perform no productive task and are merely needed to overcome contractual externalities between investors.

This special role assigned to banks is also present in Holmstrom and Tirole (1997), where banks distinguish themselves from other investors by their ability to reduce moral hazard through monitoring. A similar ability often attributed to banks is the ability to gain information over a relationship, thereby reducing adverse selection. Rajan (1992) weighs off the costs and benefits of this informational advantage: on the one hand, banks can use their informational advantage to extract rents, but on the other it does allow for more efficient continuation and liquidation decisions. Diamond and Rajan (2001) analyze how deposit-taking institutions, because of the possibility of runs, have the right balance sheet structure to mitigate potential rent-seeking.

Another paper close to this one is the paper by Dewatripont and Maskin (1995), that studies the trade-off between centralized bank financing and decentralized market financing in a model with refinancing: banks might be inclined to refinance too often and decentralized financiers do not refinance often enough. This, however, takes place in a world where financiers cannot commit ex ante to refinancing or not, and where refinancing costs are public information. Furthermore, the causal interpretation is almost reversed. In Dewatripont and Maskin (1995), overinsurance occurs because of bank financing, whereas in this paper, overinsurance occurs because of other reasons and banks are the only institution able to provide this excessive insurance.

3.2 Model

This section introduces the full model. In order to prove the necessity of intermediaries in this economy, I first restrict the agents, so that investors can only contract directly with a borrower and vice versa. Under this restriction, no trade can take place. Then I go on to show that with intermediaries, trade can take place. I do this first for a simple example (section 3.3) to develop the intuition and then for the more general model (section 3.4).

There are three dates, $t = 0, 1, 2$, and a single good, called money. All agents are risk neutral and do not discount future cash flows. There is a large sector of investors financing a single borrower. There is a single *borrower* and M *investors*, with $M > 1$. The investors are indexed by $i = 1, 2, \ldots, M$ and each investor has the same endowment W. The assumption that there are multiple investors, but only one borrower, guarantees that if contracting is exclusive, investors make zero profits and maximize the borrower's surplus.

Borrower and Projects The borrower has access to a project that requires an input of money at two different dates: at $t = 0$ an initial investment is needed and at $t = 1$ there is a *liquidity shock*. To keep the model simple, I assume that the initial investment is fixed and requires $I = 1$ units of money. Furthermore, the borrower cannot divert money or spend it on anything other than his project, so that any money the borrower raises at $t = 0$ in excess of I, does not change his utility.

The liquidity shock is an exogenous cost that realizes at $t = 1$ and that needs to be financed in order for the borrower to be able to continue the project. This shock can be thought of as a repair cost after a technical failure or an investment that is needed and of which the costs weren't certain at the inception of the project. If a firm cannot pay this liquidity shock, it cannot continue the project, i.e. there is no possibility of scaling down if only part of the necessary funds can be raised. This assumption is not unrealistic in heavy industries, where a technical failure of one reactor, smelter, or blast furnace can shut down an entire production process. Similarly, one could think of a mining project or a railroad, where natural obstacles have to be cleared in order for the project to operate.

This liquidity shock is a random variable ϑ, drawn from a distribution $F(\cdot)$ with density $f(\cdot)$, and with compact support Θ. If the borrower can raise ϑ, she will do so in order to finance the project and if he can't, the project has to be abandoned and will yield zero. Following Holmström and Tirole (1998), I assume that there are no investment or consumption opportunities at $t = 1$ and that the borrower cannot divert cash, so any money withdrawn in excess of ϑ is wasted. The only possible thing the borrower can do with excess liquidity is "burn" it, so that it only matters for the borrower whether the total amount of cash he can raise is larger than ϑ or not. I make this assumption both in the exclusive benchmark case and in the non-exclusive case that is the topic of study in this paper. What happens when one relaxes this assumption, is studied in the exclusive context in Tirole (2006, Chapter 5.3). However, to make the analysis clearer and to isolate the effects of non-exclusivity and common agency in this context, I keep the simplified assumption.

If the project is continued, it will yield a transferable return R and a private benefit $B > 0$. Both are fixed, given continuation. One could think of $R + B$ being the borrower's total income, of which only R is pledgable, or, alternatively, of R as the pecuniary return on the project and B as a private, non-pecuniary utility that the borrower enjoys from operating the project, such as perk consumption, status or experience. If the project is continued, the return and private benefit are not dependent upon any action by the borrower, so that *given continuation*, there is no moral hazard. The borrower has limited liability and B cannot be appropriated by outsiders, so at $t = 1$, the firm would always rather continue than liquidate, even if continuation entails bankruptcy. It is important to note that B does not play the same role as the private benefit in the Holmström and Tirole (1998) framework: given continuation, the return and private benefit are fixed and not dependent upon an effort choice by the borrower.

On the distribution of ϑ, I assume that for every $\tilde{\vartheta} < \sup \Theta$, one has

$$\mathbf{P}\left(\tilde{\vartheta} < \vartheta < \tilde{\vartheta} + \min(R, W)\right) > 0,$$

which means that for any level of the liquidity shock inside the range of possible values, there are other values of the liquidity shock not "too" far away, which can occur. For continuous distributions it is sufficient to assume that $F(\cdot)$ is increasing, i.e. that there are no "holes" in the support of ϑ. I make the more general, and more technical, assumption to be able to deal with discretely distributed shocks, which I need in simplified examples. In popular terms, the general assumption allows for holes in the support of ϑ, as long as they are not too big.

Contracts At $t = 0$, each investor i offers a contract C_i to the borrower. After observing the contracts, the borrower chooses a subset $\mathcal{I} \subset \{1, 2, \ldots, M\}$ with whom she trades. Each contract C_i is a triple consisting of

- an up-front transfer J_i from the investor to the borrower,
- a maximum liquidity provision $\overline{L}_i$, at $t = 1$. At $t = 1$, the borrower can demand an amount of liquidity L_i up to $\overline{L}_i$. It is assumed that the borrower has no investment or consumption opportunities at $t = 1$ (as in Holmström and Tirole, 1998), so that the borrower only cares whether or not the aggregate amount of money it attracts is large enough to cover the liquidity shock, and
- a debt repayment $D_i(\cdot)$ that can be a function of the demanded liquidity L_i. In any subgame perfect equilibrium, this function needs to be non-decreasing, as otherwise the borrower would have an incentive to withdraw too much liquidity in some states of the world, wasting the unnecessary amount.

Default and Bankruptcy The borrower is protected by limited liability. If the borrower promises an aggregate repayment that exceeds R, a total of R has to be paid to the investors. Throughout this paper, it is assumed that the division of the proceeds among creditors is *pro rata,* i.e. every investor obtains a share of the proceeds weighed by the size of his own claim on the borrower's $t = 2$ income. This pro rata division makes it possible for investors to dilute one another and thus to free-ride on one another's liquidity provisions. However, this "dilution" could also come from moral hazard (cf. Boxtel, Castiglionesi, and Feriozzi, 2013) or from a speculative tranche of the borrower's revenue. This, however, is not the main concern here.

Notation In order to discuss equilibrium allocations, it is useful to develop some notation for aggregate quantities. Denote by $\overline{\vartheta}$ the maximum amount of liquidity the borrower can obtain. For every $\vartheta \in [0, \overline{\vartheta}]$, one can now define $D(\vartheta)$, the total debt repayment the borrower has to take on in order to finance a shock ϑ.

Parameter Assumptions The number of borrowers, M, is large in the sense that

$$M > \frac{\sup \Theta}{R} \tag{3.1}$$

This assumption guarantees that there are enough investors to ensure free-entry-like competition between investors. The liquidity shock possibly exceeds each individual investor's endowment:

$$\sup \Theta > W_i \text{ for all } i, \tag{3.2}$$

so that no single investor can offer full insurance to the borrower. Nonetheless, each investor's endowment is not too small, as it could provide financing and liquidity up to a level that optimizes the total surplus from the project:

$$W_i \geq 1 + \min \operatorname*{argmax}_{\vartheta \in \Theta} \left\{ F(\vartheta)(R + B) - \int_0^{\vartheta} \vartheta f(\vartheta) \mathrm{d}\vartheta \right\} \text{ for all } i. \tag{3.3}$$

These assumptions reflect that even though each investor's endowment is of a roughly similar size to the borrower's project, the potential costs that a project might run up are relatively large. Note that as a direct consequence of these two assumptions, $\sup \Theta > R$, so that under full insurance, the firm's repayment capacity is smaller than the maximum liquidity provision. This would imply that over at least some range of values for the liquidity shock, the price that the firm pays for a marginal unit of liquidity is smaller than one.

The ex post aggregate liquidity shock is bounded in such a way that the endowments of all investors together suffice to provide full insurance to the firm:

$$1 + \sup \Theta \leq \sum_{i=1}^{M} W_i \tag{3.4}$$

This assumption says two things: first, there is abundant cash in the system, so that market power in principle lies with the borrower. Second, aggregate risk is present, but limited, so that even ex post there is no cash shortage. Ex ante, even with full insurance, the project generates enough expected pledgable income to recoup expected costs.

$$R \geq 1 + \mathbf{E}\vartheta. \tag{3.5}$$

This means that there is no credit rationing, even when the project is fully insured.

Strategies and Equilibrium Concept The assumption is that investors play pure strategies. In the version of the model where investors are restricted to only directly trade with borrowers, contracts are offered simultaneously, after which all borrowers can accept any subset of the offered contracts. Later on in this paper, when I study intermediation, I will specify the timing of the contracting stage in case investors are allowed to contract among each other. The equilibrium concept for both versions of the model is subgame perfection.

3.3 Example

First, I offer a numerical example that conveys all the intuition of the general model.

Endowments and the Project There is a number $M \geq 2$ of investors that is large enough to satisfy all the assumptions of the general model. For each investor, $W_i = 5$. The borrower's project requires an initial investment of 1 unit and has a liquidity shock that is either *low*, with $\vartheta = \vartheta_L = 2$, or *high*, with $\vartheta = \vartheta_H = 6$. Both happen with equal probability. If the project is continued, the monetary return is $R = 5$ and the private benefit is $B = \frac{1}{2}$.

Assumptions Now, I can verify whether all assumptions from the main text are satisfied. First of all, two investors are already enough in the sense that condition (3.1) is satisfied. The maximum liquidity shock of 6 exceeds all investors' individual endowments, so that condition (3.2) is satisfied. I will check assumption (3.3) when I look at the first best case in the next subsection.

The maximum total cost of the project, $1 + \sup \Theta$, equals 7, so that the endowments of the investors, equaling $\sum_{i=1}^{M} W_i = 5M \geq 10$ are enough to finance the project, even under full insurance, so that condition 3.4 is satisfied. Under full insurance, the expected pledgable income from the project equals $R = 5$, which is equal to the expected total cost of $1 + \frac{1}{2}\vartheta_L + \frac{1}{2}\vartheta_H = 5$, so that assumption (3.5) is satisfied.

3.3.1 First Best and Exclusive Competition

In order to study efficiency, I focus on the total surplus that is generated by the project. The only variable that has an effect on total surplus, is the level up to which liquidity shocks are financed. In this particular case one can have either $\overline{\vartheta} = 2$, in which case the total surplus generated equals

$$\frac{1}{2}(R + B) - (1 + \frac{1}{2}\vartheta_L) = \frac{3}{4},$$

or $\overline{\vartheta} = 6$, in which case the total surplus equals

$$(R + B) - (1 + \frac{1}{2}\vartheta_L + \frac{1}{2}\vartheta_H) = \frac{1}{2},$$

so that it is optimal to have $\overline{\vartheta} = 2$. This means that also means that condition (3.3) is satisfied.

Now, I look at a "classic" exclusive competition setting, i.e. one where the different investors offer contracts of which the borrower can accept only one. For this, assume the following slight modification of the extensive form: after all investors have offered their contracts, the borrower only chooses one investor (instead of a subset of arbitrary size) with whom to do business. Under this arrangement a classic (Bertrand, 1883) competition result holds.

Proposition 3.1. *Under exclusive competition, liquidity is provided up to* $\overline{\vartheta} = 2$, *repayment equals* $D = 4$ *and the borrower receives the full surplus of* $\frac{3}{4}$.

The intuition is as follows: whenever an investor offers the borrower a contract under which the borrower's profits are not maximized, another investor can undercut and offer a more favourable contract. Thus, any equilibrium must have the borrower receiving the total surplus, and this surplus be maximized to $\frac{3}{4}$, with $\overline{\vartheta} = 2$. This will be implemented if two or more investors offer the contract $C^* = \left(J^*, \overline{L}^*, D^*\right)$, with $J^* = 1$, $\overline{L}^* = 2$ and $D^* = 4$, and the borrower takes one of these contracts.

3.3.2 Market Failure

As mentioned before, in case contracting is non-exclusive, two major problems arise that, together, make trade with bilateral contracts between investors and the borrower impossible. First of all, a problem similar to that in Boxtel, Castiglionesi, and Feriozzi (2013) arises: if the borrower can obtain a maximum liquidity provision from any number of incumbent investors that is below the largest possible liquidity shock, another investor can offer to extend the liquidity supply, free-riding on the incumbents' provision. The borrower will accept this extension as it will allow her to continue and at least obtain a private benefit if the high shock hits. Second of all, if multiple investors finance the borrower together, each of them has an incentive to deviate by changing his contract in such a way that the other investors provide more liquidity than he does. Each investor can do this by changing the marginal price of liquidity in such a way that the borrower would rather get his liquidity first from the other investors.

Extension of Liquidity Supply

Assume now the borrower is not restricted to deal with only one investor, but instead can choose to deal with any number of investors, without investors able to control the contracts the borrower trades with other investors. In case the aggregate liquidity supply is limited to $\overline{\vartheta} = 2$, entrant investors can come in and offer to extend the liquidity supply.

Assume, for example, the borrower trades the aforementioned competitive contract C^*, with a single investor i. Another investor j can offer to finance the difference between the low and the high liquidity shock in exchange for a sufficiently diluting share, but make sure that the borrower will not use this investor j's liquidity facility in case of the low liquidity shock. This can be implemented by a contract C_j with $\overline{L}_j = 4$, $D_j(0) = 0$, and with $D_j(L) > 16$ for any $L > 0$. [4] If the borrower accepts contract C_j on top of C^*, nothing changes if the low shock hits. However, the borrower now has the possibility to continue even in case the liquidity shock is high. She can do so by getting 2 from investor i and the promised 4 from investor j. In these states of the world, she can continue and at least have the private benefit. This means that, ex ante, the borrower's utility is enhanced by $p_H B = \frac{1}{4}$, meaning that the borrower will accept this contract offer.

For the deviating investor, the contract changes nothing if $\vartheta = \vartheta_L$, but if $\vartheta = \vartheta_H$, he will have to provide the promised 4 units of liquidity, but will obtain a diluting share which is the total revenue from the project diluted by his own claim in proportion to total claim, i.e.

$$\tilde{D} = \frac{D_j(4)}{D_i(2) + D_j(4)} R.$$

This is greater than 4 by construction, so that in the states of the world where ϑ is high, the deviating investor receives a positive profit. This would thus constitute a profitable deviation for investor j.

The deviation above is very similar to the type of deviation that is central in Boxtel, Castiglionesi, and Feriozzi (2013). The presence of the deviation described above entails that any allocation in which the liquidity provision is limited to ϑ_L cannot be sustained in equilibrium. This result is formalized in the following lemma, the proof of which is in the appendix.

Lemma 3.2. *No allocation with $\overline{\vartheta} = 2$ can be sustained in equilibrium.*

The Common Agency Problem

Lemma 3.2 gives that if competition is non-exclusive, liquidity insurance must be provided up to $\vartheta = 6$. In that case, since $W_i = 5$ for all investors i, the endowment of any single investor is not sufficient to provide liquidity in case of the high shock, meaning that the borrower must be financed by at least two investors at the same time. In this case a very particular problem arises: each of the investors will want to deviate by offering a contract that makes the other investors responsible for providing liquidity.

In order to illustrate this problem, I will show how the deviation works for a specific set of contracts. Note that, even though this is a specific example, this same intuition works for any general set of contracts. Suppose there are two active

[4] This deviation is rather extreme. Since I am dealing with a discrete distribution, the entrant has to finance the very large difference between ϑ_H and ϑ_L. Results still hold with smaller dilutions, if one assumes that instead, ϑ follows a continuous distributions. The discrete distribution in this example is chosen for expositional purposes.

investors, $i = \{1, 2\}$, each financing exactly half the project and offering identical contracts, i.e.

- $J_i = \frac{1}{2}$ for both i,
- $\overline{L}_i = 3$ for both i, and
- $D_i(L) = 2\frac{1}{2}$, for both i and for all L,

Then the borrower will get 3 units of money from each investor when the high liquidity shock hits, and chooses to obtain liquidity evenly from both investors in case of the low shock, i.e. $L_1^L = L_2^L = 1$.[5]

Note that both investors need to break even precisely, as they cannot demand more than the investor's pledgable income and they can only break even if they demand the total pledgable income in all states of the world. This allocation would thus give both investors a utility of precisely zero, would allow the borrower to continue in all states of the world and give the borrower a utility of $B = \frac{1}{2}$. Now investor 1 has an incentive to deviate by altering the contract to one with a slightly higher marginal price of liquidity. He can offer the contract $\tilde{C}_1 = \left(\tilde{J}_1, \tilde{L}_1, \tilde{D}_1(\cdot)\right)$ with

- an unchanged initial payment $\tilde{J}_1 = J_1 = \frac{1}{2}$
- an unchanged maximum liquidity provision $\tilde{L}_1 = \overline{L}_1 = 3$,
- a slight discount for the borrower if she chooses not to withdraw any liquidity: $\tilde{D}_1(0) = 2\frac{1}{2} - \varepsilon$ for some small enough $0 < \varepsilon < 1$, and with
- the "old" repayment in case of a higher withdrawal of liquidity: $\tilde{D}_1(L) = 2\frac{1}{2}$ for $L \leq 3$.

