Essays in Banking Theory

by

Tamas Vadasz

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Warwick Business School

September 2018

THE UNIVERSITY OF WARWICK

Contents

List of Figures

Acknowledgments

First and foremost, I would like to thank my supervisors, Prof. John Thanassoulis and Dr. Kebin Ma to guide me through the four years of the Finance PhD programme at Warwick Business School. I am indebted for all your advice, coaching and guidance, as well as for setting the bar for me so that I can constantly expand my boundaries. I am also thankful to all friends and fellow PhD students at Warwick Business School for their constant encouragement and support. I am grateful to the Finance Group at Warwick Business School for their financial support and funding.

Declarations

This thesis is submitted to the University of Warwick in support of the requirements for the degree of Doctor of Philosophy. I confirm that I have not submitted the thesis for a degree at another university. All work contained in the thesis is my own. Chapter 2 of this thesis is co-authored with Dr. Kebin Ma, while Chapters 3 with Professor John Thanassoulis.

Abstract

This thesis consists of three essays on banking theory.

In the first essay (joint with Dr. Kebin Ma) we analyse a possible informational impact of banking regulations. Banks can take costly actions (such as higher capitalization, liquidity holding, and advanced risk management) to fend off runs. While such actions directly affect bank risks, they also carry informational content as signals of the banks' fundamentals. A separating equilibrium due to such signalling, however, involves two types of inefficiency: the high type chooses excessively costly signals, whereas the low type is vulnerable to runs. This provides a novel rationale for financial regulations: by restricting banks' actions, regulators can maintain a pooling equilibrium where the cross-subsidy among types promotes financial stability. We build a theoretical model to illustrate the point and also obtain supporting evidence from the US capital and liquidity regulations.

The second essay (joint with Prof. John Thanassoulis) seeks to provide a theoretical explanation of the variety of pricing schemes and product bundling observed in personal current account (PCA) markets. The main motivating fact is the widespread proliferation of 'free-if-in-credit' (FIIC) current accounts in certain countries (US, UK), in contrast to some other European countries (France, Italy, Hungary), where even basic current account services are subject to excessive monthly fees. Existing evidence is consistent with the possibility that FIIC current accounts are cross-subsidized by exploitative and complicated fee structures on connected products, in particular by the excessive usage of overdraft facilities. In this research we propose a novel approach to model competitive aftermarkets, and demonstrate how certain sources of market power, namely customer naiveté and adverse selection interact in equilibrium. This helps to better understand why some markets are more likely to develop FIIC pricing than others.

In the last chapter I demonstrate how illiquidity is determined endogenously during crises as a result of equilibrium behaviour of financial institutions subject to leverage constraints. I show in a simple and intuitive framework that asset liquidation decisions exhibit similar characteristic to a Prisoners' dilemma: although financial institutions are given the possibility to dampen the cost of fire-sale spillovers, the only Nash-equilibrium is where banks 'defect', and end up coordinating on selling the more liquid common asset, which in turn becomes illiquid. This reduces welfare compared to the socially optimal de-leveraging rule.

Chapter 1

Introduction

This thesis consists of three essays on banking theory. The common theme of these research papers, and thereby the organizing principle of my thesis is the presence of various frictions in banking markets, which call for the attention of policy makers, and potentially require regulatory intervention. The essays are related to relevant and recently actively debated policy issues concerning the stability of financial markets, micro- and macroprudential regulation, as well as non-competitive distortions in the banking sector.

The dominant approach in this thesis is theoretical, and is based on applied game-theory. As an applied theorist I believe that an important purpose of theoretical research is to help to understand various real-world problems, and answer practical questions. Mathematical rigorousness is a convenient tool to enlighten non-trivial perspectives, which can eventually lead to better decision making, and a more smoothly functioning financial system. In this spirit, the three chapters investigate the following topics: Chapter 2 discovers a novel, information-based impact of regulations in banking, Chapter 3 explores exploitative and welfare-reducing overdraft pricing practices in retail banking, while Chapter 4 points out a potential 'liquidity trap' in a financial crisis situation due to asset commonalities. These issues all require attention, monitoring, and potentially, intervention by the relevant regulatory bodies.

Chapter 2 (*"Bank Signalling, Risk of Runs and the Informational Impact of Regulations", joint with Dr. Kebin Ma*) is motivated by post-crisis changes in banking regulation, and presents a novel perspective on the potential impact of those regulations. Financial institutions often make use of costly actions (such as higher capitalization, liquidity holding, and advanced risk management) to fend off runs. While such actions directly affect bank risks, they also carry informational content as signals of the banks' fundamentals. We show that a separating equilibrium of this signalling game between the bank and its creditors involves two types of inefficiency: the high type chooses excessively costly signals, whereas the low type is revealed, and is vulnerable to runs. This provides a novel rationale for financial regulations. Placing restrictions on banks' actions in the form of a conventional microprudential regulatory tool, such as capital requirements or the recently introduced Liquidity Coverage Ratio, affects the value of information conveyed by risk management actions as a signalling device, and in turn, the incentives to engage in discretionary risk management. Our theoretical model shows that a minimum quantitative requirement can eliminate the separating equilibrium by making it more difficult to signal private information, thereby regulators can maintain pooling, where the cross-subsidy across types promotes financial stability.

Our model provides a novel perspective on financial regulation, which fundamentally differs from the traditional ones that emphasize its role in mitigating moral hazard or containing potential negative externalities on the real economy. By eliminating a way in which markets create information, regulators create ignorance, which is efficient, as it leads to greater financial stability and higher social welfare. Our mechanism also provides an explanation for financial institutions' reaction to the introduction of a new regulation: a sufficiently restrictive regulatory threshold induces pooling, and institutions initially not constrained by the to-be introduced quantitative regulation, optimally decrease their level of risk management towards the new regulatory limit, which now serves as a focal point. This aspect of our theory emphasizes a latent link between microprudential and macroprudential regulatory perspective: regulating some individual institutions changes the prevailing equilibrium, and thereby the behaviour of other market participants, affecting the stability of the system as a whole.

The model leads to testable empirical predictions. If financial regulations do squeeze out separating equilibria, we would expect relatively high dispersion of risk management measures among banks before the introduction of pertinent financial regulations, and clustering of observations after the introduction of the regulations. We test this hypothesis on two data sets: cash holdings of US Bank holding companies (BHC's), using a difference-in-difference method which exploits the recent introduction of Liquidity Coverage Ratio (LCR), as well as changes in capital ratios around the introduction of Basel I regulatory capital regime. We find two distinct patterns, both consistent with the predictions of our theory: first, the dispersion of cash ratios for BHC's subject to the new regulation decreased significantly more sharply than those which were not subject to the new regulation.

This is consistent with a successful elimination of separating equilibrium. For the case of capital regulation, we find an increase of the number of institutions with large equity ratios, which might be the result of an insufficient regulatory minimum, being unable to squeeze out, but boosting the signals required to maintain separating equilibrium.

The model combines signalling with a stylized bank-run game. Our methodological contribution is to develop a novel technique to analyse a global game (which is necessary to arrive at a unique equilibrium prediction) embedded into a signalling game in a tractable way which also facilitates welfare analysis. We illustrate our approach with a linear regime switching function which leads to a closed-form solution, and generalize to a larger class of models satisfying a single-crossing property.

Chapter 3: A different policy issue and market friction is at the centre of Chapter 3, titled *"Current account and Overdraft Pricing in Retail Banking" (joint with Prof. John Thanassoulis)*, which brings us to the search for a pricing equilibrium in retail banking. At first look, the market structure of banking in developed countries is a puzzle: although it possesses many key characteristics of competitive, or at least contestable markets, evidence of monopolistic behaviour in certain segments is widespread. We study a particular market segment: the markets for overdrafts. In some countries, especially in the US and the UK, the predominant personal current account pricing scheme is the so-called 'free-if-in-credit' (FIIC) pricing, sometimes loosely referred to as 'free banking'. Under this price schedule, banks charge zero monthly or regular fee for the access to the account and for basic services. However, accounts are usually bundled together with an overdraft-service - essentially a short-term borrowing facility - which allows the customer to go into debit in her account, incurring extensive charges in the form of service fees or interest payments. In contrast, in some other European countries (for example France, Italy, Hungary) even basic current account services are subject to - sometimes quite expensive - monthly fees, while overdraft charges are less important sources of banks' profit. The purpose of this research is to better understand the equilibrium consequences of some underlying market frictions, such as customer naiveté and information asymmetry, on prices, profits, and on the market structure.

Many observers point out that FIIC-pricing is consistent with a cross-subsidy across business lines as well as across various groups of customers, with significant potential welfare consequences. According to the common narrative, hidden and expensive overdraft fees exploit poorer households — more likely in need of extra liquidity — to support more sophisticated, wealthier clients, and to generate industry rents. As FIIC-pricing essentially amounts to below-marginal-cost base-good, and above-marginal-cost add-on prices, to the extent different groups of customers differ in their use of the two services, it leads to potential distributional concerns and welfare effects. To be specific, if overdraft users tend to be poorer households more likely in financial trouble and in the need of short-term financing, FIIC-pricing implies a wealth-transfer 'from the poor to the rich', leading to an overall reduction in social welfare. This view has generated a massive policy debate on the potential role and shape of regulatory intervention. However, existing theories fall short of explaining why FIIC-pricing develops in certain countries only.

In this paper we analyse a two-stage duopoly model of overdraft pricing which captures some relevant characteristics of the retail banking sector to illustrate how FIIC-pricing can emerge as an equilibrium phenomenon. Deviating from much of the theoretical literature, which treats overdrafts as a monopolistic aftermarket, we assume that competition for customers is possible even in the overdraft-stage. This captures the idea that customers' lock-in is not perfect in a financial services context. Indeed, the nature of an existing relationship to one's bank is distinctively different from a typical industrial aftermarket situation - the most cited example being the market for printers and cartridges - where technological linkages tie customers strongly to the primary-market supplier, while high initial investment costs prevent switching to another provider. Strictly speaking there is no such technological reason to link payment services (the primary market) to the provision of short-term customer credit (the aftermarket), yet the market shows similar characteristic: lack of switching behaviour, and 'exploitative' aftermarket pricing.

Our results demonstrate that, despite the possibility of competition in the second-stage, the presence of naive customers turns to be an important source of market power and economic profit. Specifically, we show that for an arbitrary low number of naive customers, prices deviate from marginal-cost pricing, and the unique Nash-equilibrium of the Bertrand-game on the overdraft market is a mixed-strategy Nash-equilibrium, where both insider and outsider banks earn positive profit. This makes overdraft a profitable business, and induces competition in the first stage (on the market for PCA) to expand market share. When primary markets are sufficiently competitive, there exist a symmetric 'FIIC-equilibrium' for a significant subset of the parameter space - specifically, FIIC-pricing can prevail even with relatively modest number of naive customers, and on highly competitive markets.

In the rest of the paper we extend the baseline setup with adverse selection, and show how the combination of adverse selection and customer naiveté, being the two most important frictions on banking markets, affect equilibrium pricing, profits, and customer behaviour. The presence of adverse selection makes it harder for the outsider to enjoy the benefits on the aftermarket. As a consequence, the profits will be tilted towards the bank's role as an insider. This increases the incentives to obtain more customers on the primary market, so it decreases first-period PCA prices even further — making it even more likely that an FIIC-pricing prevails as an equilibrium. In our model differences in the mass of myopic customers is not the only possible explanation anymore for the observed differences across countries: 'free banking' may or may not develop depending on the severity of adverse selection, or the extent of primary market competition as well.

Chapter 4: Endogenous reaction to a systemic shock in an interconnected banking environment is the subject of Chapter 4 in this dissertation, *"Fire sale in a liquidation game with leverage requirements"*. In the aftermath of the financial crisis, the view that interconnectedness is an important determinant of financial stability became conventional wisdom among academics and policy-makers. A large theoretical and empirical literature on systemic risk started to emphasize how various forms of business relations in the financial sector can turn to a transmission channel through which shocks propagate in the financial system, eventually leading to systemic bank failures and causing real economic losses. One potential layer of interconnectedness, which is the subject of this paper, is indirect linkages through common investments, or *'asset commonalities'*. If an investor is forced to liquidate their asset due to some funding pressure, prices may depart from fundamental values. Mark-to-market evaluation of portfolios forces other investors of the same asset to re-evaluate their portfolio, which decreases equity value. In turn, the drop in equity induces additional funding pressure, and those - otherwise healthy - institutions may be forced to engage in further asset liquidation.

In this paper I explicitly model the asset liquidation *decision* of financial institutions under funding pressure, when multiple asset classes are available to adjust the portfolio. In the model, the 'funding pressure', which is the key market friction behind this phenomenon, is captured by a leverage constraint: following a (systemic) asset-price shock, the banking system may be forced to engage in systemic deleveraging to restore leverage targets by selling assets and repaying debt. The novelty of the analysis is to focus on banks' optimal, equilibrium decision instead of accounting for rule-based, deterministic de-leveraging spillover effects usually studied in the academic literature. The investment portfolio on banks' balance sheets differ in ex-ante liquidity, measured as the market price impact following an asset sale during 'normal times'. Equity-maximizer financial institutions adjust their portfolio by choosing to sell assets so that the impact on equity is minimized. In the presence of asset commonalities, if all banks end up selling the same asset class ('commonality'), liquid assets suddenly may appear illiquid and can be sold only at a significant fire-sale discount, a phenomenon which was widely observed during the financial crisis. The endogenous determination of the fire-sale price has to be taken into consideration by rational financial institutions.

The joint deleveraging decision of interlinked financial institutions induces a non-cooperative game which we dub '*the liquidation game*'. The main result of this paper is that as long as the equilibrium liquidation decision of the banks is non-trivial in the sense that liquidating only one single asset does not strictly dominate, the emerging unique Nash-equilibrium is not Pareto-optimal. Individual banks could achieve higher ex-post equity value by choosing another feasible liquidation strategy, which, however, cannot be maintained as an equilibrium. The market outcome in equilibrium is reminiscent to a Prisoner's dilemma: cooperation, which in this context would mean self-restraint in selling the more liquid asset commonalities and relying more on idiosyncratic but less liquid assets to restore leverage, could increase the overall payoffs for each players, but cannot be maintained as an equilibrium. In the unique Nash-equilibrium banks 'defect', and over-liquidate the commonality.

The comparison of the equilibrium and the social planner's optimal solution reveals an even more striking feature: the potential loss from the inefficient equilibrium may even be larger, if markets appear to be ex-ante more liquid. Intuitively, more liquid commonality raises the incentives to tilt the liquidation strategy towards that asset class, which leads to an even larger equilibrium price effect, and further diminishes equity. This finding has slightly uncomfortable consequences for financial stability: higher liquidity, although almost unanimously called for by policy-makers after the crisis, can even be detrimental in highly integrated markets, if fire-sale decisions following a potential shock are jointly determined in an equilibrium.

Chapter 2

Bank Signalling, Risk of Runs, and the Informational Impact of Regulations

2.1 Introduction

Financial institutions often make use of costly actions to communicate private information about their fundamentals.[1] A particularly important class of such signalling devices is banks' quantitative risk management choices, such as the amount of capital, or high-quality liquid assets, whose primary goal is to tackle solvency risk and liquidity risk respectively. These risk management actions are heavily regulated, requiring banks to maintain adequate levels of such risk measures in circumstances where externalities or other market imperfections would prevent to reach socially optimal outcomes as a decentralized solution in a laissez-fair environment.

In this paper we study the interaction between such regulations and the bank's incentives to signal private information to its creditors. In our model, placing restrictions on banks' behaviour in the form of a conventional microprudential regulatory tool, such as capital requirements or the recently introduced *Liquidity Coverage Ratio* (LCR), affects the value of information conveyed by risk management actions as a signalling device, and in turn, the incentives to engage in discretionary risk management. In a no-regulation environment, a separating equilibrium due to signalling involves two types of inefficiency: the high-type chooses excessively

[1] A classic example would be banks maintaining high dividend payouts and executive compensation during the crisis, in the endeavour to convince the market of their relatively strong financial positions.

costly risk management, whereas the low type is revealed and becomes vulnerable to runs. In this sense, the private information creation can bear potential social inefficiency. We show that a minimum quantitative regulation can eliminate separating equilibrium and enforce pooling by making it more difficult to signal private information, thereby mitigating the related inefficiencies.

Our model provides a novel perspective on financial regulation, which fundamentally differs from the traditional ones that emphasize its role in mitigating moral hazard or containing potential negative externalities on the real economy.[2] By eliminating a way in which markets create information, regulators create ignorance, which is efficient, as it leads to greater financial stability and higher social welfare. Our mechanism also provides an explanation for financial institutions' reaction to the introduction of a new regulation: a sufficiently restrictive regulatory threshold induces pooling, and institutions initially not constrained by the to-be introduced quantitative regulation, optimally decrease their level of risk management towards the new regulatory limit, which now serves as a focal point. This aspect of our theory emphasizes a latent link between microprudential and macroprudential regulatory perspective: regulating some individual institutions changes the prevailing equilibrium, and thereby the behaviour of other market participants, affecting the stability of the system as a whole.

Our model combines signalling with a stylized bank-run game, where the unique equilibrium of the coordination problem of creditors is determined by global games techniques. We parameterize the strength of the bank by two distinct fundamental variables: the bank's innate and its financial fundamental. While insiders have private information regarding the bank's innate ability to effectively make use of costly risk management tools to fend off runs, a lack of common knowledge regarding the financial fundamental drives the global game equilibrium selection in the second stage of the game.

This model can be solved analogously to a conventional signalling game where the receivers' (creditors) unique aggregate responses to any on- and off-equilibrium action are determined by global-games techniques. Our methodological contribution is to develop a novel technique to analyse a global game embedded into a signalling game in a tractable way which also facilitates welfare analysis. We illustrate our approach with a linear regime switching function which leads to a closed-form solution, and generalize to a larger class of models satisfying a single-crossing property.

[2]For example, minimum capital regulation is often justified to correct moral hazard and risk-taking incentives of shareholders, while the recent introduction of quantitative liquidity regulation is motivated by decreasing reliance on 'public liquidity' and building up sufficient private cushions to withhold liquidity shocks.