If the borrower accepts $\tilde{C}_1$ and C_2 and the high shock hits, nothing changes for the borrower, or any of the investors: the borrower will obtain 3 units of money from each of the investors, continue and then repay its full pledgable income, divided evenly between the two investors.

However, if the low shock hits, the borrower can get up to 3 units of money from investor 2, and would find it cheaper to do so, as this would cost only $5 - \varepsilon$ instead of 5. Ex ante, the borrower would thus be $\frac{1}{2}\varepsilon$ better off with these two contracts. For investor 1, the new contract means getting ε less repayment in the low state, but also needing to provide 1 unit less liquidity. Ex ante, this improves his utility by $\frac{1}{2}(1 - \varepsilon)$. Thus, this is a profitable deviation. Of course, investor 2 has the same possible deviation.

Each of the two investors has an incentive to deviate, so the arrangement above cannot be sustained in equilibrium. In any allocation with $\overline{\vartheta} = 6$, a similar deviation is possible for at least one of the investors, giving the following lemma.

[5] Of course the borrower is indifferent between getting his money from one and from the other investor, but getting money evenly would be the only subgame perfect option, as otherwise one of the investors would make a loss.

Lemma 3.3. *No allocation with $\overline{\vartheta} = 6$ can be sustained in equilibrium.*

Lemmas 3.2 and 3.3 together have the rather dramatic implication that if investors can only contract with the borrower directly and not among each other, and without some sort of institution to coordinate investment, no equilibrium with investment can be sustained, which I state here as a proposition.

Proposition 3.4. *No equilibrium exists in which investors trade only directly with borrowers and vice versa.*

3.3.3 Intermediaries

Having established that no equilibrium exists with investors only trading directly with the borrower, I now see whether there is one if I allow for the possibility of investors contracting with one another. Without changing the information structure, an equilibrium can be established in which one investor essentially becomes an intermediary.

Consider the following arrangement: investor 1 borrows 2 units of money from investor 2 at date 0. He can then offer the borrower the contract with

- $J_1 = 1$
- $L_1^L = 2$ and $L_1^H = 6$
- $D_1^L = D_1^H = 5$

Subsequently, at $t = 2$ investor 1 pays 2 units of money to investor 2, in either state of the world. Note that, even though there is aggregate risk in the model, the depositor's investment is completely risk-free.

These contracts constitute a rational set of strategies for investors: as in a classic competition set-up, the threat of free entry prevents investor 1 from making a profit, but also, by the threat of entrants expanding the liquidity supply (as in lemma 3.2), he is forced to, inefficiently, provide full liquidity insurance. He can finance this from the funds obtained from investor 2. As all investors make zero profits, investor 2 is indifferent between depositing or not. This entails that with intermediaries, trade can still happen. Having an intermediary does not solve the exclusivity problem and thereby cannot restore the first best, as intermediaries are also subject to potential free-riding. However, the intermediary can restore trade by becoming the sole financier of the borrower. I state this result here as a proposition.

Proposition 3.5. *If $M \geq 4$, and investors can become intermediaries, an equilibrium exists, with $\overline{\vartheta} = 6$.*

3.4 Solving the General Model

In this section, I aim to analyze the general model and show that the intuition from the example carries over to a more general setting. First I analyze the first best and the classic cases of monopoly and exclusive competition. Then the model is analyzed with the restriction that borrowers only deal directly with investors, and it is shown that the market breaks down. After that I address how trade can be restored, first by introducing a social planner who acts as an intermediary and then by allowing investors to become intermediaries.

3.4.1 Benchmark

In order to properly study the effects of competition, I first address, as a benchmark, which allocations maximize the total surplus generated by the project. This total surplus equals

$$F(\overline{\vartheta})(R+B) - \left(1 + \int_0^{\overline{\vartheta}} \vartheta f(\vartheta)\mathrm{d}\vartheta\right).$$

The one choice parameter in this expression is $\overline{\vartheta}$. Taking the first order condition, one finds that a surplus-maximizing allocation should satisfy $\overline{\vartheta} = R + B$.[6]

In order to highlight the difference between the explicit modeling of competition in this paper and the more classical "exclusive" competition model, I model what happens if the borrower is restricted to deal with only one investor at the time. In this case, any offer an investor makes that leaves him any part of the surplus, or that doesn't give the borrower maximum utility, will be undercut by competitors. Thus, the optimal contract under exclusive competition optimizes the borrower's net expected utility

$$F(\overline{\vartheta})(R+B) - \int_0^{\overline{\vartheta}} D(\vartheta) f(\vartheta)\mathrm{d}\vartheta$$

subject to the investor's participation (break-even) constraint

$$1 + \int_0^{\overline{\vartheta}} \vartheta f(\vartheta)\mathrm{d}\vartheta \leq \int_0^{\overline{\vartheta}} D(\vartheta) f(\vartheta)\mathrm{d}\vartheta$$

This break-even constraint is binding, so that the optimization problem reduces to optimizing the total surplus as before. This means that under classic "exclusive" competition, the first best can be attained through direct one-on-one contracting between borrowers and investors.

3.4.2 Market Failure

As in the numerical example, two problems arise if investors contract only directly with borrowers: again, any finite liquidity supply can be extended through offering some additional liquidity in exchange for a diluting share. On the other hand, if multiple investors finance a borrower together, a common agency problem arises: each investor will want to make sure the other ones provide liquidity rather than himself.

[6] This also covers the cases in which $R + B$ is not in the support of ϑ. In those cases, setting $\overline{\vartheta} = R + B$ is equivalent to setting $\overline{\vartheta}$ to the largest value in the support of ϑ that is smaller than or equal to $R + B$, which would be optimal.

Extension of the Liquidity Supply

However, as in Boxtel, Castiglionesi, and Feriozzi (2013), the scope for dilution creates a free-riding opportunity as long as the aggregate supply of liquidity is limited. If a borrower only receives a limited supply of liquidity, an investor could offer to extend the liquidity provision to the borrower by a small amount. With this additional liquidity provision on top of the one provided by the incumbents, the borrower has a larger set of states of the world in which she can continue and obtain at least her private benefit. The investor can offer this additional liquidity in exchange for a small, possibly diluting, share of the project's proceeds, only to be paid in case the additional liquidity is used. The possibility of this deviation leads to the impossibility of limiting liquidity provision, as stated in the following proposition.

Proposition 3.6. *In any equilibrium with non-exclusive competition, one must have* $\overline{\vartheta} = \sup \Theta$.

The intuition behind the proof is, as in section 3.3, that investors who don't have a large share in financing the borrower can always come in and offer a bit of liquidity in exchange for a share of the project, possibly diluting the incumbent investors, while free-riding upon their liquidity provision.

Common Agency Problem

Proposition 3.6 would imply that in equilibrium, any borrower should have any shock financed. As investors' endowments are limited, this means the borrower needs to be financed by a large enough group of investors, as too small a group of investors cannot finance any arbitrarily large shock. This financing arrangement naturally creates a common agency situation in which each investor prefers to adjust the pricing of their liquidity provision in such a way that the borrower prefers getting its $t = 1$ funds from other investors first. This can be done in much the same way as in the example: by giving the borrower a slight discount when he chooses to withdraw less liquidity. As yet, there is no proof yet for a general distribution, but the proof holds for the following special case:

Lemma 3.7. *Let* ϑ *follow a dichotomous distribution, i.e.* $\vartheta = \vartheta_L$ *with probability* p_L *and* $\vartheta = \vartheta_H$ *with probability* $p_H := (1 - p_L)$*, then if*

$$\frac{R - (1 + \mathbf{E}\vartheta)}{p_L} < \vartheta_H - W \tag{3.6}$$

and

$$W \geq \frac{2}{3}, \tag{3.7}$$

then there is no equilibrium with $\overline{\vartheta} = \vartheta_H$.

The intuition behind the above statement is as follows: as the expected income that the borrower can pledge is only limited, and the liquidity support exceeds this limited income, the amount of expected income the borrower pays —the "price"— of a marginal unit of liquidity support, will at some point be smaller than one.

This means that investors receive less than their marginal cost from providing an additional unit of liquidity. Investors will then want to give the borrower an incentive to obtain cheap liquidity from other investors. They can do so by making their liquidity slightly more expensive than the liquidity provided by the other investors, so that the borrower will first obtain his liquidity from the others.

In more mathematical terms, if $\overline{\vartheta} > R$, there must be intervals where the slope of $D^j(\vartheta)$ is smaller than 1. To see how this matters, imagine there are only two investors, say 1 and 2, financing a borrower. As the slope of the aggregate repayment is smaller than one, there are at least two values of ϑ, say ϑ^L and ϑ^H, with $\vartheta_L < \vartheta_H$, for which

$$D(\vartheta^H) - D(\vartheta^L) < \vartheta^H - \vartheta^L$$

The mathematical intuition behind the lemma is easiest if one imagines these two values of ϑ to correspond to two positive probability states: the *H(igh)* and the *L(ow)* state. Call the amounts of liquidity supplied by investors i in each of these two states L_i^L and L_i^H, and the corresponding debt repayments D_i^L and D_i^H,In that case, as the investors together make up the total liquidity supply, it holds that

$$L_1^L + L_2^L = \vartheta^L$$
$$L_1^H + L_2^H = \vartheta^H.$$

Similarly, the aggregate repayments are the repayments to both of the investors combined.

$$D_1^L + D_2^L = D(\vartheta^L)$$
$$D_1^H + D_2^H = D(\vartheta^H)$$

This also entails that at least for one of the investors (say investor 1), it must hold that

$$D_1^H - D_1^L < L_1^H - L_1^L,$$

i.e. at least one of the investors (in casu, investor 1) sells the additional liquidity between L_1^H and L_1^L at a discount. In that case, investor 2 can change the values for the L state to $(\tilde{L}_2^L, \tilde{D}_2^L)$ with

$$\tilde{L}_2^L = \vartheta^L - L_1^H$$

and

$$\tilde{D}_2^L = D_2^L - (D_1^H - D_1^L)$$

i.e. investor 2 asks the borrower to obtain the high liquidity from investor 1, yet compensating the borrower for the additional cost of obtaining this higher liquidity.

In the H state nothing changes, but in the L state, the borrower will still be able to continue by getting L_1^H from investor 1 and $\tilde{L}_2^L$ from investor 2. As can be seen easily, the borrower's utility in this state is unchanged. Investor 2 will, however, now have a utility of $\tilde{D}_2^L - \tilde{L}_2^L$ in the low state, instead of $D_2^L - L_2^L$. Now

$$\begin{aligned}
\tilde{D}_2^L - \tilde{L}_2^L &= D_2^L - (D_1^H - D_1^L) - \left(\vartheta^L - L_1^H\right) \\
&= D_2^L - (D_1^H - D_1^L) - \left(L_2^L + L_1^L - L_1^H\right) \\
&= D_2^L - L_2^L + \left(L_1^H - L_1^L\right) - (D_1^H - D_1^L) \\
&> D_2^L - L_2^L
\end{aligned}$$

so investor 2 will be better off. This gives a mathematical idea of a profitable deviation that exists in any allocation with $\overline{\vartheta}^j > R$ for some j. In this illustration, investor 1 free-rides on the high potential provision of investor 2 in case the low shock hits. When there are more than two investors, one could imagine investor 1 deviating through free-riding in the same way on the aggregate liquidity provision of all the other active investors combined.

The reasoning above is that between two states, if one lender, or a group of lenders, provides incremental liquidity at a discount, other lenders can save on their liquidity by letting the borrower obtain additional liquidity from other lenders. However, in the example with a dichotomous shock, in the notation of the analysis above, it might be so that the jump of $D_1^H - D_1^L$, and thus the discount that lender 2 needs to give, is relatively large and exceeds the potential savings that investor 2 might make by providing less liquidity in the L state, which is bound by L_2^L. Condition (3.6) effectively puts an upper bound on$D(\vartheta^H) - D(\vartheta^L)$ and makes sure that this is not a problem in the proof. Condition (3.7) is needed in this context to make the intuition from the example with two lenders carry over to any case

Conditions (3.6) and (3.7) do not represent any deep economic insight: both conditions are technical and arise as an artefact of the dichotomous distribution. The issues in this would not be a problem if there are no "holes" in the distribution. This reasoning gives the following conjecture.

Conjecture 3.8. *If ϑ follows a continuous distribution with $F(\cdot)$ strictly increasing, any allocation in which $\overline{\vartheta} > R{+}B$, and with more than one lender providing liquidity, is not sustainable in equilibrium.*

No Trade Result

A direct corollary of conjecture 3.8 would be that no allocation in which unlimited liquidity is provided, can be sustained in equilibrium: any such allocation would necessarily have more than one investor providing liquidity and would also have $\overline{\vartheta} = \sup \Theta > R + B$. Together with propostion 3.6, this would entail that *no* equilibrium with trade can exist if lenders deal directly with borrowers.

For the dichotomous case, this gives the following proposition

Proposition 3.9. *Under the conditions of lemma 3.7, no equilibrium with trade exists.*

The way this proposition follows as a corollary from proposition 3.6 and lemma 3.7 can be stated informally as "one is not enough, two is too many:" according to proposition 3.6, in any equilibrium, investors need to provide such a potentially large supply of liquidity, that any borrower needs more than one investor for liquidity provision. However, in case a borrower with a large potential liquidity shock is financed by more than one investor, these investors will always have an incentive to change the contracts they trade.

3.5 Financial Intermediation

In this section, I aim to cover how, in the general model, trade can be restored by an intermediary, thus bypassing the common agency problems: an intermediary, being the only one dealing with a firm, is not bothered by common agency problems. Moreover, by pooling all the funds of different investors, she can provide full insurance, and is therefore not susceptible to the problem of other investors free-riding by extending the liquidity supply.

3.5.1 The Benevolent Bank

Having established that no trade exists under direct contracting between the borrower and investors, I now see whether a social planner can restore trade in this economy. As it turns out, a social planner can restore trade without needing any informational advantage with respect to other participants in the economy. The planner can do so if she can set up an institution that has a first-mover advantage in offering contracts, both to investors and to borrowers. The planner can intervene by pooling the funds of all the investors together and contracting investment, credit lines, and repayments with the borrower. Effectively, she becomes an intermediary.

To see whether this institutional arrangement works, it is important to formally define the timing of the game with the social planner intervening. At $t = 0$ the following events happen. After the institution has played, his timing is the same

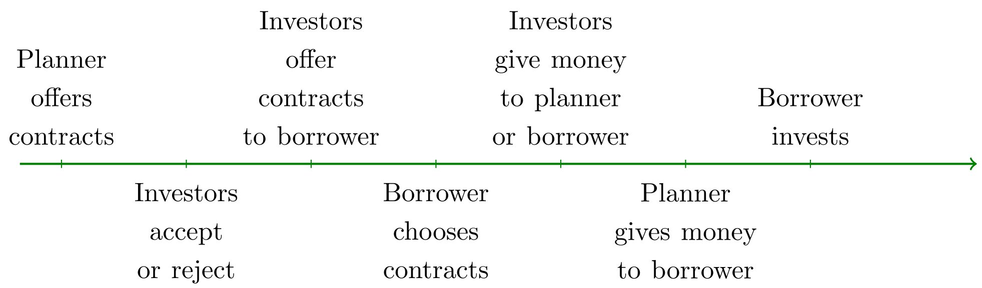

Figure 3.1: Close-up of the timing for the contracting stage with the social planner.

as before and explicitly allows for investors to privately offer contracts to borrowers *after* observing the planner's offers. The elements that are added before and after this serve to enable the planner to offer a mechanism that coordinates competition and to make sure money is passed through the institution at the right moment.

The institution proposes a contract to the borrower providing one unit of money up front, with unlimited liquidity support, and with a repayment that precisely recoups the expected cost of financing and insuring the project. The borrower will be willing to accept this contract. Waiting one round for the investors to offer contracts will not yield a better offer for the borrower, because the investors in the next round, if not depositing, would be competing in the exact same manner as in the non-intermediated case analyzed above. Thus the borrower cannot expect obtaining any better contract by waiting.

To investors, the institution can propose very simple contracts that just offer a unit of money at $t = 2$ in exchange for each unit of money invested at $t = 0$. This makes investors indifferent between depositing at the social planner's institution on the one hand and storing or consuming on the other. For simplicity, I assume as a tie-breaker that when indifferent, investors prefer bringing their money to the institution.

Offering these contracts, the social planner can restore trade. I state this result as a lemma.

Lemma 3.10. *The social planner can offer a contract to the borrower and contracts to investors that establish trade, offer investment and unlimited liquidity insurance to the borrower and have investors break even.*

Even though welfare would be higher with a limited liquidity provision, the planner cannot prevent the borrower from seeking additional liquidity from the lenders. This makes it harder to implement an allocation with limited liquidity provision. However, the planner could "bribe" lenders into depositing their endowment instead of keeping it to potentially offer as an additional liquidity supply to the borrower. To what extent this is possible is dependent upon assumptions on the extent to which the lender can base contracts on what part of their endowments lenders actually deposit, instead of keep. However, whatever these assumptions, if the potential gains by the lenders from not depositing and potentially free-riding are large enough, the total amount that the lender needs in order to bribe all lenders into depositing is larger than the gains from supplying less liquidity. In the dichotomous case, one can therefore easily derive a condition under which the supply of liquidity by the planner must always be unlimited (i.e. $\overline{\vartheta} = \vartheta^H$):

Lemma 3.11. *Under a dichotomous distribution, if*

$$(M-1)p_H\left(R - \left(\vartheta^H - \vartheta^L\right)\right) > p_L R + (1 - p_L\vartheta_L) \tag{3.8}$$

the planner can only implement an equilibrium with $\overline{\vartheta} = \vartheta^H$.

3.5.2 Investors as Intermediaries

Now I examine whether an equilibrium with trade can be sustained if investors can endogenously decide to become intermediaries. For the dichotomous shock case, I can prove the following lemma.

Lemma 3.12. *An equilibrium exists in which one lender becomes an intermediary and provides an unlimited supply of liquidity.*

The proof of this lemma is entirely analogous to the proof of proposition 3.5: one lender becomes an intermediary and collects enough money from other investors to be able to finance up to the high shock. The investors will get their money back at $t = 2$.

3.6 Robustness

3.6.1 Bankruptcy Arrangements

In this paper, full insurance is necessary because entrants can always provide additional insurance, free-riding on the insurance provided by others. This free-riding is made possible by dilution, which, in this paper, happens very explicitly through the pro rata distribution upon default. However, "dilution" can happen through other kinds of externalities between investors. Moral hazard can have a diluting effect (Bizer and DeMarzo, 1992; Boxtel, Castiglionesi, and Feriozzi, 2013) or investors can offer additional liquidity in exchange for increasingly speculative tranches of the borrower's revenue, if the revenue's support is unbounded or very large.

3.6.2 Exclusive Competition with "Small" Investors

Alternatively, one can assume there is no exclusivity problem. In this case, the common agency problem can still give rise to the need for financial intermediation, as long as the optimal liquidity provision exceeds the endowment of each of the investor. Consider again an example such as the one from section 3.3, but assume now that $B = 2$ and that $R = 5 + \varepsilon$ for some "small enough" ε.[7]

As before, any allocation with only one investor financing the borrower, and with no intermediaries, can at most have $\overline{\vartheta} = 2$, as any one investor cannot finance liquidity shocks up to $\vartheta = 6$, due to the limited endowments. However, in any allocation with $\overline{\vartheta} = 2$, the total surplus, which is an upper bound for the borrower's utility, is at most

$$\frac{1}{2}(R + B) - \left(1 + \frac{1}{2}\vartheta_L\right) = 1\frac{1}{2} + \frac{1}{2}\varepsilon.$$

Now, if trade between investors is possible, one inactive investor, say i, can offer the following two contracts. From another inactive investor he asks 2 units of money at $t = 0$, offering to return $2 + \frac{1}{3}\varepsilon$ at $t = 2$. To the borrower he offers a contract with $J_i = 1$, with $L_i^L = 2$ and $L_i^H = 6$. As for the repayment, he can ask $D_i^L = 5 + \frac{1}{3}\varepsilon$ and $D_i^H = 5 + \varepsilon$. If the borrower accepts this contract, she will reveal the liquidity shock at $t = 1$ and continue in all cases, giving her a liquidity

$$R + B - \left(\frac{1}{2}D_i^L + \frac{1}{2}D_i^H\right) = 2 + \frac{1}{3}\varepsilon,$$

which, as ε is "small enough", is greater than the maximum utility the borrower can have if $\overline{\vartheta} = 2$, so that the borrower will accept this contract.

The depositing investor will then have a utility increase of $\frac{1}{3}\varepsilon$, meaning he will be willing to accept this deviation. The deviating investor will have a utility of

$$\frac{1}{2}D_i^L + \frac{1}{2}D_H^i - \left(J_i + \frac{1}{2}L_i^L + \frac{1}{2}L_i^H\right) + 2 - \left(2 + \frac{1}{3}\varepsilon\right) = \frac{1}{3}\varepsilon.$$

[7] As will be shown later on, this ε is just needed for certain deviations to be present. The assumption that it is there can also be taken away by tie-breakers. As long as it is positive, ε can be as close to zero as we want it to be, and the analysis below will still hold.

This entails that this deviation is profitable for investor i, giving the following result (the proof of which is exactly the above analysis).

Lemma 3.13. *In case banks can trade among one another, any allocation with* $\overline{\vartheta} = 2$ *cannot be sustained in equilibrium.*

As in section 3.3, an allocation with two investors separately financing the borrower is also not possible, as the proof of lemma 3.3 carries over exactly to this context. This entails that no equilibrium with investors independently financing the borrower can exist. Also, no equilibrium with only one investor can exist, as this will need to have $\overline{\vartheta} = 2$. A candidate equilibrium is suggested by the deviating investor's strategy in the proof of lemma 3.13: one investor can become an intermediary for another investor. This gives us the following equilibrium.

Proposition 3.14. *An equilibrium exists with one investor becoming an intermediary collecting at least* 2 *units of money from other investors, offering a zero interest rate, and fully insuring the borrower's liquidity shocks.*

3.6.3 Type of Shock

The model above very much dealt with liquidity provision. The intermediate date shock is intended to represent a more generic piece of private information that plays a role at an intermediate stage, such as profitability of future cash flows.

3.7 Conclusion

An interesting axis for further research could be to add a trade-off between bank financing and decentralized financing and see if there could be some further light shed on the emergence of different types of financial systems in different times and in different countries. Also, it would be very interesting to see whether banks in a set-up like this one can extract rents, which would then feed back into a trade-off.

This paper does not offer so much an alternative to Diamond (1984) as a complementary explanation. Indeed, in many ways the two explanations feed back into each other. An intermediary as described in this paper would get his profits very strongly tied to those of a firm it is financing, thus giving this intermediary a very strong incentive to monitor. It would be very interesting to disentangle, both empirically and theoretically, the "organizing" role of intermediaries described in this paper and the delegated monitoring role in Diamond (1984).

It should be noted that in this paper, it is always one of the investors who becomes an intermediary. In principle, nothing stops borrowers from becoming intermediaries. Whereas the former has a natural interpretation of an investor becoming a bank, the latter can be interpreted naturally as the emergence of a trust. Theoretically, in the model in this paper, J.P. Morgan was just as likely to emerge as U.S. Steel.

Throughout the paper, I have followed Holmström and Tirole (1998) in assuming the borrower cannot withdraw more investment or liquidity than needed. Of course this assumption becomes more problematic when the liquidity provision is unlimited.

So monitoring or relationship banking in order for investors to verify the actual use of the liquidity provision becomes more of an issue in the non-exclusive set-up, in which unlimited liquidity needs to be provided. This might explain why universal banking models have been more successful in the context considered in this paper: banks that sit on boards and hold equity stakes are more likely to know private information and are more prone to monitor. This might merit some study of its own.

In this paper, the intermediary can always meet its liquidity promises. An interesting point for further research is to see what happens when liquidity potentially becomes scarce. Competition, by stimulating overinsurance, could lead to new (systemic) instabilities.

Bibliography

Acharya, V., S.A. Davydenko, and I.A. Strebulaev. 2012. "Cash holdings and credit risk." *Review of Financial Studies* :hhs106. 57

Agrawal, A. and G.N. Mandelker. 1987. "Managerial Incentives and Corporate Investment and Financing Decisions." *The Journal of Finance* 42 (4). 22

Alighieri, D. and J.D. Sinclair. 1961. *The Divine Comedy: Volume 1: Inferno*. A Galaxy book ; GB65-67. Oxford University Press, USA. 41

Allen, F. and G.B. Gorton. 1993. "Churning Bubbles." *The Review of Economic Studies* 60 (4). 22, 23

Arnott, R.J. and J.E. Stiglitz. 1991. "Equilibrium in competitive insurance markets with moral hazard." Tech. rep., National Bureau of Economic Research. 61, 62

Attar, A., C. Casamatta, A. Chassagnon, and J.P. Décamps. 2010. "Multiple Lenders, Strategic Default and the role of Debt Covenants." . 43, 53, 62

Attar, A. and A. Chassagnon. 2009. "On moral hazard and nonexclusive contracts." *Journal of Mathematical Economics* 45 (9-10):511–525. 43, 62

Attar, A., T. Mariotti, and F. Salanié. 2011. "Nonexclusive competition in the market for lemons." *Econometrica* 79 (6):1869–1918. 43, 63

Attar, A., Th. Mariotti, and F. Salanié. 2014. "Nonexclusive competition under adverse selection." *Theoretical Economics* 9 (1):1–40. 43, 63

Bannier, C.E., E. Feess, and N. Packham. 2012. "Competition, bonuses, and risk-taking in the banking industry." *Review of Finance* :rfs002. 25

Bates, T.W., K.M. Kahle, and R.M. Stulz. 2009. "Why do US firms hold so much more cash than they used to?" *The Journal of Finance* 64 (5):1985–2021. 41

Beales, R. 2008. "Bad assets become bonuses at Credit Suisse." *The New York Times* . 24

Beaudry, P. and M. Poitevin. 1995. "Competitive screening in financial markets when borrowers can recontract." *The Review of Economic Studies* 62 (3):401–423. 43

Becht, M. and C.D. Ramírez. 2003. "Does bank affiliation mitigate liquidity constraints? Evidence from Germany's universal banks in the pre-World War I period." *Southern Economic Journal* :254–272. 64

Bénabou, R. and J. Tirole. 2013. "Bonus Culture: Competitive Pay, Screening, and Multitasking." Tech. rep., National Bureau of Economic Research. 24, 25

Berk, J.B. and R.C. Green. 2004. "Mutual Fund Flows and Performance in Rational Markets." *The Journal of Political Economy* 112 (6). 23

Bernard, A.B., J.Y. Campbell, J.H. Cochrane, D.W. Diamond, D. Duffie, A.K. Kashyap, F. Mishkin, R.G. Rajan, R.J. Shiller, M.J. Slaughter, J.C. Stein, and R.M. Stulz. 2010. "Regulation of Executive Compensation in Financial Services." 20, 23

Bernheim, B.D. and M.D. Whinston. 1986. "Common agency." *Econometrica: Journal of the Econometric Society* :923–942. 61, 63

Bertrand, J. 1883. "Review of Cournot (1838)." *Journal des Savants* 67:499–508. 62, 70

Bhattacharya, S. and P. Pfleiderer. 1985. "Delegated portfolio management." *Journal of Economic Theory* 36 (1). 22, 24

Bijlsma, M., J. Boone, and G. Zwart. 2012. "Competition for traders and risk." . 24, 25

Bisin, A. and D. Guaitoli. 2004. "Moral hazard and nonexclusive contracts." *RAND Journal of Economics* :306–328. 43, 62

Bizer, D.S. and P.M. DeMarzo. 1992. "Sequential banking." *Journal of Political Economy* :41–61. 43, 61, 62, 80

Bolton, P., H. Mehran, and J.D. Shapiro. 2010. "Executive Compensation and Risk Taking." *SSRN eLibrary* . 21

Bolton, P., T. Santos, and J.A. Scheinkman. 2011. "Cream Skimming in Financial Markets." Working Paper 16804, National Bureau of Economic Research. 19

Boot, A.W.A. 1999. "European lessons on consolidation in banking." *Journal of Banking & Finance* 23 (2):609–613. 61

Boxtel, A.A. van, F. Castiglionesi, and F. Feriozzi. 2013. "Credit Market Competition and Firm Liquidity." Working paper — *also appears as Chapter 2 of this thesis*. 60, 61, 63, 68, 70, 71, 75, 80

Braggion, F. and S. Ongena. 2011. "A century of British firm-bank relationships: why the transition to multiple banking." Tech. rep., Working Paper, Tilburg University. 42

Calomiris, C.W. 1993. "Corporate-finance benefits from universal banking: Germany and the United States, 1870-1914." Tech. rep., National bureau of economic research. 64

———. 1995. "The costs of rejecting universal banking: American finance in the German mirror, 1870-1914." In *Coordination and information: Historical perspectives on the organization of enterprise*. University of Chicago Press, 257–322. 64

———. 1999. "Gauging the efficiency of bank consolidation during a merger wave." *Journal of Banking & Finance* 23 (2):615–621. 61

Calomiris, C.W. and J. Karceski. 2000. "Is the bank merger wave of the 1990s efficient? Lessons from nine case studies." In *Mergers and productivity*. University of Chicago Press, 93–178. 61

Castiglionesi, F. and W.B. Wagner. 2013. "On the efficiency of bilateral interbank insurance." *Journal of Financial Intermediation* 22 (2):177–200. 43, 61, 63

Chandler, A.D. 1954. "Patterns of American Railroad Finance, 183050." *Business History Review* 28:248–263. 64

Chen, C.R., Th.L. Steiner, and A.M. Whyte. 2006. "Does stock option-based executive compensation induce risk-taking? An analysis of the banking industry." *Journal of Banking and Finance* 30 (3):915 – 945. 22

Cheng, I.H., H. Hong, and J.A. Scheinkman. 2010. "Yesterday's Heroes: Compensation and Creative Risk-Taking." Working Paper 16176, National Bureau of Economic Research. 21

Créquy, P. 2009. "Sarkozy : pas de bonus pour les traders sans malus." *Le Figaro* . 24

Cuñat, V. and M. Guadalupe. 2009. "Executive compensation and competition in the banking and financial sectors." *Journal of Banking and Finance* 33 (3). 22

Da Rin, M. 1996. "Understanding the development of the German Kreditbanken, 1850-1914: an approach from the economics of information." *Financial History Review* 3:29–48. 64

Da Rin, M. and Th. Hellmann. 2002. "Banks as catalysts for industrialization." *Journal of Financial Intermediation* 11 (4):366–397. 64

Das, S.R. and R.K. Sundaram. 2002. "Fee Speech: Signaling, Risk-Sharing, and the Impact of Fee Structures on Investor Welfare." *Review of Financial Studies* 15 (5). 23, 24

Dasgupta, A. and A. Prat. 2008. "Information aggregation in financial markets with career concerns." *Journal of Economic Theory* 143 (1). 23

Davis, L.E. 1963. "Capital immobilities and finance capitalism: A study of economic evolution in the United States 1820-1920." *Explorations in entrepreneurial history : EEH* 1 (1). 64

DeLong, J.B. 1991. "Did JP Morgans Men Add Value? An Economist's Perspective on Financial Capitalism." In *Inside the business enterprise: Historical perspectives on the use of information*. University of Chicago Press, 1991, 205–250. 64

Dewatripont, M. and E.S. Maskin. 1995. "Credit and efficiency in centralized and decentralized economies." *The Review of Economic Studies* 62 (4):541–555. 65

DeYoung, R., E.Y. Peng, and M. Yan. 2013. "Executive compensation and business policy choices at US commercial banks." Tech. Rep. 01. 21

Diamond, D.W. 1984. "Financial intermediation and delegated monitoring." *The Review of Economic Studies* 51 (3):393–414. 59, 64, 65, 81

Diamond, D.W. and P.H. Dybvig. 1983. "Bank runs, deposit insurance, and liquidity." *The Journal of Political Economy* :401–419. 59, 65

Diamond, D.W. and R.G. Rajan. 2001. "Liquidity Risk, Liquidity Creation, and Financial Fragility: A Theory of Banking." *Journal of Political Economy* 109 (2). 59, 65

Diamond, P. 1998. "Managerial incentives: On the near linearity of optimal compensation." *Journal of Political Economy* 106 (5):931–957. 21, 25

Duchin, R. 2010. "Cash holdings and corporate diversification." *The Journal of Finance* 65 (3):955–992. 41

Fahlenbrach, R. and R.M. Stulz. 2011. "Bank CEO incentives and the credit crisis." *Journal of Financial Economics* 99 (1). 22

Foster, D.P. and H.P. Young. 2010. "Gaming Performance Fees By Portfolio Managers." *The Quarterly Journal of Economics* 125 (4). 22

Froot, K.A., D.S. Scharfstein, and J.C. Stein. 1992. "Herd on the Street: Informational Inefficiencies in a Market with Short-Term Speculation." *The Journal of Finance* 47 (4). 23

Gerschenkron, Alexander et al. 1962. "Economic backwardness in historical perspective." *Economic backwardness in historical perspective.* . 63, 64

Goldoni, C. 1745. *Il servitore di due padroni*. 9, 13

Grinblatt, M.S., M. Keloharju, and J.T. Linnainmaa. 2012. "IQ, trading behavior, and performance." *Journal of Financial Economics* 104 (2):339–362. 23

Gruber, M.J. 1996. "Another Puzzle: The Growth in Actively Managed Mutual Funds." *The Journal of Finance* 51 (3). 23

Guinnane, T.W. 2002. "Delegated monitors, large and small: Germany's banking system, 1800-1914." *Journal of economic Literature* :73–124. 64

Hodder, J.E. and J.C. Jackwerth. 2007. "Incentive Contracts and Hedge Fund Management." *Journal of Financial and Quantitative Analysis* 42 (4). 23

Holmstrom, B. and J. Tirole. 1997. "Financial intermediation, loanable funds, and the real sector." *The Quarterly Journal of Economics* 112 (3):663. 65