The first main result of the paper provides conditions for the existence of a separating equilibrium in which a high-type bank chooses an excessively high signal. We show that the existence of this equilibrium, as well as the magnitude of inefficiency, is inherently linked to the precision of receivers' noisy private signals, with higher precision leading to larger inefficiencies. Intuitively, higher precision leads to more correlated aggregate behaviour of creditors in any states of the world, in turn, leading to more pronounced aggregate responses to (perceived) changes in unknown parameters. This encourages the low-type to mimic the high-type more aggressively, who, in turn needs to send a higher signal to maintain separation.

Second, we show that a regulator can eliminate inefficient *high signals* by setting a *minimum* threshold on the bank's risk management action. Under such regulations, the minimax payoff for the low-type (i.e. the payoff she could get irrespectively of the other type's behaviour) decreases, which in turn increases the critical signal that is required for the high-type to maintain separation. For a sufficiently restrictive regulation, this action is too costly, and it is no longer incentive-compatible for the high-type to maintain the separation. This *critical* regulatory minimum also changes with precision: when precision is high, the separating signal is already high, therefore a relatively low regulatory threshold is sufficient to induce pooling. An implication is that signalling may emerge during turbulent times (characterized by low precision of observation of the fundamentals): a certain level of regulation which is just sufficient to maintain pooling during normal times might not be able to prevent wasteful signalling during turbulent times.

We perform a preliminary welfare analysis tailored specifically to the context of financial regulations. We show that a minimum ratio regulation can indeed increase ex-ante welfare by squeezing out separating equilibria. A separating equilibrium in the model leads to two types of inefficiency: the high-type chooses a signal that is excessively high and costly, whereas the low-type is identified as weak and becomes vulnerable - resulting in more runs and greater financial instability. In contrast, in a pooling equilibrium, the high-type will cross-subsidise the low-type, and the economy can feature greater financial stability as well as a reduction in costly signalling. In this sense, financial regulations reduce the information available in private markets, and the resulting ignorance is efficient. The ex-ante improvement in expected profits implies that it can be incentive-compatible for banks to accept financial regulations.

Our model leads to testable empirical predictions. If financial regulations do squeeze out separating equilibria, we would expect relatively high dispersion of risk management measures among banks before the introduction of pertinent financial regulations, and clustering of observations after the introduction of the regulations. We test this hypothesis on two data sets: cash holdings of US Bank holding companies (BHC's), using a difference-in-difference method which exploits the recent introduction of Liquidity Coverage Ratio (LCR), as well as changes in capital ratios around the introduction of Basel I regulatory capital regime. We find two distinct patterns, both consistent with the predictions of our theory: first, the dispersion of cash ratios for BHC's subject to the new regulation decreased significantly more sharply than those which were not subject to the new regulation. This is consistent with a successful elimination of separating equilibrium. For the case of capital regulation, we find an increase of the number of institutions with large equity ratios, which might be the result of an insufficient regulatory minimum, being unable to squeeze out, but boosting the signals required to maintain separating equilibrium.

The methodology of our paper is most related to Angeletos et al. [2006], and Angeletos and Pavan [2013]. In their pioneering work, the authors consider a perfectly informed policy maker (sender) who tries to defend a regime with possible policy intervention and show that once the signalling effect of the policy intervention is taken into consideration, multiple equilibria will re-surface in a global-game setting due to the endogeneity of the attackers' (receivers') information set. The specific form of multiplicity in the (semi-) separating equilibrium arises due to the fact that there is no uncertainty regarding the regime outcome from the sender's perspective. As a consequence, any positive policy intervention signals the survival of the regime, which makes 'no attack' a dominant strategy, and the global game is played out over a truncated posterior distribution on the range of fundamentals when intervention does not occur.

In contrast to this work, as well as a growing literature on persuasion with multiple receivers (Inostroza and Pavan [2017], Goldstein and Huang [2016]) the sender in our model only imperfectly observes the fundamentals. In the context of banking, insiders such as bank equity holders or managers (i) can have an informational advantage over their creditors regarding the bank's resilience to shocks but (ii) still face uncertainty regarding the fate of the bank. Apart from being more realistic in the context of banking risk management, the modelling role of residual uncertainty on the sender's side is crucial: despite her informational advantage, this additional uncertainty keeps the sender uncertain regarding the fate of the regime, therefore policy intervention cannot make even the highest type bank completely 'run-proof', although it changes incentives to run.

Related literature. The idea that simple risk management measures such as capital or liquidity can signal private information beyond the fact that higher values can mechanically protect the bank against shocks has been proposed in the literature before. For example, Hughes and Mester [1998] writes: *"Since financial capital constitutes the bank's own bet on its management of risk, it conveys a credible signal to depositors of the resources allocated to preserving capital and insuring the safety of their deposit"*. The signalling role of capital is also well recognized generally in the corporate finance literature, albeit with somewhat inconclusive predictions (Ross [1977],Brealey et al. [1977] and Harris and Raviv [1991]). Malherbe [2014] interprets the bank's liquid holdings as a signal of the underlying reason for asset sales, so that a higher liquidity might increase adverse selection in asset markets. In other papers, asymmetric information concerns the quality of assets, and banks are sending credible signals either through proper security design (Nicolo and Pelizzon [2008]) or by retention (He [2009]). An extensive literature in accounting surveys the signalling role of loan loss provisioning (LLP), with many papers arguing that higher LLP credibly signals a prudent risk management, and management's intention to resolve problem debt situations[3].

Our model predicts a positive relationship between the level of risk management measures and the value of the bank in case of first-best as well as whenever separating equilibrium still prevails on the markets. In the case of capital, this is consistent with Mehran and Thakor (2010), who present an elegant theory and strong empirical support for a positive cross-sectional correlation between bank capital and market value. In their model, increased capital has two effects: it increases the probability of survival and in turn, incentives to monitor (direct effect), while increased loan monitoring enhances the value of the portfolio (indirect effect). The overall impact of the two effects is that banks with lower monitoring costs will have a higher marginal benefit of capital, and in turn, find it optimal to hold more.

The paper is also related to the large literature on bank runs. Since the seminal contribution of Diamond and Dybvig [1983], it is well known that liquidity transformation makes banks vulnerable to runs driven by agents' self-fulfilling beliefs regarding the behaviour of other agents. In the more recent follow-up literature, global games theory[4] has been routinely used to resolve the equilibrium selection problem in the Diamond-Dybvig framework (Goldstein and Pauzner [2005]). It has also been pointed out that liquidity, as well as capital – ceteris paribus – can serve as a buffer, thereby dampening the probability of distress and increasing financial stability (Diamond and Rajan [2000], Diamond and Kashyap [2016]).

[3]Some early contributions are: Beaver and Engel [1996], Scholes et al. [1990], Grammatikos and Saunders [1990], Griffin and Wallach [1991]

[4]Carlsson and Van Damme [1993] and Morris and Shin [1998]

The paper is organized as follows: we introduce our model in Section 2. In Section 3, we analytically solve the model for a no-regulation equilibrium with stylized functional forms. Section 4 analyses the impact of a minimum quantitative regulation and discusses the most important welfare trade-offs. Section 5 provides some empirical insights. Finally, we generalize some of our results to a larger class of functional forms in the Appendices.

2.2 Model setup

We consider a two-period game played by two groups of players: a bank and its creditors. In our model, a continuum of creditors of mass 1 hold demandable claims of the bank, and simultaneously decide whether to run on the bank or not. The bank, on the other hand, is incentivized to defend itself from runs by implementing costly risk management practices.

The bank's fundamental strength is determined by two random variables, θ_1 and θ_2, which we assume to be independently distributed. We interpret θ_1 as the bank's inherent ability to manage its risks, that we dub as skill. The bank can be of low-quality, $\theta_1 = \theta_1^L$, with probability p, or high-quality, $\theta_1 = \theta_1^H$, with probability $(1-p)$, where $0 < \theta_1^L < \theta_1^H$. Parameter θ_2 captures the financial fundamental of the bank and is drawn from a uniform distribution with support on the interval $\left[\underline{\theta}_2, \overline{\theta}_2\right]$.

In period 1, the bank privately observes the realization of θ_1, which is the bank's Harsanyi-type. Upon the observation, but before the realization of θ_2, the bank chooses a costly, non-negative risk management action $s \in (0, +\infty)$ to enhance its ability to survive runs. The bank's strategy, therefore, specifies a choice of s for each possible realization of θ_1. The risk management action s influences the ability to survive runs directly, and also serves as an informative public signal for the bank's type.