Holmström, B. and J. Tirole. 1998. "Private and Public Supply of Liquidity." *Journal of Political Economy* :1–40. 17, 41, 42, 44, 46, 49, 50, 59, 66, 67, 81, 89

Holmström, B. and J. Tirole. 2011. *Inside and outside liquidity.* The MIT Press. 42, 44, 46

Hoshi, T., A.K. Kashyap, and D.S. Scharfstein. 1991. "Corporate structure, liquidity, and investment: Evidence from Japanese industrial groups." *The Quarterly Journal of Economics* 106 (1):33–60. 64

Houston, J.F. and C. James. 1995. "CEO compensation and bank risk Is compensation in banking structured to promote risk taking?" *Journal of Monetary Economics* 36 (2). 22

Hubbard, R.G. and D. Palia. 1995. "Executive pay and performance Evidence from the U.S. banking industry." *Journal of Financial Economics* 39 (1). 22

Jarque, A. and E.S. Prescott. 2010. "Optimal bonuses and deferred pay for bank employees: implications of hidden actions with persistent effects in time." Working Paper 10-16, Federal Reserve Bank of Richmond. 23

Jeidels, O. 1905. *Das Verhältnis der deutschen Grossbanken zur Industrie: mit besonderer Berücksichtigung der Eisenindustrie*, vol. 24. Duncker & Humblot. 63

Kahn, C.M. and D. Mookherjee. 1998. "Competition and incentives with nonexclusive contracts." *The Rand Journal of Economics* :443–465. 43

Karceski, J., S. Ongena, and D.C. Smith. 2005. "The impact of bank consolidation on commercial borrower welfare." *The Journal of Finance* 60 (4):2043–2082. 61

Kneer, C. 2013. "The Absorption of Talent into Finance: Evidence from US Banking Deregulation." Tech. rep., Netherlands Central Bank, Research Department. 19

Lewis, W.A. 1978. *Growth and Fluctuations 1870-1913 (Routledge Revivals).* Routledge. 64

Malkiel, B.G. 1995. "Returns from Investing in Equity Mutual Funds 1971 to 1991." *The Journal of Finance* 50 (2). 23

———. 2003. "The Efficient Market Hypothesis and Its Critics." *The Journal of Economic Perspectives* 17 (1). 23

Martimort, D. and L. Stole. 2002. “The revelation and delegation principles in common agency games.” *Econometrica* 70 (4):1659–1673. 61, 63

Murphy, K.M., A. Shleifer, and R.W. Vishny. 1991. “The Allocation of Talent: Implications for Growth.” *The Quarterly Journal of Economics* :503–530. 19

Palomino, F. and A. Prat. 2003. “Risk Taking and Optimal Contracts for Money Managers.” *The RAND Journal of Economics* 34 (1). 22, 24

Parlour, C.A. and U. Rajan. 2001. “Competition in loan contracts.” *American Economic Review* :1311–1328. 43, 62

Pauly, M.V. 1974. “Overinsurance and public provision of insurance: the roles of moral hazard and adverse selection.” *The Quarterly Journal of Economics* 88 (1):44–62. 62

Peters, M. 2001. “Common agency and the revelation principle.” *Econometrica* 69 (5):1349–1372. 63

Rajan, R.G. 1992. “Insiders and Outsiders: The Choice between Informed and Arm’s-Length Debt.” *The Journal of Finance* 47 (4). 65

Ramirez, C.D. 1995. “Did JP Morgan’s men add liquidity? Corporate investment, cash flow, and financial structure at the turn of the twentieth century.” *The Journal of Finance* 50 (2):661–678. 64

Riesser, J. 1910. *Die deutschen Großbanken und ihre Konzentration*. 63

Steinbeck, J.E. 1939. *The Grapes of Wrath*. Penguin Classics. 59

Stracca, L. 2006. “Delegated Portfolio Management: A Survey of the Theoretical Literature.” *Journal of Economic Surveys* 20 (5). 23

Thanassoulis, J.E. 2011. “Bankers’ Pay Structure And Risk.” Economics Series Working Papers 545, University of Oxford, Department of Economics. 21, 23, 25

———. 2012. “The case for intervening in bankers? pay.” *The Journal of Finance* 67 (3):849–895. 21

Tirole, J. 2006. *The theory of corporate finance*. Princeton University Press. 44, 46

Appendix A

Proofs for Chapter 2

Proof of Proposition 2.2. The proof of the competition benchmark case proceeds in three steps: first, we note that, because of exclusive competition, the allocation is necessarily the one that maximizes the firm's total surplus, subject to the lender's break-even constraint. The argument is a simple and classic one: in any other allocation, an entrant would have the possibility to undercut the allocation and the firm would switch to doing business with the entrant. Because of exclusive competition, there is no way for the firm to improve on this by seeking extra financing, as it is restricted to deal with only one lender at a time. Then, we find the optimal allocation $\left(I, \overline{\vartheta}, D(\cdot)\right)$, where $D(\cdot)$ is defined on $\left[0, \overline{\vartheta}\right]$. The difference to the proof in Holmström and Tirole (1998) is that we allow $D(\cdot)$ to be any non-decreasing function. In order to show that, as in Holmström and Tirole (1998), the solution with the constant $D(\vartheta) = \left(R - \frac{B}{\Delta p}\right) I$ is the optimal one, we have to make a small limiting argument: for $D(\cdot)$ we take a class of piecewise constant functions on equidistant intervals, then arguing that, for arbitrarily small intervals, the solution remains the same with $\overline{\vartheta} = \vartheta^*$, with $D(\vartheta) = \left(R - \frac{B}{\Delta p}\right) I$ for all ϑ, and with

$$I = \frac{A}{F(\vartheta^*)(\vartheta^* - \rho_0)}.$$

The type of function we choose for $D(\cdot)$ is, for an arbitrary integer N, characterized by the parameters $\underline{\vartheta}$, the liquidity level up to which the total repayment is incentive compatible, and $\overline{\vartheta}$, the maximum liquidity level. It is further characterized by D_n^L, for $n = 1$ through N, the repayment on interval $\left[\frac{(n-1)\underline{\vartheta}}{N}, \frac{n\underline{\vartheta}}{N}\right]$, and by D_n^H, for $n = 1$ through N, the repayment on interval $\left[\underline{\vartheta} + \frac{(n-1)\left(\overline{\vartheta}-\underline{\vartheta}\right)}{N}, \underline{\vartheta} + \frac{n\left(\overline{\vartheta}-\underline{\vartheta}\right)}{N}\right]$. Thus, $D(\cdot)$ can be written in the form

$$\begin{aligned} D(\vartheta) = &\sum_{n=1}^{N} D_n^L \mathbf{1}\left(\vartheta \in \left[\frac{(n-1)\underline{\vartheta}}{N}, \frac{n\underline{\vartheta}}{N}\right]\right) \\ &+ \sum_{n=1}^{N} D_n^H \mathbf{1}\left(\vartheta \in \left[\underline{\vartheta} + \frac{(n-1)\left(\overline{\vartheta}-\underline{\vartheta}\right)}{N}, \underline{\vartheta} + \frac{n\left(\overline{\vartheta}-\underline{\vartheta}\right)}{N}\right]\right), \end{aligned}$$

where **1** is an indicator, equal to one if its argument is true and zero if it is not. The firm's utility can then be written as

$$\sum_{n=1}^{N} p_H \left(RI - D_n^L\right) \left(F\left(\frac{n\underline{\vartheta}}{N}\right) - F\left(\frac{(n-1)\,\underline{\vartheta}}{N}\right)\right) + \sum_{n=1}^{N} \left((p_L R + B)\, I - p_L D_n^H\right) \cdots$$
$$\cdots \left(F\left(\underline{\vartheta} + \frac{n\left(\overline{\vartheta} - \underline{\vartheta}\right)}{N}\right) - F\left(\underline{\vartheta} + \frac{(n-1)\left(\overline{\vartheta} - \underline{\vartheta}\right)}{N}\right)\right).$$

Before we address the maximization problem, we note that $\underline{\vartheta} > 0$ and that $F\left(\underline{\vartheta}\right) > 0$, as otherwise the NPV of the project is negative, and no surplus is divided at all. Also note that the proposed solution generates a positive surplus, so the optimal allocation must also generate a positive total surplus. The lender's costs equal $\left(1 + \int_0^{\overline{\vartheta}} \vartheta f(\vartheta)\, \mathrm{d}\vartheta\right) I - A$, and his revenues equal

$$\sum_{n=1}^{N} p_H D_N^L \left(F\left(\frac{n\underline{\vartheta}}{N}\right) - F\left(\frac{(n-1)\,\underline{\vartheta}}{N}\right)\right)$$
$$+ \sum_{n=1}^{N} p_L D_N^H \left(F\left(\underline{\vartheta} + \frac{n\left(\overline{\vartheta} - \underline{\vartheta}\right)}{N}\right) - F\left(\underline{\vartheta} + \frac{(n-1)\left(\overline{\vartheta} - \underline{\vartheta}\right)}{N}\right)\right).$$

We can now state the maximization problem as

$$\max_{I,\underline{\vartheta},\overline{\vartheta},(D_n^L)_{n=1}^N,(D_n^H)_{n=1}^N} \left\{ \sum_{n=1}^{N} p_H \left(RI - D_n^L\right) \left(F\left(\frac{n\underline{\vartheta}}{N}\right) - F\left(\frac{(n-1)\,\underline{\vartheta}}{N}\right)\right) \right.$$
$$+ \sum_{n=1}^{N} \left((p_L R + B)\, I - p_L D_n^H\right) \cdots$$
$$\left. \cdots \left(F\left(\underline{\vartheta} + \frac{n\left(\overline{\vartheta} - \underline{\vartheta}\right)}{N}\right) - F\left(\underline{\vartheta} + \frac{(n-1)\left(\overline{\vartheta} - \underline{\vartheta}\right)}{N}\right)\right) \right\},$$

subject to the lender's break-even constraint

$$\left(1 + \int_0^{\overline{\vartheta}} \vartheta f(\vartheta)\, \mathrm{d}\vartheta\right) I - A$$
$$\leq \sum_{n=1}^{N} p_H D_n^L \left(F\left(\frac{n\underline{\vartheta}}{N}\right) - F\left(\frac{(n-1)\,\underline{\vartheta}}{N}\right)\right)$$
$$+ \sum_{n=1}^{N} p_L D_n^H \left(F\left(\underline{\vartheta} + \frac{n\left(\overline{\vartheta} - \underline{\vartheta}\right)}{N}\right) - F\left(\underline{\vartheta} + \frac{(n-1)\left(\overline{\vartheta} - \underline{\vartheta}\right)}{N}\right)\right),$$

to which we associate the Lagrange multiplier λ, the N different incentive compatibility constraints

$$p_H D_n^L \leq p_H \left(R - \frac{B}{\Delta p}\right) I, \text{ for } n = 1, 2, \ldots, N,$$

with the associated Lagrange multipliers $(\mu_n)_{n=1}^N$, the N different feasibility constraints

$$p_L D_n^H \leq p_L RI, \text{ for } n = 1, 2, \ldots, N,$$

with the associated Lagrange multipliers $(\nu_n)_{n=1}^N$, and finally the constraint that

$$\underline{\vartheta} \leq \overline{\vartheta}$$

to which we associate the Lagrange multiplier ξ. With respect to I, the first order condition is

$$\begin{aligned}
&\sum_{n=1}^{N} p_H R \left(F\left(\frac{n\underline{\vartheta}}{N}\right) - F\left(\frac{(n-1)\underline{\vartheta}}{N}\right) \right) \\
&\quad + \sum_{n=1}^{N} (p_L R + B) \left(F\left(\underline{\vartheta} + \frac{n(\overline{\vartheta} - \underline{\vartheta})}{N}\right) - F\left(\underline{\vartheta} + \frac{(n-1)(\overline{\vartheta} - \underline{\vartheta})}{N}\right) \right) \\
&\qquad = \lambda \left(1 + \int_0^{\overline{\vartheta}} \vartheta f(\vartheta)\, \mathrm{d}\vartheta\right) - \sum_{n=1}^{N} p_H R \mu_n - \sum_{n=1}^{N} p_L R \nu_n,
\end{aligned}$$

which, by condensing the telescoping sum[1], can be rewritten as

$$\begin{aligned}
&p_H R F(\underline{\vartheta}) + (p_L R + B)\left(F(\overline{\vartheta}) - F(\underline{\vartheta})\right) \\
&\qquad = \lambda \left(1 + \int_0^{\overline{\vartheta}} \vartheta f(\vartheta)\, \mathrm{d}\vartheta\right) - \sum_{n=1}^{N} p_H \left(R - \frac{B}{\Delta p}\right) \mu_n - \sum_{n=1}^{N} p_L R \nu_n.
\end{aligned} \tag{A.1}$$

With respect to $\underline{\vartheta}$, the first order condition is

$$\begin{aligned}
&(p_H R - (p_L R + B)) f(\underline{\vartheta}) I \\
&\quad = \sum_{n=1}^{N} (1-\lambda) p_H D_n^L \left(\frac{n}{N} f\left(\frac{n\underline{\vartheta}}{N}\right) - \frac{n-1}{N} f\left(\frac{(n-1)\underline{\vartheta}}{N}\right) \right) \\
&\qquad + \sum_{n=1}^{N} (1-\lambda) p_L D_n^H \cdots \\
&\qquad\quad \cdots \left(\frac{N-n}{N} f\left(\underline{\vartheta} + \frac{n(\overline{\vartheta} - \underline{\vartheta})}{N}\right) - \frac{N-(n-1)}{N} f\left(\underline{\vartheta} + \frac{(n-1)(\overline{\vartheta} - \underline{\vartheta})}{N}\right) \right) \\
&\qquad + \xi.
\end{aligned} \tag{A.2}$$

[1] i.e., $\sum_{n=1}^{N} (a_n - a_{n-1}) = a_N - a_0$. Henceforth, in this proof, we will use this identity a number of times without explicitly mentioning it.

With respect to $\overline{\vartheta}$, the first order condition is

$$\begin{aligned}
&(p_L R + B) f\left(\overline{\vartheta}\right) I \\
&= \lambda \overline{\vartheta} f\left(\overline{\vartheta}\right) I + \sum_{n=1}^{N} (1-\lambda) p_L D_n^H \cdots \\
&\cdots \left(\frac{n}{N} f\left(\underline{\vartheta} + \frac{n\left(\overline{\vartheta} - \underline{\vartheta}\right)}{N} \right) - \frac{n-1}{N} f\left(\underline{\vartheta} + \frac{(n-1)\left(\overline{\vartheta} - \underline{\vartheta}\right)}{N} \right) \right) \\
&- \xi.
\end{aligned} \tag{A.3}$$

With respect to each of the D_n^L, there is the following equation:

$$(\lambda - 1) p_H \left(F\left(\frac{n\underline{\vartheta}}{N}\right) - F\left(\frac{(n-1)\underline{\vartheta}}{N}\right) \right) = p_H \mu_n, \tag{A.4}$$

and likewise with respect to each of the D_n^H, there is

$$(\lambda - 1) p_L \left(F\left(\underline{\vartheta} + \frac{n\left(\overline{\vartheta} - \underline{\vartheta}\right)}{N} \right) - F\left(\underline{\vartheta} + \frac{(n-1)\left(\overline{\vartheta} - \underline{\vartheta}\right)}{N} \right) \right) = p_L \nu_n. \tag{A.5}$$

Getting the expressions for μ_n and ν_n from equations (A.4) and (A.5), and substituting these into (A.1), we obtain (recalling the notation $\rho_1 := p_H R$ for total income and $\rho_0 := p_H \left(R - \frac{B}{\Delta p}\right)$ for pledgeable income)

$$\begin{aligned}
&\rho_1 F\left(\underline{\vartheta}\right) + (p_L R + B)\left(F\left(\overline{\vartheta}\right) - F\left(\underline{\vartheta}\right)\right) \\
&= \lambda \left(1 + \int_0^{\overline{\vartheta}} \vartheta f(\vartheta)\, \mathrm{d}\vartheta\right) - (\lambda - 1)\rho_0 F\left(\underline{\vartheta}\right) - (\lambda - 1) p_L R \left(F\left(\overline{\vartheta}\right) - F\left(\underline{\vartheta}\right)\right),
\end{aligned}$$

which can be rewritten as

$$\begin{aligned}
&(\rho_1 - \rho_0) F\left(\underline{\vartheta}\right) + B\left(F\left(\overline{\vartheta}\right) - F\left(\underline{\vartheta}\right)\right) \\
&\qquad = \lambda \left(1 + \int_0^{\overline{\vartheta}} \vartheta f(\vartheta)\, \mathrm{d}\vartheta - \rho_0 F\left(\underline{\vartheta}\right) - p_L R \left(F\left(\overline{\vartheta}\right) - F\left(\underline{\vartheta}\right)\right)\right),
\end{aligned}$$

so

$$\lambda = \frac{(\rho_1 - \rho_0) F\left(\underline{\vartheta}\right) + B\left(F\left(\overline{\vartheta}\right) - F\left(\underline{\vartheta}\right)\right)}{1 + \int_0^{\overline{\vartheta}} \vartheta f(\vartheta)\, \mathrm{d}\vartheta - \rho_0 F\left(\underline{\vartheta}\right) - p_L R \left(F\left(\overline{\vartheta}\right) - F\left(\underline{\vartheta}\right)\right)}. \tag{A.6}$$

As stated before, the total surplus must be positive under the optimal solution, so that

$$\rho_1 F\left(\underline{\vartheta}\right) + (p_L R + B)\left(F\left(\overline{\vartheta}\right) - F\left(\underline{\vartheta}\right)\right) > 1 + \int_0^{\overline{\vartheta}} \vartheta f(\vartheta)\, \mathrm{d}\vartheta,$$

which gives

$$\begin{aligned}
&(\rho_1 - \rho_0) F\left(\underline{\vartheta}\right) + B\left(F\left(\overline{\vartheta}\right) - F\left(\underline{\vartheta}\right)\right) \\
&\qquad > 1 + \int_0^{\overline{\vartheta}} \vartheta f(\vartheta)\, \mathrm{d}\vartheta - \rho_0 F\left(\underline{\vartheta}\right) - p_L R \left(F\left(\overline{\vartheta}\right) - F\left(\underline{\vartheta}\right)\right),
\end{aligned}$$

so that necessarily $\lambda > 1$ (i.e., the break-even constraint for the firm must bind) and thus $\mu_n > 0$ for all n. Note that we cannot draw the same conclusion about ν_n, as potentially $\underline{\vartheta} = \overline{\vartheta}$. As $\mu_n > 0$, we also have $D_n^L = \left(R - \frac{B}{\Delta p}\right) I$. This means that condition (A.2) can be rewritten as:

$$\begin{aligned}
&\left(\left(\rho_1 - \rho_0\right) - \left(p_L R + B\right)\right) f\left(\underline{\vartheta}\right) I \\
&\quad = \xi - \lambda \rho_0 f\left(\underline{\vartheta}\right) I + \sum_{n=1}^{N} \left(1 - \lambda\right) p_L D_n^H \cdots \\
&\qquad \cdots \left(\frac{N-n}{N} f\left(\underline{\vartheta} + \frac{n\left(\overline{\vartheta} - \underline{\vartheta}\right)}{N}\right) - \frac{N-(n-1)}{N} f\left(\underline{\vartheta} + \frac{(n-1)\left(\overline{\vartheta} - \underline{\vartheta}\right)}{N}\right)\right).
\end{aligned} \tag{A.7}$$

Now we need to show that $\underline{\vartheta} = \overline{\vartheta}$. In order to do so, assume by contradiction that $\underline{\vartheta} < \overline{\vartheta}$. In that case, all the ν_n's must be greater than zero, so that by equations (A.5) and $D_n^H = RI$ for all n, and also ξ must be zero, so that the above equation becomes

$$\left(\left(\rho_1 - \rho_0\right) I - \left(p_L R + B\right) I\right) f\left(\underline{\vartheta}\right) = -\lambda \rho_0 f\left(\underline{\vartheta}\right) I + \left(\lambda - 1\right) p_L R f\left(\underline{\vartheta}\right) I.$$

Dividing out $f\left(\underline{\vartheta}\right) I$, we get

$$\left(\rho_1 - \rho_0\right) - B = -\lambda \left(\rho_0 - p_L R\right)$$

Furthermore, because $\lambda > 0$, the break-even constraint binds and we get that

$$\lambda = -\frac{\left(\rho_1 - \rho_0\right) - B}{\rho_0 - p_L R}. \tag{A.8}$$

Observing that $\rho_1 - \rho_0 = \frac{p_H}{\Delta p} B$ and using the other expression (A.6) for λ, we find that

$$\frac{\frac{p_H}{\Delta p} B F\left(\underline{\vartheta}\right) + B\left(F\left(\overline{\vartheta}\right) - F\left(\underline{\vartheta}\right)\right)}{c\left(\overline{\vartheta}\right) F\left(\overline{\vartheta}\right) - \rho_0 F\left(\underline{\vartheta}\right) - p_L R\left(F\left(\overline{\vartheta}\right) - F\left(\underline{\vartheta}\right)\right)} = -\frac{\frac{p_L}{\Delta p} B}{\rho_0 - p_L R}.$$

Taking out the B and cross-multiplying, we find that

$$\begin{aligned}
&-\frac{p_L}{\Delta p}\left(c\left(\overline{\vartheta}\right) F\left(\overline{\vartheta}\right) - \rho_0 F\left(\underline{\vartheta}\right) - p_L R\left(F\left(\overline{\vartheta}\right) - F\left(\underline{\vartheta}\right)\right)\right) \\
&\qquad = \left(\rho_0 - p_L R\right)\left(\frac{p_H}{\Delta p} F\left(\underline{\vartheta}\right) + \left(F\left(\overline{\vartheta}\right) - F\left(\underline{\vartheta}\right)\right)\right).
\end{aligned}$$

Multiplying both sides by Δp and evaluating, we get

$$\begin{aligned}
&-p_L c\left(\overline{\vartheta}\right) F\left(\overline{\vartheta}\right) + p_L \rho_0 F\left(\underline{\vartheta}\right) + \left(p_L\right)^2 R F\left(\overline{\vartheta}\right) - \left(p_L\right)^2 R F\left(\underline{\vartheta}\right) \\
&\quad = \rho_0\left(p_H F\left(\underline{\vartheta}\right) - \Delta p F\left(\underline{\vartheta}\right) + \Delta p F\left(\overline{\vartheta}\right)\right) - p_L R\left(p_H F\left(\underline{\vartheta}\right) - \Delta p F\left(\underline{\vartheta}\right) - \Delta p F\left(\overline{\vartheta}\right)\right),
\end{aligned}$$

giving

$$\begin{aligned}
&-p_L c\left(\overline{\vartheta}\right) F\left(\overline{\vartheta}\right) + p_L \rho_0 F\left(\underline{\vartheta}\right) + \left(p_L\right)^2 R F\left(\overline{\vartheta}\right) - \left(p_L\right)^2 R F\left(\underline{\vartheta}\right) \\
&\qquad = \rho_0 p_L F\left(\underline{\vartheta}\right) + \rho_0 \Delta p F\left(\overline{\vartheta}\right) - p_L R p_L F\left(\underline{\vartheta}\right) - \Delta p F\left(\overline{\vartheta}\right) p_L R.
\end{aligned}$$

All the terms with $F(\underline{\vartheta})$ can be eliminated from this equation. Then dividing by $p_L F\left(\overline{\vartheta}\right)$ and rearranging, we obtain

$$p_H R - \frac{\Delta p}{p_L}\rho_0 = c\left(\overline{\vartheta}\right).$$

With $\xi = 0$, equation (A.3) boils down to

$$\lambda = \frac{B}{\overline{\vartheta} - p_L R}.$$

Combining this with (A.8) gives

$$p_H R - \frac{\Delta p}{p_L}\rho_0 = \overline{\vartheta},$$

so that $\overline{\vartheta} = c\left(\overline{\vartheta}\right)$ and $\overline{\vartheta} = \vartheta^*$, which would also mean that $\vartheta^* = p_H R - \frac{\Delta p}{p_L}\rho_0$, which already leads to a contradiction for all sets of primitives for which $\vartheta^* \neq p_H R - \frac{\Delta p}{p_L}\rho_0$, so that when $\vartheta^* \neq p_H R - \frac{\Delta p}{p_L}\rho_0$, the solution must satisfy $\underline{\vartheta} = \overline{\vartheta}$. This gives that, adding equation (A.3) to (A.7) and dividing by I, we obtain

$$(\rho_1 - \rho_0) f\left(\overline{\vartheta}\right) = \lambda\left(\overline{\vartheta} - \rho_0\right) f\left(\overline{\vartheta}\right).$$

Furthermore, equation (A.1) simplifies to

$$(\rho_1 - \rho_0) F\left(\overline{\vartheta}\right) = \lambda\left(c\left(\overline{\vartheta}\right) - \rho_0\right) F\left(\overline{\vartheta}\right),$$

so that $\overline{\vartheta} = c\left(\overline{\vartheta}\right)$ and $\overline{\vartheta} = \vartheta^*$. This, together with the break-even constraint, gives that the optimal allocation is the one with

$$I = \frac{A}{F(\vartheta^*)(\vartheta^* - \rho_0)},$$

with $\overline{\vartheta} = \vartheta^*$ and with

$$D(\vartheta) = \left(R - \frac{B}{\Delta p}\right) I$$

for all ϑ.

Having established the results for all sets of primitives with $\vartheta^* \neq p_H R - \frac{\Delta p}{p_L}\rho_0$, we now turn our attention to the special case with $\vartheta^* = p_H R - \frac{\Delta p}{p_L}\rho_0$. QED

Proof of Proposition 2.3. . In the monopoly case, the lender wants to offer an allocation $\left(I, \overline{\vartheta}, D(\cdot)\right)$ that maximizes his profits under the firm's participation constraint. This gives rise to an optimization problem that can be solved in a manner entirely analogous to the one used in the proof of Proposition 2.2 and will thus also give a constant, a repayment that is always incentive compatible, and a liquidity level that minimizes the effective cost of investment. However, in this case, it is not the lender's break-even constraint that binds, but the firm's participation constraint, so that:

$$I = \frac{U^A}{\rho_1 - \rho_0},$$

$\overline{\vartheta} = \vartheta^{*}$, and

$$D(\vartheta) = \left(R - \frac{B}{\Delta p}\right) I$$

for all ϑ. QED

Proof of Proposition 2.4. . For this proposition, we assume an equilibrium concept akin to trembling-hand perfection. We assume that, for any offered contract other than the subgame perfect one, there is a small chance ε that the firm accepts this contract, where we let ε tend to zero. This precludes loss-giving contracts from being offered, even if they would not be accepted on the equilibrium path.

The proof proceeds by contradiction: assume that the equilibrium is characterized by the exclusive competition allocation $(I^C, \vartheta^{*}, D^C(\cdot))$, with $D^C(\vartheta) := D^C := \left(R - \frac{B}{\Delta p}\right) I^C$. An inactive lender, say i, can offer the contract $(J_i, L_i(\cdot), D_i(\cdot))$ with $J_i = 0$, with

$$L_i(m) = \begin{cases} m - \overline{\vartheta} & \text{if } \overline{\vartheta} < m \leq \overline{\vartheta} + \Delta\vartheta \\ 0 & \text{otherwise} \end{cases}$$

and with

$$D_i(m) = \begin{cases} \Delta D := RI - D^C & \text{if } \vartheta < m \leq \overline{\vartheta} + \Delta\vartheta \\ 0 & \text{otherwise} \end{cases}$$

for some $\Delta\vartheta$ small enough in the sense that

$$\mathbf{E}(\vartheta - \overline{\vartheta} \mid \overline{\vartheta} < \vartheta < \overline{\vartheta} + \Delta\vartheta) < p_L(RI - D).$$

Note that such $\Delta\vartheta$ exists since the distribution of ϑ is continuous and increasing. Accepting this offer on top of the set of contracts giving $(I^C, \vartheta^{*}, D^C(\cdot))$, the firm will be able to attract liquidity in case $\overline{\vartheta} < \vartheta < \overline{\vartheta} + \Delta\vartheta$, and then reap at least the private benefits. This raises the firm's ex ante expected utility by $\left(F(\overline{\vartheta} + \Delta\vartheta) - F(\overline{\vartheta})\right) BI$, and giving the entrant an expected profit equal to

$$\left(F(\overline{\vartheta} + \Delta\vartheta) - F(\overline{\vartheta})\right)\left(p_L(RI^C - D^C) - \mathbf{E}(\vartheta \mid \overline{\vartheta} < \vartheta < \overline{\vartheta} + \overline{\vartheta}_{\tilde{N}+1})\right)$$

This constitutes a profitable deviation for the entrant. This gives that the allocation $(I^C, \vartheta^{*}, D^C(\cdot))$ cannot be sustained in equilibrium.

Now it remains to be proven that this type of deviation cannot be prevented by the existence of so-called *latent contracts* that are not traded, but are offered. Assume by contradiction that there is a set of contracts that together would constitute an allocation $(I', \bar{\vartheta}', D\prime(\cdot))$. For any allocation $(I, \bar{\vartheta}, D(\cdot))$, denote the firm's utility by $U(I, \bar{\vartheta}, D(\cdot))$ and the lenders' aggregate profits by $\Pi(I, \bar{\vartheta}, D(\cdot))$. Note that if $D'(\vartheta') + \Delta D < RI'$, deviations such as the above would not be a problem, so we can assume that $D'(\vartheta') + \Delta D \vec{\geq} RI'$. The type of deviation above is not a profitable strategy only if this set of contracts is chosen after the above deviation for any $\Delta\vartheta$, but not before the deviation. For this to be the case one must have on the one hand that

$$U(I', \bar{\vartheta}', D\prime(\cdot)) \leq U(I^C, \vartheta^{*}, D^C(\cdot)),$$

but on the other hand that for all $\Delta\vartheta$

$$\begin{aligned} &U(I', \bar{\vartheta}', D\prime(\cdot)) + (F(\vartheta' + \Delta\vartheta) - F(\vartheta'))\, BI' \\ &\quad > U(I^C, \vartheta^*, D^C(\cdot))U(I', \bar{\vartheta}', D\prime(\cdot)) + (F(\vartheta^* + \Delta\vartheta) - F(\vartheta^*))\, BI^C. \end{aligned}$$

This means that

$$U(I', \bar{\vartheta}', D\prime(\cdot)) = U(I^C, \vartheta^*, D^C(\cdot)),$$

but $(I', \bar{\vartheta}',\ D\prime(\cdot))$ is different from $(I^C, \vartheta^*,\ D^C(\cdot))$. However, $\left(I^C, \vartheta^*, D^C(\cdot)\right)$ is the allocation that maximizes $U(I, \bar{\vartheta},\ D(\cdot))$ subject to $\Pi(I, \bar{\vartheta},\ D(\cdot)) \geq 0$. This entails that $\Pi(I', \bar{\vartheta}', D\prime(\cdot)) < 0$, meaning that at least one lender offers a contract that would give him a loss if accepted. If there is any positive probability that the firm mistakenly accepts this contract, this lender would suffer a loss. Therefore, this lender would be better off not offering any contract. Thus, this cannot be the case in any equilibrium satisfying our notion of trembling hand perfection and the deviation mentioned above remains a profitable strategy. QED

Proof of Proposition 2.5. . Assume by contradiction that the equilibrium is characterized by the allocation $(I, \overline{\vartheta},\ D(\cdot))$, with $\overline{\vartheta} < \infty$. An entrant lender, say i, can offer the contract $(J_i, L_i(\cdot), D_i(\cdot))$ with $J_i = 0$, with

$$L_i(m) = \begin{cases} m - \overline{\vartheta} & \text{if } \overline{\vartheta} < m \leq \overline{\vartheta} + \Delta\vartheta \\ 0 & \text{otherwise} \end{cases}$$

and with

$$D_i(m) = \begin{cases} \overline{D} & \text{if } \vartheta < m \leq \overline{\vartheta} + \Delta\vartheta \\ 0 & \text{otherwise} \end{cases}$$

for some $\Delta\vartheta$ small enough and $\overline{D}$ large enough in the sense that

$$\mathbf{E}(\vartheta - \overline{\vartheta} \mid \overline{\vartheta} < \vartheta < \overline{\vartheta} + \Delta\vartheta) < p_L \frac{\overline{D}}{D\left(\overline{\vartheta}\right) + \overline{D}} RI.$$

Note that such $\Delta\vartheta$ exists since the distribution of ϑ is continuous and increasing. Accepting this offer on top of the offers by incumbent lenders, the firm will be able to attract liquidity in case $\overline{\vartheta} < \vartheta < \overline{\vartheta} + \Delta\vartheta$, and then reap at least the private benefits. This raises the firm's ex ante expected utility by $\left(F(\overline{\vartheta} + \Delta\vartheta) - F(\overline{\vartheta})\right) BI$, and gives the entrant an expected profit equal to

$$\left(F(\overline{\vartheta} + \Delta\vartheta) - F(\overline{\vartheta})\right) \left(p_L \frac{\overline{D}}{D\left(\overline{\vartheta}\right) + \overline{D}} RI - \mathbf{E}\left(\vartheta \mid \overline{\vartheta} < \vartheta < \overline{\vartheta} + \overline{\vartheta}_{\tilde{N}+1}\right)\right).$$

This constitutes a profitable deviation for the entrant. Thus, $(I, \overline{\vartheta}, D(\cdot))$ cannot be sustained in equilibrium.

A simple argument shows that any contracts that are offered, but not traded, would not change this result: if the deviation mentioned above would trigger any combination of these "latent" contracts to be traded, the aggregate of these contracts would still be susceptible to the same kind of deviation. QED

Proof of Proposition 2.6. . To see that no entrant has an incentive to offer a contract that the firm would accept on top of C_0, we first note that no entrant has an incentive to only increase the liquidity supply to the firm, as this supply is already unlimited. Because of the clause withholding liquidity insurance, the entrant, say i, needs to supply liquidity.