The bank fails if sufficiently many creditors decide to run. In particular, we capture the failure of the bank with a continuous, differentiable, and real-valued regime switching function $\mathcal{R}(\theta_1, \theta_2, s, \alpha)$. The bank fails whenever $\mathcal{R}(\theta_1, \theta_2, s, \alpha) < 0$, where α denotes the mass of creditors who run on the bank. We assume that the regime switching function is such that the bank's survival is more likely if any of the fundamentals or the risk management action are higher, and less likely if the mass of creditors who run is greater. Furthermore, the more skilled bank benefits more from the risk management action for any given level of s,[5] meaning the fundamental θ_1

[5]These requirements imply that the derivatives of the regime switching function satisfy $\mathcal{R}_{\theta_2} > 0$, $\mathcal{R}_{\theta_1} > 0$, $\mathcal{R}_s > 0$, $\mathcal{R}_\alpha < 0$, and $\mathcal{R}_{s\theta_1} > 0$.

is a measure of how effective the bank's pre-emptive intervention can be in avoiding bankruptcy.[6] For the main part of the paper, we solve the model analytically for the following functional form.

$$\mathcal{R}(\theta_1, \theta_2, s,\alpha) := \theta_1 s + \theta_2 - \alpha$$

The bank's payoff is specified as follows:

$$U(\theta_1, \theta_2, s, \alpha) = \begin{cases} k - c \cdot s & \text{if} \quad \mathcal{R}(\theta_1, \theta_2, s, \alpha) \geq 0 \\ 0 & \text{if} \quad \mathcal{R}(\theta_1, \theta_2, s, \alpha) < 0 \end{cases}$$

where k is the benefit of surviving the bank run, c is the unit cost of risk management, and the payoff conditional on failure is normalized to zero. One interpretation of the payoff structure is that k captures the charter value of the bank which would be lost if the bank fails, while the zero payoff in case of failure reflects the fact that banks are protected by limited liability. For simplicity, k and c are exogenous constants.

In period 2, the unit mass of creditors perfectly observe the public signal of s, and each of them receives a private, noisy signal of the bank's financial fundamental θ_2. In particular, a creditor $i \in [0, 1]$ receives private signal $x_i = \theta_2 + \sigma\epsilon_i$, where $\epsilon_i \sim U(-1, 1)$ is independently and identically distributed across creditors. The parameter $\sigma > 0$ captures the accuracy of the private signals. Based on the information, the creditors simultaneously decide whether to run on the bank or not, the two actions we denote by RUN and WAIT. We focus on the symmetric strategy equilibrium, since creditors are ex ante identical. A creditor's payoff from choosing action RUN is normalized to a constant $t \in (0, 1)$, while the payoff from action WAIT depends on whether the bank survives the runs, and is specified as follows.

$$u(\theta_1, \theta_2, s, \alpha) = \begin{cases} 1 & \text{if } \mathcal{R}(\theta_1, \theta_2, s, \alpha) \geq 0 \\ 0 & \text{if } \mathcal{R}(\theta_1, \theta_2, s, \alpha) < 0 \end{cases}$$

We will assume throughout that fundamentals are such that even the strongest banks can fail for any risk management actions which is taken in equilibrium, while for a sufficiently high realization of the fundamental, a bank survives even if all creditors run on the bank. These assumptions guarantee the existence of *dominance*

[6] One may interpret θ_1 as the bank's skill in screening loans and s as the bank's capital ratio. A better screening skill helps the bank to maintain a higher asset quality that is privately observed, and the same level of capital helps a bank better if the bank has higher asset quality. Alternatively, one may consider s being the amount of liquid asset held by the bank and θ_1 being the market liquidity of those asset in a crisis, or θ_1 being the bank's human capital in advanced risk modelling and s being IT infrastructure required in its implementation.

regions, where creditors find it dominant to run (resp. wait).

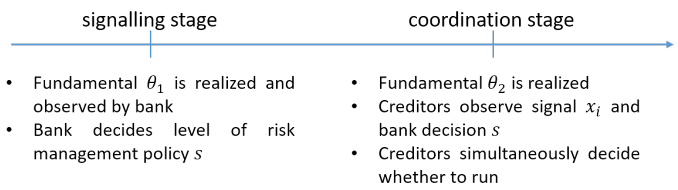
Assumption 1 *We assume that the following parameter restrictions are satisfied:*

- *For every* $\theta_1 \in \Theta$*, and for every* $s \in \mathcal{S}^\star$*,* $\underline{\theta}_2 < -\theta_1 s := \theta_\star$ *where* $\mathcal{S}^\star$ *is the set of equilibrium values of signal* s
- *For every* $\theta_1 \in \Theta$*, and for every* $s \in \mathcal{S}^\star$*,* $\overline{\theta}_2 > 1 - \theta_1 s := \theta^\star$

Assumption 1 implies that the creditors' bank run game has two dominance regions: there exist $\theta_\star$ and $\theta^\star$ such that when $\theta_2 < \theta_\star$, RUN is a dominant strategy for creditors independent of their private signals. And when $\theta > \theta^\star$, WAIT is a dominant strategy for creditors independent of their private signals.

The timeline is summarized in Figure 2.1. Note that the sequence of the game has a natural interpretation: a bank's skill for risk management can be slow-moving and affects the bank's risk management decisions. In the model, this is reflected in the exogenous type θ_1 and the timing that s is chosen based on the private information of θ_1. The financial fundamental (e.g., the default rate of the loan portfolio), on the other hand, can fluctuate more frequently and therefore not (fully) revealed to the bank when the risk management decision is made. While the creditors' decision to run on the bank may be instantly triggered by contemporaneous changes of the financial fundamental, the bank may not be able to change its amount of liquidity or capital equally fast in the presence of financial market frictions.

Figure 2.1: Timeline



The simple functional forms for the regime switching function as well as for the payoff functions are selected to sharpen the intuition and emphasize the interaction between the signalling and coordination stages of the game.[7] In 2.B, we analyse some of the consequences of these modelling choices and generalize the results to a broader set of functional forms.

[7]Indeed, the coordination stage can be straightforwardly recast as a backbone global game of regime change where imperfectly informed atomistic players (creditors) play a game with strategic complementarities whether to attack (WAIT) or not (RUN) a regime (the bank) whose survival depends on its fundamentals (θ_1, θ_2), actions (s), and the mass of atomistic players attacking the regime (α).

2.3 Equilibrium analysis

We use Perfect Bayesian Equilibrium as a solution concept. Let $s(\theta_1) \in [\underline{s}, \infty)$ denote the strategy of the bank, i.e. the risk management action chosen by type θ_1, $a(x_i, s)$ denote the action of an agent receiving private signal x_i and publicly observing risk management s, and $\alpha(\theta_1, \theta_2, s)$ denote the mass of creditors who run. We define an equilibrium as follows.

Definition 1 *A symmetric, Perfect Bayesian Equilibrium of the signalling-global game consists of (1) a strategy $s^\star(\theta_1) : \{\theta_1^L, \theta_1^H\} \rightarrow [\underline{s}, \infty)$ for the bank, (2) a strategy $a^\star(x_i, s) : \mathbb{X} \times [\underline{s}, \infty) \rightarrow \{RUN, WAIT\}$ for creditors; (3) posterior beliefs on $\{\theta_1, \theta_2\}$ for creditors upon observing $\{x_i, s\}$: $\mu_i(\theta_1, \theta_2 | x_i, s) : \mathbb{X} \times [\underline{s}, \infty] \rightarrow [0, 1]$, such that*
(1) Bank's strategy is profit-maximizing given aggregate runs:

$$s^\star(\theta_1) \in \arg\max_s \mathbb{E}U(\theta_1, \theta_2, s, \alpha(\theta_1, \theta_2, s))$$

(2) Creditors' decision whether to run is profit-maximizing given their information set

$$a^\star(x_i, s) \in \arg\max_a \mathbb{E}u(t, \theta_1, \theta_2, s | x_i, s)$$

where aggregate attack α is consistent with individual decisions

$$\alpha(\theta_1, \theta_2, s) = \int_0^1 a^\star(x_i, s) di$$

(3) Beliefs are updated using Bayes' rule whenever possible. Strategies are sequentially rational, and consistent with beliefs.

We solve the model backwards: First, we solve the coordination game for any given risk management level s, then determine the bank's (sender) optimal choice of risk management, given the second stage equilibrium.

2.3.1 Symmetric information benchmark

Our model parsimoniously captures the combination of two informational frictions: first, there is informational asymmetry between the bank and the creditors regarding the value of θ_1. Second, there is incomplete information leading to strategic uncertainty regarding creditors' beliefs about each others' actions.

To set up a benchmark, we start the analysis with a version of the game where the value of θ_1 is observed by the creditors as well. Without information asymmetry regarding θ_1, there is no signalling role for risk management action.

Notice that in the complete information version of the game, i.e. in the one where θ_2 is also perfectly observed by the creditors, there would be multiple equilibria of the subgame whenever the value of θ_2 is outside of the dominant regions, that is, $\theta_2 \in (\theta_\star, \theta^\star)$. This is caused by the strategic complementarities in the coordination problem of creditors. We obtain uniqueness of the equilibrium using standard global games refinement (Morris and Shin [1998], Morris and Shin [2001]) in Lemma 1.

Lemma 1 *When parameter θ_1 is perfectly observed by the creditors, the unique equilibrium of the coordination-stage that survives iterated elimination of strictly dominated strategies is characterized by two thresholds*[8]

$$\hat{x}\,(\theta_1, s) = t - \theta_1 s + 2\sigma t - \sigma \tag{2.1}$$

$$\hat{\theta}_2\,(\theta_1, s) = t - \theta_1 s \tag{2.2}$$

such that creditor i runs if and only if $x_i < \hat{x}$, and the bank fails if and only if $\theta_2 < \hat{\theta}_2$.

Both thresholds are decreasing in type θ_1 and in action s, which implies that higher fundamental, as well as higher intervention, make survival of the bank more likely.

Proof. See Appendix ■

As stated by Lemma 1, the equilibrium of the subgame is characterized by a pair $\{\hat{\theta}_2, \hat{x}\}$ which jointly solves two equations: (i) a creditor who receives signal $\hat{x}$ is just indifferent between RUN and WAIT, and (ii) the bank just fails at $\hat{\theta}_2$. The proof in the Appendix derives these two conditions and proves that the solutions constitute the unique equilibrium of the subgame.

Next, we solve for optimal risk management, given that the bank anticipates correctly the equilibrium in the second stage. For any choice of s, the equilibrium quantities $\{\hat{\theta}_2(\theta_1, s), \hat{x}(\theta_1, s)\}$ determine the mass of agents who run on the bank (α), and in turn, the probability of survival. Therefore the expected profit, which is the bank's objective function, can be expressed as a function of the exogenous type θ_1 and the endogenous risk management action s:

$$\pi(\theta_1, s) := \mathbb{E}[U|\theta_1, s] = Pr[\theta_2 > \hat{\theta}_2(\theta_1, s)](k - cs)$$

The first order condition of bank's optimal action $s^\star$ trades off the higher cost of risk management with an increased probability of survival. Lemma 2 establishes

[8] Following the literature, we will refer to (variants of) $\hat{\theta}_2$ as 'fundamental threshold', while $\hat{x}$ as 'strategic threshold'.

optimal risk management action $s^\star(\theta_1)$ and the associated payoff.

Lemma 2 *When type θ_1 is public information, a bank of type θ_1 optimally sets $s^\star(\theta_1)$ and obtains payoff $\pi^\star(\theta_1)$ as defined below:*

$$s^\star(\theta_1) = \frac{1}{2}\left(\frac{k}{c} - \frac{\overline{\theta}_2 - t}{\theta_1}\right) \tag{2.3}$$

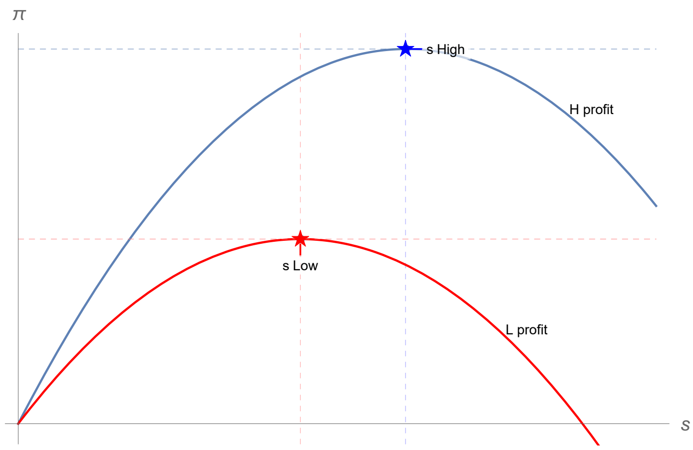
$$\pi^\star(\theta_1) = \frac{\left(c\left(\overline{\theta}_2 - t\right) + \theta_1 k\right)^2}{4c\theta_1} \tag{2.4}$$

Provided that $\frac{k}{c} > \frac{\overline{\theta}_2 - t}{\theta_1}$, both the optimal action and the optimal expected profit increase in parameter θ_1.

Proof. See Appendix ■

Figure 2.2 illustrates expected profits as a function of the risk management action s. The optimal s under complete information is higher for the H-type, which is a direct consequence of the higher marginal benefit of action. The condition in Lemma 2 is a necessary and sufficient condition for the positivity of the optimal action, which we will assume in order for the problem to be interesting.

Figure 2.2: Expected profit as a function of risk management action



Expected profit is increasing in type, and the optimal risk management action satisfies $s^\star_L < s^\star_H$

2.3.2 Separating equilibrium

Moving towards analysing asymmetric information, we first characterize separating equilibrium. In any pure-strategy separating equilibrium the two types send different signals

$$s(\theta_1) := \begin{cases} s_L & \text{if} \quad \theta_1 = \theta_1^L \\ s_H & \text{if} \quad \theta_1 = \theta_1^H \end{cases}$$

and the chosen signal reveals the type perfectly to creditors. Separating equilibrium can be maintained if no sender has incentives to deviate. Before formalizing the incentive compatibility constraints, the first step is to compute payoffs off-the-equilibrium path.

2.3.2.1 Off-equilibrium payoffs

First, consider the case where a bank of type L chooses an off-equilibrium action and mimics type H by sending the signal s_H. The creditors - believing that they are facing a H-type bank - behave accordingly as if they were facing H-type with certainty. Therefore, their optimal response given these beliefs is described by the strategic threshold defined under Lemma 1 Equation 2.1 for type H:

$$\hat{x}^H := \hat{x}(\theta_1^H, s_H) = t - \theta_1^H s_H + 2\sigma t - \sigma$$

This implies that the mass of agents who would run upon any realization of θ_2 is exactly the same as if it is under type H, that is, $\alpha(\hat{x}^H, \theta_2)$[9]. The off-equilibrium fundamental threshold for the deviating type, denoted by $\hat{\theta}_2^{L.H}$, is the value of θ_2 which solves

$$\theta_1^L s_H + \theta_2 - \alpha(\hat{x}^H, \theta_2) = 0$$

which implies, after substituting the expression for α and rearranging terms:

$$\hat{\theta}_2^{L.H} = \frac{\hat{x}^H + \sigma - 2\sigma s_H \theta_1^L}{2\sigma + 1}$$

After substituting $\hat{x}^H$, defining the type difference $\Delta\theta_1 = (\theta_1^H - \theta_1^L)$, and introducing the notation $\hat{\theta}_2^L := \hat{\theta}_2(\theta_1^L)$ and $\hat{\theta}_2^H := \hat{\theta}_2(\theta_1^H)$, we obtain

$$\hat{\theta}_2^{L.H} = \hat{\theta}_2^H + \frac{2\sigma s_H \Delta\theta_1}{1 + 2\sigma} = \hat{\theta}_2^L - \frac{s_H \Delta\theta_1}{1 + 2\sigma} \tag{2.5}$$

[9] The expression for that can be found in the proof of Lemma 1, Equation 2.3.

Analogously, by replacing indices but keeping the definition of $\Delta\theta_1 = (\theta_1^H - \theta_1^L)$ fixed, it is possible to define off-equilibrium thresholds for the case when H mimics L[10]

$$\hat{\theta}_2^{H.L} = \hat{\theta}_2^L - \frac{2\sigma s_L \Delta\theta_1}{1+2\sigma} = \hat{\theta}_2^H + \frac{s_L \Delta\theta_1}{1+2\sigma} \tag{2.6}$$

To conclude, by mimicking the other type's action in an off-the-equilibrium path of the candidate separating equilibrium, the bank can influence the behaviour of creditors and induce them to behave according to the strategy what they would follow under the other type. However, he cannot achieve the same fundamental threshold, since the true type enters directly into the regime change function $\mathcal{R}$, which determines the threshold.

Before characterizing the equilibrium, we discuss an alternative interpretation of the off-equilibrium thresholds (2.5) and (2.6). For any given (not necessarily equilibrium) s, the *functions* $\hat{\theta}_2^{L.H}(s)$ and $\hat{\theta}_2^{H.L}(s)$ can be understood as failure thresholds for type L (respectively H), if its creditors believe it to be the other type. These functions define an additive decomposition of the difference between the two types' *complete information* fundamental thresholds for a given s. For example, using Equation (2.5), we can write

$$\hat{\theta}_2^L(s) - \hat{\theta}_2^H(s) = \underbrace{\frac{2\sigma s \Delta\theta_1}{1+2\sigma}}_{\text{Direct effect}} + \underbrace{\frac{s\Delta\theta_1}{1+2\sigma}}_{\text{Indirect effect}} \tag{2.7}$$

Equation 2.7 decomposes the difference between the fundamental thresholds under complete information into a sum of a *direct effect*, attributable to the fundamental difference between types L and H, and an *indirect effect*, which is solely due to creditors' beliefs. As we show in the next section, the larger the indirect effect is, the more a low-type can potentially benefit from mimicking the high type, and similarly, the larger is the potential loss for a high type for not being able to distinguish himself from a low type.