Denote by $C_i = \left(\tilde{J}_i, \tilde{L}_i(\cdot), \tilde{D}_i(\cdot)\right)$ the contract that the entrant offers, by $\tilde{I} := A + J_0 + \tilde{J}_i$ the newly proposed investment size and by $\tilde{\vartheta} := \frac{\max \tilde{L}_i(\cdot)}{\tilde{I}}$ the maximum liquidity shock that can be financed when the firm accepts C_0 and $\tilde{C}_i$. Because the entrant causes the repayment to be the full value of the potential revenue upon increasing the project size, low effort will always be induced. For any repayment D_i that the entrant wants, he can thus get at most $p_L R\tilde{I}\frac{D_i}{D_i+RI} < p_L R\tilde{I}$ and the firm gets at most $B\tilde{I}$. This means that the firm's utility upon accepting C_0 and $\tilde{C}_i$ equals $\tilde{U}_i = F\left(\tilde{\vartheta}\right)B\tilde{I}$. Now this should be greater than the utility from acccepting only C_0, so that we must have

$$F\left(\tilde{\vartheta}\right)B\tilde{I} \geq (\rho_1 - \rho_0)(A + J_0).$$

As $\rho_1 - \rho_0 = p_H\frac{B}{\Delta p}$, this then becomes the very simple

$$F\left(\tilde{\vartheta}\right)\tilde{I} \geq \frac{p_H}{\Delta p}(A + J_0). \tag{A.9}$$

As said, for all ϑ, the entrant's revenue is smaller than $p_L R\tilde{I}$. This means that the entrant's total profits are smaller than

$$F\left(\tilde{\vartheta}\right)p_L R\tilde{I} - \left(\left(1 + \int_0^{\tilde{\vartheta}} \vartheta f(\vartheta)\,\mathrm{d}\vartheta\right)\tilde{I} - (A + J_0)\right),$$

which equals

$$(A + J_0) - F\left(\tilde{\vartheta}\right)\left(c\left(\tilde{\vartheta}\right) - p_L R\right)\tilde{I}.$$

By the assumption that $\vartheta^* - p_L R \geq \frac{\Delta p}{p_H}$, this is smaller than $(A + J_0) - F\left(\tilde{\vartheta}\right)\frac{\Delta p}{p_H}\tilde{I}$. Combining this with condition (A.9), we find that the profits of the entrant are negative for any acceptable contract. Without unlimited liquidity supply entrants would have an incentive to deviate. Also, given the unlimited liquidity supply, this contract optimizes the firm's utility subject to the lender's break-even constraint (as can be shown by methods analogous to those in the proof of Proposition 2.2), so that no entrant has an incentive to deviate by undercutting Lender 0's offer. QED

Proof of Proposition 2.7. . Let lender 0 offer the contract from Proposition 2.6. Then no lender other than lender 0. has an incentive to deviate, as noted in Proposition 2.6. Lowering the liquidity supply leaves an incentive to deviate for the entrants, who can then extend the liquidity supply. Any contract that gives Lender 0 a higher profit while keeping the liquidity supply unlimited will give entrants an incentive to undercut. Thus Lender 0 also has no incentive to deviate. QED

Proof of Proposition 2.8. . This is a direct consequence of Propositions 2.5 and 2.7. First of all, by Proposition 2.7 an equilibrium exists characterized by the mentioned allocation. Then, by Proposition 2.5, any equilibrium must satisfy $\overline{\vartheta} = \infty$. Furthermore, the allocation must optimize the firm's utility given the restriction of unlimited liquidity support and given the lenders' break-even constraint. Again, it can be proven, analogously to the proof of Proposition 2.2, that this optimal allocation is the one with

$$D(\vartheta) = \left(R - \frac{B}{\Delta p}\right) I$$

and

$$I = \frac{A}{1 + \mathbf{E}\vartheta - \rho_0}.$$

QED

Appendix B

Chapter 2 with a Finite Number of Principals

In this section, we assume that there are $N < \infty$ lenders, indexed $i = 1, 2, \ldots, N$, simultaneously offering contracts. We derive results for a subgame perfect equilibrium. First, we establish the equivalent to Proposition 2.4.

Lemma B.1. *The exclusive competition allocation* $(I^C, \vartheta^*, D^C(\cdot))$*, with* $D^C(\vartheta) := D^C := \left(R - \frac{B}{\Delta p}\right) I^C$ *cannot be sustained in equilibrium.*

Proof. For this proposition, we assume the same equilibrium concept as in Proposition 2.4. The proof proceeds by contradiction: assume that the equilibrium is characterized by the exclusive competition allocation $(I^C, \vartheta^*, D^C(\cdot))$, with $D^C(\vartheta) := D^C := \left(R - \frac{B}{\Delta p}\right) I^C$.

First, if there is an inactive lender, then this lender has an incentive to deviate by offering the type of contract mentioned in the proof of Lemma 2.4. If there is no inactive lender, consider a lender i with

$$i \in \arg\min_j \left\{ \max_m L_j(m) \right\},$$

i.e., the lender who provides the smallest amount of liquidity to the firm at the firm's highest liquidity need. Then

$$L_i\left(m_i\left(\overline{\vartheta}\right)\right) \leq \frac{\overline{\vartheta I}}{N} < p_L\left(RI - D^C\right).$$

Now lender i can change the contract to $\left(\tilde{J}_i, \tilde{L}_i(\cdot), \tilde{D}_i(\cdot)\right)$with $\tilde{J}_i = J_i$, with

$$\tilde{L}_i(m) = \begin{cases} L_i(m) \text{ if } 0 \leq m \leq \overline{\vartheta} \\ m - \overline{\vartheta} \text{ if } \overline{\vartheta} < m \leq \overline{\vartheta} + \Delta\vartheta \end{cases}$$

for some small enough $\Delta\vartheta$, and with

$$\tilde{D}_i(m) = \begin{cases} D_i(m) \text{ if } 0 \leq m \leq \overline{\vartheta} \\ D_i\left(m_i\left(\overline{\vartheta}\right)\right) + RI - D^C \text{ if } \overline{\vartheta} < m \leq \overline{\vartheta} + \Delta\vartheta \end{cases}$$

Accepting this contract increases the firm's utility by $\left(F\left(\overline{\vartheta}+\Delta\vartheta\right)-F\left(\overline{\vartheta}\right)\right)BI$ and lender i's utility by

$$\left(F\left(\overline{\vartheta}+\Delta\vartheta\right)-F\left(\overline{\vartheta}\right)\right)\left(p_L\left(RI-D^C\right)\right)-L_i\left(m_i\left(\overline{\vartheta}\right)\right)-\mathbf{E}\left(\vartheta-\overline{\vartheta}|\overline{\vartheta}<\vartheta\leq\overline{\vartheta}+\Delta\vartheta\right)\right),$$

which is positive for $\Delta\vartheta$ small enough. This means that lender i has an incentive to deviate. QED

The intuition behind the proof is simple. If there is an inactive lender in equilibrium, he will want to behave just like the entrant in Proposition 2.4. If all lenders are active, and if the number of lenders is large enough, at least one lender will have such a small stake in the firm that he will effectively want to act as an entrant. A direct corollary is the following:

Corollary B.2. *If $N>\vartheta^*\frac{\Delta p}{p_L B}$, none of the benchmark allocations can be sustained in equilibrium.*

Proof. If $N>\vartheta^*\frac{\Delta p}{p_L B}$, then both benchmark allocations satisfy $\frac{\overline{\vartheta I}}{N}<p_L\left(RI-D(\overline{\vartheta})\right)$, so that they cannot be sustained in equilibrium. QED

Now we will establish the equivalents to the results in Section 2.5.2. We also assume that the investment size I is observable at both $t=1$ and $t=2$, so that the lender can penalize the firm for seeking additional up-front investment. The size is used in the contract in the exact same way as it is in the free-entry model. However, throughout the proofs, we suppress the explicit size clauses in the contract for expositional purposes. Again we assume that strategic default is possible and division is pro rata. This gives us the following result, equivalent to Proposition 2.5:

Proposition B.3. *No equilibrium with $\frac{\overline{\vartheta}}{N}<p_L R$ can exist.*

Proof. Assume there is an equilibrium with $\frac{\overline{\vartheta I}}{N}<p_L RI$. First, if there is an inactive lender, then this lender has an incentive to deviate by offering the type of contract mentioned in the proof of Lemma 2.4. If there is no inactive lender, consider a lender i with

$$i\in\arg\min_j\left\{\max_m L_j(m)\right\},$$

i.e., the lender who provides the smallest liquidity to the firm at the firm's highest liquidity need. Then

$$L_i\left(m_i\left(\overline{\vartheta}\right)\right)\leq\frac{\overline{\vartheta I}}{N}<p_L RI.$$

Now lender i can change its contract to $\left(\tilde{J}_i,\tilde{L}_i\left(\cdot\right),\tilde{D}_i\left(\cdot\right)\right)$with $\tilde{J}_i=J_i$, with

$$\tilde{L}_i(m)=\begin{cases}L_i\left(m\right)\text{ if }0\leq m\leq\overline{\vartheta}\\ m-\overline{\vartheta}\text{ if }\overline{\vartheta}<m\leq\overline{\vartheta}+\Delta\vartheta\end{cases}$$

for some small enough $\Delta\vartheta$, and with

$$\tilde{D}_i(m)=\begin{cases}D_i\left(m\right)\text{ if }0\leq m\leq\overline{\vartheta}\\ \tilde{D}\text{ if }\overline{\vartheta}<m\leq\overline{\vartheta}+\Delta\vartheta.\end{cases}$$

Accepting this contract increases the firm's utility by $\left(F\left(\overline{\vartheta}+\Delta\vartheta\right)-F\left(\overline{\vartheta}\right)\right)BI$ and lender i's utility by

$$\left(F\left(\overline{\vartheta}+\Delta\vartheta\right)-F\left(\overline{\vartheta}\right)\right)\left(p_L\frac{\tilde{D}}{D(\overline{\vartheta})}RI-L_i\left(m_i\left(\overline{\vartheta}\right)\right)-\mathbf{E}\left(\vartheta-\overline{\vartheta}|\overline{\vartheta}<\vartheta\leq\overline{\vartheta}+\Delta\vartheta\right)\right),$$

which is positive for $\Delta\vartheta$ small enough and $\tilde{D}$ large enough. This means that lender i has an incentive to deviate. QED

Again, the intuition is that, if at least one lender is inactive, he will behave like an entrant. If all lenders are active, one of them is small enough to want to behave like an entrant. This proposition gives a necessary condition for the maximum liquidity shock that has to be insured in equilibrium, namely that

$$\overline{\vartheta}\geq Np_LR.$$

Assuming that the set of possible levels of the liquidity shock Θ is bounded, one has the following corollary:

Corollary B.4. *If $N>\frac{\sup\Theta}{p_LR}$, then no allocation with $\overline{\vartheta}<\sup\Theta$ can be sustained in equilibrium.*

Proof. If $N>\frac{\sup\Theta}{p_LR}$ and $\overline{\vartheta}<\sup\Theta$, then $\frac{\overline{\vartheta}}{N}<p_LR$, so that an allocation with $\overline{\vartheta}<\sup\Theta$ cannot be sustained. QED

This corollary says that, when N is large enough with respect to the potential liquidity shock, then liquidity support in equilibrium is essentially unbounded. The intuition is that, if the largest potential liquidity shock is not too large, and the number of investors is large enough, then even at the largest liquidity shock, one investor will be contributing so little to the firm's liquidity insurance that he will want to behave as an entrant.

Establishing the existence of an equilibrium is technically more involved in the simultaneous-move game than it is under free entry, but the intuitions remain the same. In any equilibrium, the threat of entry must now be explicitly modeled by latent contracts offered by inactive lenders. We can state the following existence result.

Proposition B.5. *Assume that $N\geq 4$, that*

$$\frac{\left(\rho_1-\rho_0\right)A}{1+\mathbf{E}\vartheta-\rho_0}\geq U_A,$$

and that for all $\tilde{\vartheta}<\sup\Theta$, one has

$$\rho_0>\mathbf{E}\left(\vartheta-\tilde{\vartheta}|\vartheta\geq\tilde{\vartheta}\right);$$

then, there exists an equilibrium with the lenders breaking even, with $\overline{\vartheta}=\sup\Theta$, and with a flat, incentive-compatible repayment.

Proof. First of all, two lenders, say lenders 3 and 4, offer a contract that serves the same function as the "threat"of the entrant expanding the liquidity supply: for $i \in \{3,4\} : (J_i, L_i(\cdot), D_i(\cdot))$, with $J_i = 0$, with $L_i(m) = m$ and with

$$D_i(m) = \begin{cases} 0 \text{ if } m = 0 \\ \overline{D} \text{ if } m > 0 \end{cases}$$

for some large enough $\overline{D}$.

With this contract on offer, whenever all lenders, except for 3 and 4, together offer an acceptable allocation with $\overline{\vartheta} < \sup \Theta$, the firm accepts contract 3 on top of what the other lenders offer, as doing so increases the firm's utility by $\left(1 - F\left(\overline{\vartheta}\right)\right) BI$. Two other lenders, say lenders 1 and 2, both offer the contract $(J_i, L_i(\cdot), D_i(\cdot))$, for $i \in \{1,2\}$, with

$$J_i = \left(\frac{1}{1 + \mathbf{E}\vartheta - \rho_0} - 1\right) A,$$

with $L_i(m) = m$ for all m, i.e. with an unlimited supply of liquidity, and with $D_i = \left(R - \frac{B}{\Delta p}\right)(A + J_i)$, i.e., with a precisely incentive-compatible repayment.

All other lenders just offer the null contract. With these three contracts on offer, the firm chooses either C_1 or C_2, upon which no lender makes a profit, and the firm earns a total consumption utility of

$$\frac{(\rho_1 - \rho_0) A}{1 + \mathbf{E}\vartheta - \rho_0},$$

which, by assumption, is larger than the firm's reservation utility. To see that no lender has an incentive to deviate, first think of a deviation in which a lender, say i, offers a contract with a finite maximum supply of liquidity. Proving that the best contract of this type asks for a flat, incentive-compatible repayment is done in a manner entirely analogous to the proofs of the same result from the section with the benchmark cases, so we restrict our attention to deviating contracts of the form $\tilde{C}_i = \left(\tilde{J}_i, \tilde{L}_i(\cdot), \tilde{D}_i(\cdot)\right)$, with

$$\tilde{L}_i(m) = \begin{cases} m \text{ if } 0 \leq m \leq \overline{L}_i \\ 0 \text{ if } m > \overline{L}_i \end{cases}$$

and $\tilde{D}_i(m) = \tilde{D} := \left(R - \frac{B}{\Delta p}\right)\left(A + \tilde{J}_i\right)$. Now call $\tilde{I}_i := A + \tilde{J}_i$ and $\overline{\vartheta}_i := \frac{\overline{L}_i}{\tilde{I}_i}$. Then, if the firm accepts the contract $\tilde{C}_i$, it will also accept contract 3 or 4. Upon this, the firm will have a utility of

$$F\left(\overline{\vartheta}_i\right)(\rho_1 - \rho_0)\tilde{I}_i + \left(1 - F\left(\overline{\vartheta}_i\right)\right) B\tilde{I}_i,$$

which can be rewritten as

$$\left(F\left(\overline{\vartheta}_i\right) + \left(1 - F\left(\overline{\vartheta}_i\right)\right)\Delta p\right)(\rho_1 - \rho_0)\tilde{I}_i$$

and lender i will receive a gross revenue of

$$F\left(\overline{\vartheta}_i\right)\rho_0 \tilde{I}_i + \left(1 - F\left(\overline{\vartheta}_i\right)\right) p_L \frac{\tilde{D}}{\tilde{D} + \overline{D}} R\tilde{I}_i$$

with an expected cost of investing of

$$\left(1 + \int_0^{\overline{\vartheta}_i} \vartheta f(\vartheta)\,\mathrm{d}\vartheta + \left(1 - F\left(\overline{\vartheta}_i\right)\right)\overline{\vartheta}_i\right)\tilde{I}_i - A.$$

Combining, we find the following expression for the lender's net profit:

$$A + \left(\rho_0 - (1 + \mathbf{E}\vartheta) - \left(1 - F\left(\overline{\vartheta}_i\right)\right)\left(\rho_0 - \mathbf{E}\left(\vartheta - \tilde{\vartheta}|\vartheta \geq \tilde{\vartheta}\right) + p_L \frac{\tilde{D}}{\tilde{D} + \overline{D}} R\tilde{I}_i\right)\right)\tilde{I}_i.$$

In order for the firm to accept contract i, it must be better off than before, i.e.,

$$\left(F\left(\overline{\vartheta}_i\right) + \left(1 - F\left(\overline{\vartheta}_i\right)\right)\Delta p\right)(\rho_1 - \rho_0)\,\tilde{I}_i \geq \frac{(\rho_1 - \rho_0)\,A}{1 + \mathbf{E}\vartheta - \rho_0},$$

giving that

$$\tilde{I}_i \geq \frac{A}{\left(F\left(\overline{\vartheta}_i\right) + \left(1 - F\left(\overline{\vartheta}_i\right)\right)\Delta p\right)(1 + \mathbf{E}\vartheta - \rho_0)},$$

and lender i's profits are

$$\begin{aligned} A &+ \left(\rho_0 - (1 + \mathbf{E}\vartheta) - \left(1 - F\left(\overline{\vartheta}_i\right)\right)\left(\rho_0 - \mathbf{E}\left(\vartheta - \tilde{\vartheta}|\vartheta \geq \tilde{\vartheta}\right) + p_L \frac{\tilde{D}}{\tilde{D} + \overline{D}} R\tilde{I}_i\right)\right)\tilde{I}_i \\ &\leq A - \frac{\left(1 + \mathbf{E}\vartheta - \rho_0 + \left(1 - F\left(\overline{\vartheta}_i\right)\right)\left(\mathbf{E}\left(\vartheta - \tilde{\vartheta}|\vartheta \geq \tilde{\vartheta}\right) - \rho_0 - p_L \frac{\tilde{D}}{\tilde{D} + \overline{D}} R\tilde{I}_i\right)\right)}{(1 + \mathbf{E}\vartheta - \rho_0)\left(F\left(\overline{\vartheta}_i\right) + \left(1 - F\left(\overline{\vartheta}_i\right)\right)\Delta p\right)} A. \end{aligned}$$

Because $\overline{D}$ was assumed to be large enough, the right hand side of this equality is negative, as $\rho_0 > \mathbf{E}\left(\vartheta - \tilde{\vartheta}|\vartheta \geq \tilde{\vartheta}\right)$ and $F\left(\overline{\vartheta}_i\right) < 1$, meaning that a deviating lender would always make a loss with a deviation that is acceptable for the firm. This means that there is no possible deviation involving a lender offering a contract with only a limited supply of liquidity. To see that no lender would deviate with a contract with an unlimited supply of liquidity, one needs to observe that the contracts offered by lenders 1 and 2 are the ones offering the highest utility to the firm, given the restrictions that lenders must break even and that the liquidity supply must be unlimited. This means that there are no profitable deviations for any of the lenders other than 1 and 2. Also, any deviation by lender 1 would make the firm accept the contract by lender 2 and vice versa. The proof that these contracts are indeed the optimal ones is exactly analogous to the proofs in the benchmark cases. QED

To get an idea of when the technical prerequisites of the above proposition are satisfied, we state the following corollary:

Corollary B.6. *Let ϑ follow an exponential distribution with parameter λ; then, if $N \geq 4$, if $\rho_0 > \frac{1}{\lambda}$, and if $\frac{\rho_1 - \rho_0}{1 + \frac{1}{\lambda} - \rho_0} \geq \frac{U_A}{A}$, there exists an equilibrium with trade and an unlimited supply of liquidity.*

Proof. For the exponential distribution, we have $\mathbf{E}\left(\vartheta - \tilde{\vartheta}|\vartheta \geq \tilde{\vartheta}\right) = \mathbf{E}\vartheta = \frac{1}{\lambda}$. Filling this in gives the conditions for Proposition B.5. QED

Now we see that the results from the free-entry analysis in the main paper carry over to a more formal game with simultaneous moves and a finite number of players. By proposition B.3, we have found a lower bound on the total supply of liquidity in any equilibrium. If the support of the liquidity shock is bounded, this translates into a basically unlimited supply of liquidity by Corollary B.4. Furthermore, with proposition B.5, we have found sufficient conditions for the existence of an equilibrium with unlimited liquidity provision.