If type L is believed to be[11] type H at *some* signal s, she obtains the following expected payoff:

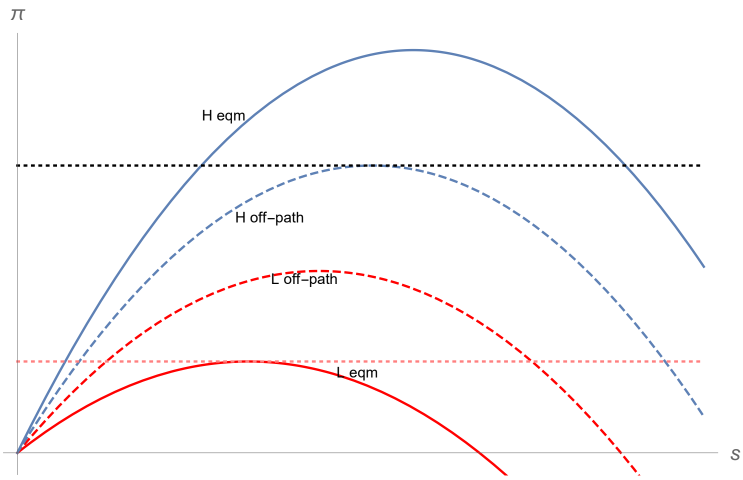
$$\pi(s, \underbrace{\theta_1^L}_{\text{true}}, \underbrace{\theta_1^H}_{\text{perceived}}) = Pr\left[\theta_2 > \hat{\theta}_2^{L.H}\right](k - c\cdot s) \tag{2.8}$$

The optimal off-equilibrium action (that is, optimal action if type L is believed to

[10] As will be clear in the equilibrium analysis, H would never want to mimic L. Yet, he might find it optimal to go 'off-path', in which case his payoffs are characterized by $\hat{\theta}_2^{H.L}$

[11] that happens if she successfully mimics type H

Figure 2.3: On- and off-path equilibrium payoffs



be type H) is derived from the first-order condition $\partial\pi/\partial s = 0$, which implies the optimum off-equilibrium intervention and expected profits:

$$s^\star_{L.H} = \frac{1}{2}\left(\frac{k}{c} + \frac{(1+2\sigma)(t-\overline{\theta}_2)}{2\sigma\theta_1^L + \theta_1^H}\right) = 0 \tag{2.9}$$

$$\pi^\star_{L.H} = \pi(s^\star_{L.H}, \theta_1^L, \theta_1^H) = \frac{\left(-c(2\sigma+1)\left(t-\overline{\theta}_2\right) + k(2\sigma\theta_1^L + \theta_1^H)\right)^2}{4c(2\sigma+1)\left(2\sigma\theta_1^L + \theta_1^H\right)} \tag{2.10}$$

Analogous expressions can be derived for $s^\star_{H.L}$ and $\pi^\star_{H.L}$. It is straightforward to show that the following relationships hold:

$$s^\star_L < s^\star_{L.H} \qquad \text{and} \qquad s^\star_H > s^\star_{H.L} \tag{2.11}$$

$$\pi(s,\theta_1^L,\theta_1^L) < \pi(s,\theta_1^L,\theta_1^H) \qquad \text{and} \qquad \pi(s,\theta_1^H,\theta_1^H) > \pi(s,\theta_1^H,\theta_1^L) \quad \forall s \tag{2.12}$$

Off-equilibrium payoffs and optimal actions are critical in analysing the existence of equilibrium. In particular, the profit $\pi^\star_{H.L}$ is H-type's minimax payoff: even with the most adverse beliefs of creditors (if all believe he is of bad type), he can obtain payoff at least $\pi^\star_{H.L}$. Therefore, in any proposed equilibrium, type H's payoff must exceed $\pi^\star_{H.L}$. Note that in contrast, the low type's (L) minimax payoff is $\pi^\star_L$.

Figure 2.3 illustrates on- and off-equilibrium payoffs as a function of an arbitrary policy intervention s. In the next section we establish the values of s which can be maintained as separating equilibrium.

2.3.2.2 Characterization of separating equilibrium

Our equilibrium concept does not place any restrictions on the beliefs off-the-equilibrium path, which are never reached to verify those beliefs. Following the standard signalling literature, we impose the following belief system: if creditors observe any off-equilibrium risk management action $s \neq \{s_L, s_H\}$, they believe that the regime is of low type. Otherwise, if they observe an equilibrium action, they believe that they are facing with the appropriate type with certainty. Under this specification, every equilibrium signal which gives at least as much profit *to both types* as they would get under 'low type' beliefs can be maintained.

Equilibrium requires that no types have incentives to deviate from the proposed equilibrium actions $\{s_L, s_H\}$. First, note that in any separating equilibrium, types are revealed, so the L-type will find it optimal to set $s_L = s_L^\star$, and obtain profit $\pi_L^\star$. A separating equilibrium in which the high-type sets some value s_H and the low type sets her optimum value $s_L^\star$ can be maintained if and only if

$$\pi^{eq}(s_L^\star, \theta_1^L) \geq \pi^{off}(s_H, \theta_1^L, \theta_1^H) \tag{IC_L}$$

$$\pi^{eq}(s_H, \theta_1^H) \geq \pi^{off}(s_{H.L}^\star, \theta_1^H, \theta_1^L) \tag{IC_H}$$

where π^{eq} is equilibrium payoff given the specified belief system, while $s_{H.L}^\star$ is the best deviation for the high-type, specified in the previous section. Let us denote by s_L^{cri} the value of s_H which solves $[IC_L]$, that is, the value of a separating signal at which the L-type is just indifferent between mimicking the high type, or setting $s_L^\star$ and obtaining her minimax profit. This is the value of s which solves

$$Pr\left[\theta_2 > \hat{\theta}_2^L(s_L^\star)\right](k - c \cdot s_L^\star) \geq Pr\left[\theta_2 > \hat{\theta}_2^{L.H}(s)\right](k - c \cdot s)$$

The right-hand-side is a quadratic function with a negative coefficient of the quadratic term and with maximum value exceeding the constant on the left-hand-side, so the corresponding equality has two solutions $\left(s_{L,1}^{cri} < s_{L,2}^{cri}\right)$. Incentive compatibility requires that $s_H \notin [s_{L,1}^{cri}, s_{L,2}^{cri}]$, otherwise L-type would have an incentive to mimic the H-type. In this case $\pi^{eq}(s_{L.1}^{cri}, \theta_1^H) < \pi^{eq}(s_{L.2}^{cri}, \theta_1^H)$ implying that the individually rational choice for the good type is to send a high signal, and the level which can maintain a separating equilibrium must fulfil $s_H \geq s_{L.2}^{cri}$.

Similarly, define s_H^{cri} to be the critical s which is incentive-compatible for type H and solves $[IC_H]$. This is the level of intervention at which the profit for a H-type in a separating equilibrium is at least as much as his best achievable profit if he is believed to be of low-type. This latter utility is the high-types' *minimax* *payoff* - irrespectively of creditors' beliefs, he can always achieve at least $\pi^{\star}_{H.L}$ by setting the off-equilibrium profit-maximizing level of s.[12] After substituting the profit functions, $[IC_H]$ leads to:

$$Pr\left[\theta_2 > \hat{\theta}_2^H(s^{\star}_H)\right](k - c \cdot s_H) \geq Pr\left[\theta_2 > \hat{\theta}_2^{H.L}(s^{\star}_{H.L})\right](k - c \cdot s^{\star}_{H.L})$$

which, by similar argument,, has two solutions: $s^{cri}_{H.1}$ and $s^{cri}_{H.2}$. We can characterize the existence of a separating equilibrium in terms of the thresholds derived above as follows.[13]

1. A separating equilibrium exists and it restores the symmetric information benchmark if and only if $s^{cri}_{L.2} \leq s^{\star}_H$.
2. A separating equilibrium exists and in this equilibrium the high-type sets inefficiently high risk management action if and only if $s^{\star}_H < s^{cri}_{L.2} \leq s^{cri}_{H.2}$.
3. A separating equilibrium does not exist if and only if $s^{cri}_{H.2} < s^{cri}_{L.2}$.

We derive closed form analytical formulas for the critical values in the Appendix along with some limiting cases, which we will use in the following discussion. Theorem 1 establishes the link between the cost of risk management, precision of private signals, and the existence of separating equilibrium.

Theorem 1 *There exists an 'efficient' separating equilibrium in which signals coincide with the symmetric information benchmark if and only if the noise in creditors' private observation is sufficiently large, that is, if and only if*

$$\sigma \geq \overline{\sigma}$$

Whenever $\sigma < \overline{\sigma}$, and $c > \hat{c}$, there exists an 'inefficient' separating equilibrium in which the bank must choose a higher-than-the-first-best risk management intervention, where $\hat{c}$ is defined as

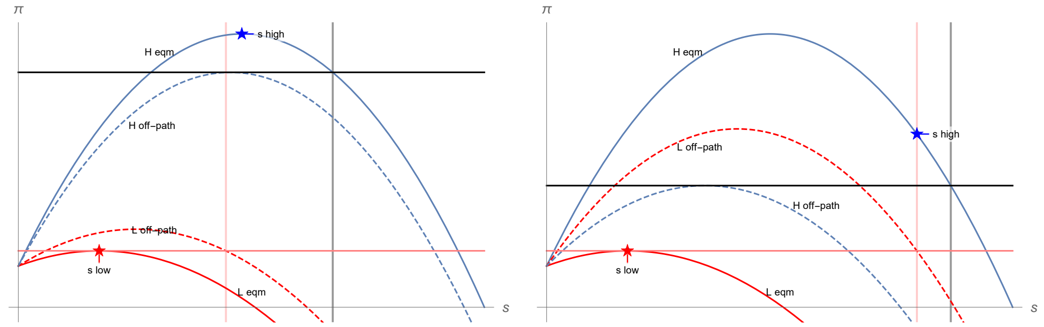
$$\hat{c} = \sqrt{\frac{\Delta\theta_1(k\theta_1^L)^2}{\left(3\theta_1^L + \theta_1^H\right)\left(\overline{\theta}_2 - t\right)}} \tag{2.13}$$

Whenever $c < \hat{c}$, there exists a lower boundary $\underline{\sigma}(c)$ such that separating equilibrium does not exists for every $\sigma < \underline{\sigma}(c)$.