Appendix C

Proofs for Chapter 3

Proof of Proposition 3.1. Assume that any allocation leaves the borrower with a surplus $U < \frac{3}{4}$. In that case, an inactive investor i can offer the contract with $J_i = 1$, with $\overline{L}_i = 2$ and with $D_i(2) = 4 + \varepsilon$, where $0 < \varepsilon < 1 + B - 2U$. In that case, when accepting the contract, the borrower will have a utility of

$$\frac{1}{2}\left(R + B - D_i(2)\right) = \frac{1}{2}(1 + B - \varepsilon) > U,$$

so the borrower will accept this contract. For the investor, the resulting utility will be

$$-1 + \frac{1}{2}(D_i(2) - 2) = \frac{1}{2}\varepsilon > 0,$$

so this is a profitable deviation for the investor. Thus any equilibrium must provide the maximum surplus of $\frac{3}{4}$ to the borrower. This also entails that $\overline{\vartheta} = 2$ and lenders must break even. An example equilibrium is one in which at least two investors offer the contract $C^* = \left(J^*, \overline{L}^*, D^*(\cdot)\right)$, with $J^* = 1$, $\overline{L}^* = 2$, and $D^* \equiv 4$. The borrower then takes one of these contracts. QED

Proof of Lemma 3.2. Assume an equilibrium with $\overline{\vartheta} = 2$. Now, there is at least one investor who supplies less than $\frac{2}{M}$ when the low liquidity shock realizes. Assume this investor is active. Call this investor i, and the amount this investor supplies when the low liquidity shock hits L_i^L. This investor can now modify his contract to one $\left(\tilde{J}, \tilde{L}, \tilde{D}(\cdot)\right)$, with $\tilde{J}_i = J_i$, with $\overline{L}_i = 4 + L_i^L$, with $\tilde{D}_i(L) = D_i(L_i^L)$ for $L \leq L_i^L$, and with $\tilde{D}_i(L) = \overline{D}$ for $L \leq L_i^L$, with $\overline{D}$ large enough. In case this investor is inactive, the same works with $\tilde{J}_i = 0$, $\overline{L}_i = 4$ and $\tilde{D}_i(L) = \overline{D}\mathbf{1}_{\{L>0\}}$.

If the borrower accepts this contract, on top of all the other ones that were traded in our conjectured equilibrium, nothing changes in case the low liquidity shock realizes. However, if the high liquidity shock realizes, it can still withdraw a total of $2 - L_i^L$ from all active investors except for i, and $4 + \overline{L}_i^L$ from investor i. This will give the borrower at least the private benefit in case the high shock realizes, so that the borrower is better off ex ante accepting this contract. For investor i, nothing changes in case the low shock hits, but if the high shock hits, he will receive $5\frac{\overline{D}}{\overline{D}+D(2)} - (4 + \overline{L}_i^L)$, which is greater than zero, because $\overline{D}$ is large enough and $\overline{L}_i^L < \frac{2}{N}$ and N was assumed to be large enough. This means that there is a profitable deviation and the assumed equilibrium can not be sustained. QED

Proof of Lemma 3.3. Denote by $\mathcal{I}$ the set of investors with which the borrower trades.

Assume $\overline{\vartheta} = 6$. In that case, at least two investors must be active in equilibrium. If in any state of the world, the total due repayment to investors were to be greater than 5, each active investor would have an incentive to increase his share of the revenue by further diluting the other investors. Therefore, this proof focuses on allocations where in each state of the world ϑ, $\sum_{i \in \mathcal{I}} D_i^{\vartheta} \leq R$.

The expected total cost of financing the borrower in equilibrium is

$$1 + \frac{1}{2}\vartheta_L + \frac{1}{2}\vartheta_H = 5.$$

This means that because the investors need to break even:

$$\frac{1}{2}\sum_{i \in \mathcal{I}} D_i^L + \frac{1}{2}\sum_{i \in \mathcal{I}} D_i^H = 5$$

This means that for both states of the world, the total repayment has to be

$$\sum_{i \in \mathcal{I}} D_i\left(L_i(\vartheta)\right) = 5.$$

Now first of all, the repayments to each of the investors must be constant over the two different states. Otherwise, there must be one investor i such that $D_i(L_i(2)) < D_i(L_i(6))$, giving also that $D_{-i}(L_{-i}(2)) > D_{-i}(L_{-i}(6))$, giving that $L_{-i}(2) > L_{-i}(6)$. This means that $L_{-i}(6) < 2$, so that $L_i(6) > 4$. Also, $L_i(6) \leq 5$, so $L_{-i}(6) \geq 1$.

Assume that $D_{-i}\left(L_{-i}(2)\right) - D_{-i}\left(L_{-i}(6)\right) < L_{-i}(2) - L_{-i}(6)$, i.e. the combined provision of lenders other than i sells the incremental units of liquidity between $L_{-i}(6)$ and $L_{-i}(2)$ at a discount. Lender i, could take advantage of this by offering a contract with C_i' with $J_i' = J_i$ $\overline{L}_i' = 6 - L_{-i}(2)$ and

$$D_i'(L) = \begin{cases} D_i\left(L_i(2)\right) & \text{if } L \leq L_i(2) \\ D_i\left(L_i(6)\right) - \left(D_{-i}\left(L_{-i}(2)\right) - D_{-i}\left(L_{-i}(6)\right)\right) & \text{if } L > L_i(2) \end{cases}$$

In that case, the borrower would be better off choosing the contracts C_i' and all the contracts in $\mathcal{I} \setminus \{i\}$. In case $\vartheta = 2$, this would not change anything, and in case $\vartheta = 6$, the firm can get $\overline{L}_i'$ from lender i and $L_{-i}(2)$ from the others at a total price of

$$D_i\left(L_i(6)\right) - \left(D_{-i}\left(L_{-i}(2)\right) - D_{-i}\left(L_{-i}(6)\right)\right) + D_{-i}\left(L_{-i}(2)\right) = D(6)$$

so that she has the same repayment in case the high shock hits. For lender i this will raise his utility in case the high shock hits by

$$L_{-i}(2) - L_{-i}(6) - \left(D_{-i}\left(L_{-i}(2)\right) - D_{-i}\left(L_{-i}(6)\right)\right) > 0,$$

so this constitutes a profitable deviation.

Assume that $D_{-i}\left(L_{-i}(2)\right) - D_{-i}\left(L_{-i}(6)\right) > L_{-i}(2) - L_{-i}(6)$, then all lenders except for i sell the incremental liquidity from $L_{-i}(6)$ and $L_{-i}(2)$ at a mark-up. This means lender i can "undercut" them for this incremental provision. He could do so by lowering his price for obtaining $2 - L_{-i}(6)$ to

$$D_i' := D_i\left(L_i(2)\right) + \left(D_{-i}\left(L_{-i}(2)\right) - D_{-i}\left(L_{-i}(6)\right)\right) - \varepsilon.$$

In this case, when $\vartheta = 2$ the borrower is better off getting $L_{-i}(6)$ from the lenders other than i, and $2 - L_{-i}(6)$ from lender i. This would mean an improvement for the borrower, and lender i will now, in case $\vartheta = 2$, instead of $D_i\left(L_i(2)\right) - L_i(2)$, obtain

$$\begin{aligned}
&D_i'\left(2 - L_{-i}(6)\right) \\
&\quad = D_i\left(L_i(2)\right) + \left(D_{-i}\left(L_{-i}(2)\right) - D_{-i}\left(L_{-i}(6)\right)\right) - \varepsilon - \left(2 - L_{-i}(6)\right) \\
&\quad = D_i\left(L_i(2)\right) + \left(D_{-i}\left(L_{-i}(2)\right) - D_{-i}\left(L_{-i}(6)\right)\right) - \varepsilon - \left(L_i(2) + L_{-i}(2) - L_{-i}(6)\right) \\
&\quad = D_i\left(L_i(2)\right) - L_i(2) - \varepsilon + \left(\left(D_{-i}\left(L_{-i}(2)\right) - D_{-i}\left(L_{-i}(6)\right)\right) - \left(L_{-i}(2) - L_{-i}(6)\right)\right),
\end{aligned}$$

which, as long as ε is small enough, is greater than $D_i\left(L_i(2)\right) - L_i(2)$. This constitutes a profitable deviation for lender i.

Now the case is left with $D_{-i}\left(L_{-i}(2)\right) - D_{-i}\left(L_{-i}(6)\right) = L_{-i}(2) - L_{-i}(6)$. In that case also $D_i\left(L_i(6)\right) - D_i\left(L_i(2)\right) = L_{-i}(2) - L_{-i}(6)$. In that case, a lender $j \neq i$ could offer the contract C_j' with $J_j' = J_{-i}$, with $\overline{L}_j' = L_{-i}(6)$ and with

$$D_i'(L) = \begin{cases} D_{-i}\left(L_{-i}(6)\right) - \varepsilon & \text{if } L = 0 \\ D_{-i}\left(L_{-i}(6)\right) & \text{if } L > 0 \end{cases}$$

If the borrower chooses contracts j and i, she will choose to obtain all liquidity from lender i in case the low shock hits, paying

$$\begin{aligned}
D_{-i}\left(L_{-i}(6)\right) - \varepsilon + D_i(2) &\leq D_{-i}\left(L_{-i}(6)\right) + D_i\left(L_i(6)\right) - \varepsilon \\
&= D(6) - \varepsilon \\
&< D(2).
\end{aligned}$$

The borrower will make the same aggregate profit as all lenders in $\mathcal{I} \setminus \{i\}$ in case the high shock hits, but in case the low shock hits, he will obtain

$$\begin{aligned}
D_{-i}\left(L_{-i}(6)\right) - \varepsilon &= D_{-i}\left(L_{-i}(2)\right) - \left(D_{-i}\left(L_{-i}(2)\right) - D_{-i}\left(L_{-i}(6)\right)\right) - \varepsilon \\
&= D_{-i}\left(L_{-i}(2)\right) - \left(L_{-i}(2) - L_{-i}(6)\right) - \varepsilon,
\end{aligned}$$

which, for ε small enough, is greater than $D_{-i}\left(L_{-i}(2)\right) - L_{-i}(2)$, so that lender j makes a profit which is strictly greater than the aggregate profit made by all active lenders except i. Now, this is profitable for lender j, as if $j \in \mathcal{I}$, then j's profit can not be higher than that of all the lenders in $\mathcal{I} \setminus \{i\}$ combined, and if $j \notin \mathcal{I}$, he will go from zero profits to a nonzero profit.

Having established that the repayments are constant over states of the world, meaning that each lender provides liquidity at a marginal price of zero, it is possible to show that there always exists a profitable deviation for at least one of the lenders.

This means that there exists an investor i that supplies a positive amount L_i^L in case $\vartheta = 2$. In that case, this investor could instead offer the contract $\tilde{C}$ with

$$\tilde{L}_i^L = \max\left\{0, 2 - \sum_{j \in \mathcal{I}\setminus\{i\}} L_j^H\right\}$$

and $\tilde{D}_i^L = D_i^L - \varepsilon$, and with $\tilde{L}_i^H = L_i^H$ and $\tilde{D}_i^H = D_i^H$. The borrower would then prefer getting as much of its liquidity as possible from the other investors in the low state. This will save the borrower $\frac{1}{2}\varepsilon$, and will save investor i an amount of

$$\frac{1}{2}\min\left\{L_i^L, L_i^L + \sum_{j \in \mathcal{I}\setminus\{i\}} L_j^H - 2\right\} > 0$$

ex ante. Thus, there exists an incentive to deviate for investor i. QED

Proof of Proposition 3.4. This is a direct consequence of lemmas 3.2 and 3.3: in any allocation with $\overline{\vartheta} = 2$, there are incentives to deviate, as there are in any allocation with $\overline{\vartheta} = 6$. QED

Proof of Proposition 3.5. Consider the following set of strategies: at least two candidate intermediaries (call the set of these intermediaries $\mathcal{B}$), offer a contract to two of the other investors (call the set of these investors $\mathcal{D}$), asking for one unit of money from each of them, for which he will return precisely one unit of money at $t = 2$. The intermediaries offer the borrower the contract with $J_1 = 1$, with $\overline{L}_1 = 6$ and with $D(L) = 5$ for all L. At least two other investors than the intermediary (call the set of these investors $\mathcal{L}$) offer the contract which offers no initial investment, offers $\overline{L}_i = 4$ and with

$$D_i(L) = \begin{cases} 0 & \text{if } L = 0 \\ D & \text{if } L > 0 \end{cases}$$

for some very large D. Note that the sets $\mathcal{D}$ and $\mathcal{L}$ are not necessarily disjoint, one could even have them be exactly the same. I only assume $\mathcal{B} \cap \mathcal{D}$ and $\mathcal{B} \cap \mathcal{L}$ are empty. With these contracts on offer, the borrower chooses to do business with investor 1 and both investors in $\mathcal{D}$ deposit their money.

Any more attractive contract for the borrower should have $\overline{L} = 2$, but if any one of the investors deviates by offering such a contract, the borrower would always take one of the contracts from $\mathcal{L}$ as well. Assume a borrower i deviates by offering a contract with $J_i = 1$ and $\overline{L}_i = 2$. The borrower would then obtain a utility of

$$B + \frac{1}{2}\left(R - D_i(2)\right)_+$$

and investor i would have a utility of

$$-1 + \frac{1}{2}\left(D_i(2) - 2\right) + \frac{1}{2}\left(\frac{D_i(2)}{D_i(2) + D}R - 2\right).$$

In order to have the borrower prefer this contract, it must have $D_i(2) < R$, so that if D is large enough, the deviating investor's utility is negative. QED

Proof of Proposition 3.6. This proof is largely analogous to the one of lemma 3.2 and proceeds by contradiction. Assume $\overline{\vartheta} < \sup\Theta$. This means that there is a value $\Delta\vartheta$ such that $F\left(\overline{\vartheta} + \Delta\vartheta\right) - F\left(\overline{\vartheta}\right) > 0$ and $0 < \Delta\vartheta < R$.

If there is an inactive lender, say lender i, he can offer a contract $\left(J_i, \overline{L}_i, D_i(\cdot)\right)$ with $J_i = 0$, with $\overline{L}_i = \Delta\Theta$ and with

$$D_i(L) = \begin{cases} 0 & \text{if } L = 0 \\ D & \text{if } L > 0 \end{cases}$$

for some large enough D. QED

Proof of Lemma 3.7. Assume that $\overline{\vartheta} = \vartheta^H$. Now, first of all note that since $W \geq \frac{2}{3}\vartheta^H$, it is always possible to divide the active lenders into two disjoint subsets $\mathcal{I}^1$ and $\mathcal{I}^2$ with $\mathcal{I}^1 \cup \mathcal{I}^2 = \mathcal{I}$, such that for both $k \in \{1, 2\}$

$$\sum_{i \in \mathcal{I}^k} L_i\left(\vartheta^H\right) \leq W.$$

This justifies considering the problem as if it were one with only two lenders. Below, I will consider equilibria with only two lenders, but every time I mention a deviating strategy by lender $k \in \{1, 2\}$, one can read that "a lender in $\mathcal{I}^k$" executes this deviating strategy. If I say that an inactive lender can play a strategy to undercut lender k, one can also read that, if there are no inactive lenders, a lender in $\mathcal{I}^k$ can execute this strategy. Thus, I will focus on the case $\mathcal{I} = \{1, 2\}$ and use the index $-i$ to indicate the quantities and contracts for the lender other than i.[1]

Note that in the case of a dichotomous ϑ, the assumptions of the model give that

$$I + \vartheta_L < I + p_L\vartheta^L + p_H\vartheta^H \leq R < R + B < \vartheta^H.$$

Furthermore, $R > \vartheta^H$, and $\vartheta_H > W\vartheta_H - \vartheta_L$. The proof is largely analogous to the one of lemma 3.3.

Furthermore, the aggregate break-even condition gives that

$$D\left(\vartheta^H\right) - p_L\left(D\left(\vartheta^H\right) - D\left(\vartheta^L\right)\right) \geq 1 + \mathbf{E}\vartheta,$$

so that, because $D\left(\vartheta^H\right) \leq R$,

$$D\left(\vartheta^H\right) - D\left(\vartheta^L\right) \geq \frac{R - (1 + \mathbf{E}\vartheta)}{p_L},$$

so that condition 3.6 gives that

$$D\left(\vartheta^H\right) - D\left(\vartheta^L\right) \geq \vartheta^H - W.$$

As a first result, I need to prove that for any i, $L_{-i}\left(\vartheta^H\right) > L_{-i}\left(\vartheta^L\right)$. In order to do so, I will, by contradiction, assume that $L_{-i}\left(\vartheta^H\right) \leq L_{-i}\left(\vartheta^L\right)$, and distinguish the cases in which

[1] i.e. quite pedantically, lender $3 - i$.

1. $D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right)-D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right)>L_{-i}\left(\vartheta^L\right)-L_{-i}\left(\vartheta^H\right)$,
2. $D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right)-D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right)<L_{-i}\left(\vartheta^L\right)-L_{-i}\left(\vartheta^H\right)$, and
3. $D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right)-D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right)=L_{-i}\left(\vartheta^L\right)-L_{-i}\left(\vartheta^H\right)$, which also includes the case in which $L_{-i}\left(\vartheta^H\right)=L_{-i}\left(\vartheta^L\right)$.

If $D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right)-D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right)>L_{-i}\left(\vartheta^L\right)-L_{-i}\left(\vartheta^H\right)$, the lender other than i overcharges for the incremental liquidity between $L_{-i}\left(\vartheta^H\right)$ and $L_{-i}\left(\vartheta^L\right)$. This means that lender i can undercut this lender for this part of the liquidity provision. To do so, he only needs to change his pricing $D_i(\cdot)$ to $D_i'(\cdot)$ in such a way that

$$D_i'\left(\vartheta^L-L_{-i}\left(\vartheta^H\right)\right)=D_i\left(L_i\left(\vartheta^L\right)\right)+\left(D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right)-D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right)\right)-\varepsilon$$

for some small enough ε. This makes that, in case the low shock hits, the borrower wants to obtain $\vartheta^L-L_{-i}\left(\vartheta^H\right)$ from lender i and $L_{-i}\left(\vartheta^H\right)$ from the other lenders. This raises the borrower's utility by ε. Lender i will now have the same utility if the high shock hits, but if the low shock hits, instead of getting $D_i\left(L_i\left(\vartheta^L\right)\right)-L_i\left(\vartheta^L\right)$, he now gets

$$\begin{aligned}
&D_i'\left(\vartheta^L-L_{-i}\left(\vartheta^H\right)\right)-\left(\vartheta^L-L_{-i}\left(\vartheta^H\right)\right)\\
&\quad=D_i\left(L_i\left(\vartheta^L\right)\right)+\left(D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right)-D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right)\right)\\
&\qquad\quad-\left(L_i\left(\vartheta^L\right)+L_{-i}\left(\vartheta^L\right)-L_{-i}\left(\vartheta^H\right)\right)-\varepsilon\\
&\quad=\left(D_i\left(L_i\left(\vartheta^L\right)\right)-L_i\left(\vartheta^L\right)\right)\\
&\qquad\quad+\left(D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right)-D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right)-\left(L_{-i}\left(\vartheta^L\right)-L_{-i}\left(\vartheta^H\right)\right)\right)-\varepsilon,
\end{aligned}$$

which, for ε small enough, is greater than $D_i\left(L_i\left(\vartheta^L\right)\right)-L_i\left(\vartheta^L\right)$, so this constitutes a profitable deviation for lender i.

If $D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right)-D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right)<L_{-i}\left(\vartheta^L\right)-L_{-i}\left(\vartheta^H\right)$, the lender other than i is selling the incremental liquidity between $L_{-i}\left(\vartheta^H\right)$ and $L_{-i}\left(\vartheta^L\right)$ at a discount. This means that lender i can profit from this cheap liquidity. He can do so by offering the contract C_i' with a maximum liquidity provision of $\overline{L}_i':=\vartheta^H-L_{-i}\left(\vartheta_L\right)$ and with

$$D_i'\left(\vartheta^H-L_{-i}\left(\vartheta_L\right)\right)=D_i\left(L_i\left(\vartheta^H\right)\right)-D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right)-D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right).$$

This would make the borrower, in the high state, opt to obtain $L_{-i}\left(\vartheta_L\right)$ from lender $-i$, and $\vartheta^H-L_{-i}\left(\vartheta_L\right)$ from lender i, not changing her utility in the high state. For lender i, this gives, instead of a profit in state H of $D_i\left(L_i\left(\vartheta^H\right)\right)-L_i\left(\vartheta^H\right)$, a profit of

$$\begin{aligned}
&D_i'\left(\vartheta^H-L_{-i}\left(\vartheta^L\right)\right)-\left(\vartheta^H-L_{-i}\left(\vartheta^L\right)\right)\\
&\quad=D_i\left(L_i\left(\vartheta^H\right)\right)-\left(D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right)-D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right)\right)\\
&\qquad\quad-\left(L_i\left(\vartheta^H\right)+L_{-i}\left(\vartheta^H\right)-L_{-i}\left(\vartheta^L\right)\right)\\
&\quad=\left(D_i\left(L_i\left(\vartheta^H\right)\right)-L_i\left(\vartheta^H\right)\right)\\
&\qquad\quad-\left(D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right)-D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right)-\left(L_{-i}\left(\vartheta^L\right)-L_{-i}\left(\vartheta^H\right)\right)\right),
\end{aligned}$$

which is greater than $D_i\left(L_i\left(\vartheta^H\right)\right) - L_i\left(\vartheta^H\right)$. This constitutes a profitable deviation.

If, however, lender $-i$ sells the additional liquidity between $L_{-i}(\vartheta^H)$ and $L_{-i}(\vartheta^L)$ at a fair price, lender i necessarily sells his incremental liquidity between $L_i(\vartheta^L)$ and $L_i(\vartheta^H)$ at a discount, as in that case

$$\begin{aligned}\left(D_i\left(L_i\left(\vartheta^H\right)\right) - D_i\left(L_i\left(\vartheta^H\right)\right)\right) - \left(L_i\left(\vartheta^H\right) - L_i\left(\vartheta^H\right)\right)& \\ = \left(D(\vartheta^H) - D(\vartheta^L) - (\vartheta^H - \vartheta^L)\right).&\end{aligned}$$

In that case, it is lender $-i$ who can free-ride on the incremental provision of lender i by setting a new contract C'_{-i} with

$$D'_{-i}\left(\max\left\{\vartheta_L - L_i\left(\vartheta^H\right), 0\right\}\right) = D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right) - \left(D_i\left(L_i\left(\vartheta^H\right)\right) - D_i\left(L_i\left(\vartheta^H\right)\right)\right)$$

If $L_{-i}(\vartheta^L) > L_i(\vartheta^H) - L_i(\vartheta^L)$, lender $-i$ saves $L_{-i}(\vartheta^L)$ in the L-state by providing less liquidity, but pays by giving the discount of

$$\begin{aligned}D_i\left(L_i\left(\vartheta^H\right)\right) - D_i\left(L_i\left(\vartheta^H\right)\right)& \\ &= D(\vartheta^H) - D(\vartheta^L) + \left(D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right) - D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right)\right) \\ &= D(\vartheta^H) - D(\vartheta^L) + \left(L_{-i}\left(\vartheta^L\right) - L_{-i}\left(\vartheta^H\right)\right) \\ &\leq \vartheta^H - W + \left(L_{-i}\left(\vartheta^L\right) - L_{-i}\left(\vartheta^H\right)\right) \\ &\leq L_{-i}\left(\vartheta^L\right).\end{aligned}$$

This last inequality follows from the fact that as $L_{-i}\left(\vartheta^H\right) = \vartheta_H - L_i\left(\vartheta^L\right) > \vartheta_H - W$. This gives a profitable deviation.

This gives that both lenders must provide more money in the high state than in the low state, so I can now study the cases with $L_{-i}\left(\vartheta^H\right) > L_{-i}\left(\vartheta^L\right)$. Here I proceed in much the same way as before. First, I look at the case in which one lender overcharges for incremental liquidity, and then at the case when both sell incremental liquidity at least at a fair price, with at least one of them selling at a discount.

If one of the two lenders, let's say lender $-i$, overcharges for the incremental liquidity between $L_{-i}\left(\vartheta^H\right)$ and $L_{-i}\left(\vartheta^L\right)$, i.e. $D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right) - D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right) > L_{-i}\left(\vartheta^H\right) - L_{-i}\left(\vartheta^L\right)$, this gives that an inactive lender j can undercut this lender[2] for this part of the liquidity provision. To do so, he gives a contract with $J_j = 0$, with $\overline{L}_j = L_{-i}\left(\vartheta^H\right) - L_{-i}\left(\vartheta^L\right)$ and with

$$D_j(L) = \begin{cases} 0 & \text{if } L = 0 \\ D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right) - D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right) - \varepsilon & \text{if } L > 0, \end{cases}$$

for some small enough ε. In case the low shock hits, nothing changes, but when the high shock hits, the borrower wants to obtain $L_i\left(\vartheta^H\right)$ from lender i, $L_{-i}\left(\vartheta^H\right) - L_{-i}\left(\vartheta^L\right)$ from lender j and $L_{-i}\left(\vartheta^L\right)$ from lender $-i$. This raises the borrower's utility by $p_H\varepsilon$. Lender j will now have a profit of zero, as he did before, in case the low shock hits, but in case of the high shock, instead of getting zero, he now gets

$$\begin{aligned}D_j\left(L_{-i}\left(\vartheta^H\right) - L_{-i}\left(\vartheta^L\right)\right) - \left(L_{-i}\left(\vartheta^H\right) - L_{-i}\left(\vartheta^L\right)\right)& \\ = \left(D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right) - D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right)\right) - \left(L_{-i}\left(\vartheta^H\right) - L_{-i}\left(\vartheta^L\right)\right) - \varepsilon,&\end{aligned}$$

[2] If all lenders are active, this condition can be replaced by one of the active lenders adding this provision to the contract he is already trading, thereby increasing his profit.

which, for ε small enough, is greater than zero, so this constitutes a profitable deviation for lender j.

If for both lenders the incremental liquidity is sold at at least a fair price, then we have an i such that $D_i\left(L_i\left(\vartheta^H\right)\right) - D_i\left(L_i\left(\vartheta^L\right)\right) \leq L_i\left(\vartheta^H\right) - L_i\left(\vartheta^L\right)$, and $D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right) - D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right) < L_{-i}\left(\vartheta^H\right) - L_{-i}\left(\vartheta^L\right)$. This means that lender i can profit from the cheap liquidity offered by lender $-i$.

He can do so by offering the contract C_i' with a maximum liquidity provision of $\overline{L}_i' := \overline{L}_i$ and with

$$D_i'\left(\max\left\{\vartheta^L - L_{-i}\left(\vartheta^H\right), 0\right\}\right) = D_i\left(L_i\left(\vartheta^L\right)\right) - \left(D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right) - D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right)\right).$$

This would make the borrower, in the low state, opt to obtain $L_{-i}\left(\vartheta_H\right)$ from lender $-i$, and $\max\left\{\vartheta^L - L_{-i}\left(\vartheta^H\right), 0\right\}$ from lender i, not changing her utility in the high state. In case $\vartheta^L > L_{-i}\left(\vartheta^H\right)$, this gives lender i, instead of a profit in state L of $D_i\left(L_i\left(\vartheta^L\right)\right) - L_i\left(\vartheta^L\right)$, a profit of

$$\begin{aligned}
&D_i'\left(\vartheta^H - L_{-i}\left(\vartheta^H\right)\right) - \left(\vartheta^H - L_{-i}\left(\vartheta^L\right)\right)\\
&\quad = D_i\left(L_i\left(\vartheta^H\right)\right) - \left(D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right) - D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right)\right)\\
&\qquad + \left(L_i\left(\vartheta^H\right) + L_{-i}\left(\vartheta^H\right) - L_{-i}\left(\vartheta^L\right)\right)\\
&\quad = \left(D_i\left(L_i\left(\vartheta^H\right)\right) - L_i\left(\vartheta^H\right)\right)\\
&\qquad - \left(D_{-i}\left(L_{-i}\left(\vartheta^L\right)\right) - D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right) - \left(L_{-i}\left(\vartheta^L\right) - L_{-i}\left(\vartheta^H\right)\right)\right),
\end{aligned}$$

which is greater than $D_i\left(L_i\left(\vartheta^L\right)\right) - L_i\left(\vartheta^L\right)$. This constitutes a profitable deviation.

If, however, $\vartheta^L < L_{-i}\left(\vartheta^H\right)$

$$\begin{aligned}
D_{-i}&\left(L_{-i}\left(\vartheta^H\right)\right) - D_{-i}\left(L_{-i}\left(\vartheta^H\right)\right)\\
&= D(\vartheta^H) - D(\vartheta^L) + \left(D_i\left(L_i\left(\vartheta^L\right)\right) - D_i\left(L_{-i}\left(\vartheta^H\right)\right)\right)\\
&\leq D(\vartheta^H) - D(\vartheta^L) + \left(L_i\left(\vartheta^L\right) - L_i\left(\vartheta^H\right)\right)\\
&< \vartheta^H - W - \left(L_i\left(\vartheta^H\right) - L_i\left(\vartheta^L\right)\right)\\
&\leq L_i\left(\vartheta^L\right).
\end{aligned}$$

This last inequality follows from the fact that as $L_i\left(\vartheta^H\right) = \vartheta^H - L_{-i}\left(\vartheta^L\right) > \vartheta^H - W$. This means that for lender i the amount he can save by this deviation, i.e. L_i^L, is greater than the discount he has to give the borrower. Note that $D_i\left(L_i(\vartheta^L)\right)$ must be greater than $L_i(\vartheta^L)$ as lender i needs to break even, so that also $D_i'\left(\vartheta^H - L_{-i}\left(\vartheta^H\right)\right) \geq 0$. QED

Proof of Lemma 3.12. Consider the following set of strategies: at least two candidate intermediaries (call the set of these intermediaries $\mathcal{B}$), offer a contract to two of the other investors (call the set of these investors $\mathcal{D}$), asking for at least $1 + \vartheta^H - W$ units of money from both of them together, for which he will return precisely one unit of money at $t = 2$.

The intermediaries offer the borrower the contract with $J_i = 1$, with $\overline{L}_i = \vartheta^H$ and with $D(L)$ such that $1 + \vartheta^H - W \leq D(\vartheta^H) \leq R$ and $p_L D(\vartheta^L) + p_H D(\vartheta^H) = 1 + \mathbf{E}\vartheta$.

At least two investors other than the intermediary (call the set of these investors $\mathcal{L}$) offer the contract which offers no initial investment, offers $\overline{L}_i = \vartheta^H - \vartheta^L$ and with

$$D_i(L) = \begin{cases} 0 & \text{if } L = 0 \\ D & \text{if } L > 0 \end{cases}$$

for some very large D. Note that the sets $\mathcal{D}$ and $\mathcal{L}$ are not necessarily disjoint, one could even have them be exactly the same. I only assume $\mathcal{B} \cap \mathcal{D}$ and $\mathcal{B} \cap \mathcal{L}$ are empty. With these contracts on offer, the borrower chooses to do business with one of the investors in $\mathcal{B}$ and both investors in $\mathcal{D}$ deposit their money.

Any more attractive contract for the borrower should have $\overline{L} = \vartheta^L$, but if any one of the investors deviates by offering such a contract, the borrower would always take one of the contracts from $\mathcal{L}$ as well. Assume a borrower i deviates by offering a contract with $J_i = 1$ and $\overline{L}_i = \vartheta^L$. The borrower would then obtain a utility of

$$B + \frac{1}{2}\left(R - D_i(2)\right)_+$$

and investor i would have a utility of

$$-1 + \frac{1}{2}\left(D_i(2) - 2\right) + \frac{1}{2}\left(\frac{D_i(2)}{D_i(2) + D}R - 2\right).$$

In order to have the borrower prefer this contract, it must have $D_i(2) < R$, so that if D is large enough, the deviating investor's utility is negative. QED