[12]Note the difference here: L-type's minimax payoff is $\pi^{\star}_L$ while type H's minimax payoff is $\pi^{\star}_{H.L}$ under the specified beliefs.

[13]We concentrate only on the 'upper' regions, which is relevant for our application. It is straightforward to extend the analysis to the lower part.

Figure 2.4: Separating Equilibrium



(a) The first-best is restored in Separating Equilibrium with high σ (low precision)

(b) High-type sends inefficiently high intervention with low σ (high precision)

Proof. See Appendix ■

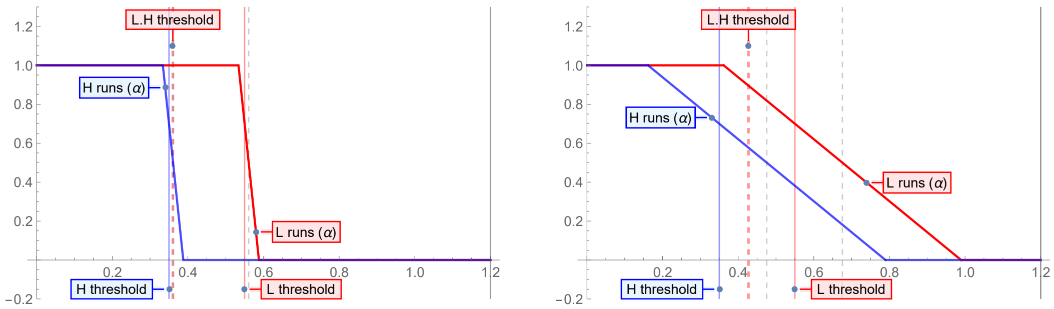
The Theorem is illustrated in Figure 2.4. The precision of creditor's signal $(1/\sigma)$ has a critical role in determining which type of equilibrium can survive. If information is less precise, the potential benefit/loss from mimicking the other type decreases. This is because if type L mimics type H, the (off-the-equilibrium-path) strategic threshold $\hat{x}$ is pinned down according to the equilibrium of type H, but the fundamental threshold, which enters directly into the integral boundary of the expected profit, is not.[14] More noisy private information pushes the realized fundamental threshold upwards, which decreases the profit which can be obtained by mimicking type H. This, in turn, decreases the critical threshold of low-types incentive compatibility constraint, $s^{cri}_{L.2}$, which approaches $s^\star_L$ as $\sigma \to \infty$. Since $\lim s^{cri}_{L.2} < s^\star_H$, due to continuity there exists an σ (denoted by $\overline{\sigma}$) at which $s^{cri}_{L.2}(\overline{\sigma}) = s^\star_H$. Consequently, if and only if the fundamental is observed with large enough noise $(\sigma > \overline{\sigma})$ the separating equilibrium is efficient (i.e. restores bank-optimal first-best). If the equilibrium is inefficient, the distortion increases as the noise becomes more precise.

Intuitively, the more precise receivers' private observation is, the more correlated is creditors' behaviour. This implies, with higher precision the *effective strategy*, defined as the probability of run for any realization of fundamental θ_2, is more 'extreme' (see Figure 2.5). In particular, with $\sigma \to 0$, the effective strategy converges to a limiting case where all creditors run if and only if $x_i \leq \hat{\theta}_1$, and no

[14] A way to think about this is: in the standard global game, the fundamental threshold is fixed when precision in varied, and the strategic threshold adapts to the changes. In contrast, when L mimics H, the strategic threshold is fixed and the realized fundamental threshold varies. When the strategic threshold of H shifts to the left due to increased noise, the fundamental threshold must shift to the left as well.

agents attack otherwise (2.5a). However, the more extreme is the effective strategy, the more important it is to 'get the other parameters right'. The aggregated *strategic error* by following a certain strategy which happens to be wrong is largest when the information which determines the strategy is the most precise. On the other hand, with lower precision the effective strategy is more 'flat', and the effect of not knowing the other parameter (θ_1) correctly is smoothed out by the relative flatness of the effective strategy (2.5b). This decreases the potential benefit of mimicking the H-type.

Figure 2.5: Equilibrium determination as a function of noise



(a) High precision leads to highly coordinated behaviour...

(b) ...while low precision maintains uncertainty regarding creditor's behaviour.

2.3.3 Regulation in separating equilibrium

In this section we show how a *minimum threshold* regulation can eliminate separating equilibria. Let us denote the regulatory minimum by s_p. Under a binding risk management regulatory regime the action space of the bank is restricted to the interval $s \in [s_p, \infty)$.

A minimum requirement changes minimax payoffs for both types. We will assume that minimum policy is high enough to be binding for both types in the sense that it exceeds their minimax strategies, i.e.

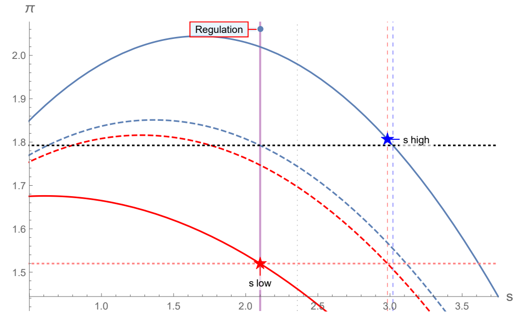
$$s_p \geq s^\star_{H.L}$$

This single condition is sufficient, since $s^\star_L < s^\star_{H.L}$, so the constraint will always be binding for L-type. Then, we can reformulate IC's for a separating equilibrium as

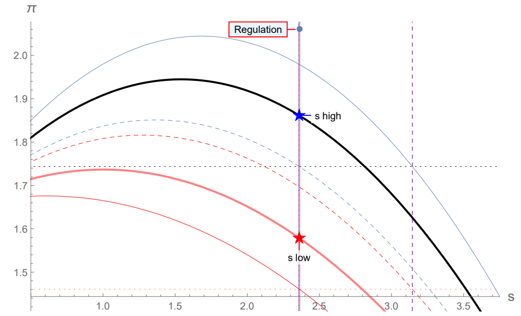
$$\pi^{eq}(s_p, \theta_1^L) \geq \pi^{off}(s_H, \theta_1^L, \theta_1^H) \qquad (IC_L)$$

$$\pi^{eq}(s_H, \theta_1^H) \geq \pi^{off}(s_p, \theta_1^H, \theta_1^L) \qquad (IC_H)$$

Figure 2.6: Impact of regulation on the equilibrium



(a) Regulation is ineffective: separating equilibrium survives



(b) Regulation is effective: separating equilibrium fails

where s_H denotes the signal an H-type sends in a separating equilibrium. Similarly to the previous (no-regulation) case, we can define $s_L^{cri}(s_p)$ and $s_H^{cri}(s_p)$ as the value of s which solves the two IC's respectively with equality, now both regarded as a function of regulatory threshold s_p. Then, the critical regulatory threshold level which guarantees that separating equilibrium does not exist is determined by the equation:

$$s_L^{cri}(s_p) = s_H^{cri}(s_p) \tag{2.14}$$

Theorem 2 *As long as the regulatory minimum policy s_p exceeds a critical regulatory minimum s_p^{cri} as defined by equality 2.14, that is, $s_p \geq s_p^{cri}$, separating equilibrium does not exist. The critical regulatory level is defined as follows:*

$$s_p^{cri} = \left(\frac{(1+2\sigma)(\bar{\theta}_2 - t)}{\Delta\theta_1} + \frac{k}{2c} \right) - \sqrt{\left[\frac{(1+2\sigma)(\bar{\theta}_2 - t)}{\Delta\theta_1} \right]^2 + \left[\frac{k}{2c} \right]^2} \tag{2.15}$$

The pooling equilibrium in which all types of banks set s_p^{cri} Pareto-dominates from the banks' perspective all other pooling equilibria.

Proof. See Appendix ■

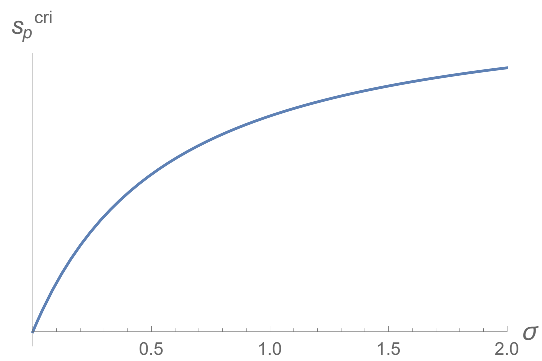
Figure 2.6a depicts a situation where, despite a quantitative minimum regulation for policy being in place, a separating equilibrium still survives, as s_L^{cri} (the policy which is just incentive-compatible for the low type) is lower than s_H^{cri} (just incentive-compatible for the high type). The regulator must increase the minimum policy to at least s_p^{cri}, where $s_L^{cri} = s_H^{cri}$, so no separating equilibrium exists anymore (figure 2.6b).

Now we state a result which has interesting implications for the effect of regulation during a crisis situation.

Corollary 1 *The critical regulatory threshold decreases in precision* $1/\sigma$

The intuition behind this result is straightforward. We have seen that with high precision, the inefficiency in a separating equilibrium is very high as high-quality institutions are sending excessively high signals to distinguish themselves from low-quality institutions. However, this is exactly the situation when it is relatively easy for a regulator to squeeze out the separating equilibrium by setting a relatively low pooling threshold. Since the incentive compatibility constraints for the H-type are already close to binding, a little bit more pressure induced by the regulator can be sufficient to break down separation. Figure 2.7 illustrates the level of critical regulatory minimum as a function of noise (σ) in private information.

Figure 2.7: Critical regulation



Critical regulatory threshold increases in the noise of private information

It is possible to interpret this result in the context of the cyclicality of banking regulations. Noisy private signals are consequences of turbulent economic periods, as increased uncertainty amplifies strategic uncertainty and information asymmetries among creditors. In this case, as we discussed before, there is relatively little to gain from mimicking the other type, so it is *easier* to maintain a separating equilibrium. On the other hand this means that a regulator must maintain a relatively strict minimum policy if he wants to squeeze out separation. In contrast, in normal times - represented by a small idiosyncratic noise in our model - there is more temptation to mimic high-types, leading to highly inefficient separating signal levels. An already high risk management signal is however relatively easy to squeeze out. This effect can explain why - observationally - signalling seems to be more prevalent during turbulent times. A regulation in place which is just sufficient to impose pooling in a normal market environment, might not be sufficient to achieve the same during turbulent times, when strategic uncertainty is greater, so financial institutions engage more and more in costly signalling as the markets shift towards a crisis period.

2.3.4 Pooling equilibrium

In any pooling equilibrium the same signal is chosen by both types of banks, which conveys no information to the creditors. Equilibrium of the coordination stage is determined analogously to a standard global game with an important twist: given creditors' strategy, represented by the threshold $\hat{x}$, which must be the same under both types of banks in a pooling equilibrium, the fundamental threshold for the two types will be different. This has to be taken into consideration by the creditors when calculating equilibrium strategies. In conclusion, any pooling equilibrium is characterized by a common strategic threshold for creditors $\hat{x}$, and a distinct fundamental threshold for each type of banks, $\hat{\theta}_2^L \neq \hat{\theta}_2^H$, such that (i) creditors run if and only if $x_i \leq \hat{x}$, and (ii) a bank of type L (resp. H) fails if and only if $\theta_2 \leq \hat{\theta}_2^L$ (resp. $\leq \hat{\theta}_2^H$). The equations determining the equilibrium of the global game change accordingly: (i) a creditor who receives private signal $\hat{x}$ should be just indifferent between actions RUN and WAIT, given that banks of type $\{L, H\}$ fails if and only if the fundamental θ_2 is below the respective threshold and the (posterior) beliefs are (p_L, p_H), and (ii) a bank of type L(resp. H) fails exactly at $\hat{\theta}_2^L$(resp. $\hat{\theta}_2^H$) if creditors run if and only if $x_i < \hat{x}$. Pooling equilibrium thresholds are characterized by Lemma 3.

Lemma 3 *In any pooling equilibrium where banks follow the same risk management strategy s_p, the equilibrium of the stage 2 subgame (coordination stage) is characterized by fundamental thresholds $\hat{\theta}_2^{L,P}$ and $\hat{\theta}_2^{H,P}$ and strategic threshold $\hat{x}$ where*

$$\hat{\theta}_2^{L,P}(s_p) = t - \frac{s_p\overline{\theta}_1}{1+2\sigma} - \frac{2\sigma s_p \theta_1^L}{1+2\sigma}$$
$$\hat{\theta}_2^{H,P}(s_p) = t - \frac{s_p\overline{\theta}_1}{1+2\sigma} - \frac{2\sigma s_p \theta_1^H}{1+2\sigma}$$

and

$$\hat{x} = 2\sigma t - \sigma + p_L\hat{\theta}_2^{L,P} + p_H\hat{\theta}_2^{H,P}$$

Proof. See Appendix ■

From Lemma 3, the thresholds can be rewritten as

$$\hat{\theta}_2^{L,P}(s) = \hat{\theta}_2^{L,FI}(s) - \frac{p_H \Delta\theta_1}{1+2\sigma}s$$
$$\hat{\theta}_2^{H,P}(s) = \hat{\theta}_2^{H,FI}(s) + \frac{p_L \Delta\theta_1}{1+2\sigma}s$$

where ^{FI} index stands for the full-information threshold. Some consequences can be seen immediately. First, for every s we obtain that $\hat{\theta}_1^{H,FI}(s) < \hat{\theta}_2^{H,P}(s) < \hat{\theta}_2^{L,P}(s) < \hat{\theta}_1^{L,FI}(s)$. In this sense, pooling among the two types of banks implements a cross-subsidy across types, as for any given level of risk management, increases L-type's, while decreases H-type's payoff. Second, as the noise of private information increases, the pooling thresholds continuously approach the full-information thresholds, so the cross-subsidy effect of pooling decreases. Intuitively, similarly to the case with separating equilibrium, with an increasing noise in the private signals, fundamental thresholds become less-and-less responsive to other parameters of the game, dampening the effect of the shift in the strategic threshold. Finally, as the ex-ante percentage of low (high) types increases, the strategic threshold in pooling approaches the full information low (high) threshold.

The following Corollary will turn out to be useful to characterize welfare effects:

Corollary 2 *The average fundamental threshold in any pooling equilibrium is a linear function of the average type $\overline{\theta_1} = p_L\theta_1^L + p_H\theta_1^H$. Precisely,*

$$\overline{\theta}_2^P(s) = t - \frac{s(p_L\theta_1^L) + p_H\theta_1^H)}{1+2\sigma} - p_L\frac{2\sigma s\theta_1^L}{1+2\sigma} - p_H\frac{2\sigma s\theta_1^H}{1+2\sigma} = t - \overline{\theta_1}s$$

Now we turn to the question of which risk management actions can be maintained in a pooling equilibrium. The complication arises from the fact that the equilibrium concept we used so far does not place any restrictions on the beliefs off-the-equilibrium path, which are never reached to verify those beliefs. First, we assume that agents' beliefs are characterized as follows:

- *(Equilibrium path)* If agents observe the pooling level intervention, s_p they play according to respective coordination game, as defined above;
- *(Off-the-equilibrium path)* if agents observe any other intervention $s \neq s_p$ they believe that the regime is of low type.

This belief system is often used in the signalling literature as a benchmark. Under this specification, every level of risk management which gives at least as much profit *to both types* as they would get under 'low type' beliefs can be maintained. Since the privately optimal levels for the two types are different, there is no Pareto-dominant pooling equilibrium.[15] However, any policy intervention $s_p \notin \left[s_p^{L\star}, s_p^{H\star}\right]$ is Pareto-dominated by some intervention level $s_p \in \left[s_p^{L\star}, s_p^{H\star}\right]$.

Banks may be pooling on various values of action s, which potentially leads to a continuum of pooling equilibria. This set of equilibria is limited by the usual incentive compatibility constraints: in any proposed equilibrium, both types should get a higher payoff than in any off-equilibrium path. Because according to the specified belief system, on any off-equilibrium path banks are perceived to be low-types, their best deviation is to implement the optimal action, given that the perceived type is 'low'. Thus, any candidate equilibrium s_p is such that

$$\pi^{\text{off}}\left(s_L^\star; \theta_1^L, \theta_1^L\right) \leq \pi^{\text{peq}}(s_p; \theta_1^L) \tag{2.16}$$

$$\pi^{\text{off}}\left(s_{H.L}^\star, \theta_1^H, \theta_1^L\right) \leq \pi^{\text{peq}}(s_p, \theta_1^H) \tag{2.17}$$

These incentive compatibility constraints select a critical value — denoted by s_p^{max} — as a maximum incentive-compatible signal in a pooling equilibrium[16].

Theorem 3 *There exist a pooling equilibrium where both types of banks send a signal s_p for every $s_p \in [s_\star, s^\star]$, where the critical values solve the incentive compatibility constraints in Equation 2.16 and 2.17.*

Proof. Follows from above. ■

2.4 Payoffs and incentive-compatible regulations

In this section we analyse ex-ante expected payoffs for the bank and the creditors separately, and draw some conclusions regarding the welfare effects of regulation. For the ease of exposition let's denote any equilibrium as $Q := \{s_L, s_H, \hat{\theta}_2^L, \hat{\theta}_2^H, \hat{x}^L, \hat{x}^H\}$, where $\{s_L, s_H\}$ are equilibrium first-stage strategies of a bank of type L, H, $\{\hat{x}^L, \hat{x}^H\}$ are strategic thresholds and $\{\hat{\theta}_2^L, \hat{\theta}_2^H, \}$ are fundamental thresholds of type L, H.[17]

[15] This is in contrast to, for example, a simple Spence-model, where zero education by both types Pareto-dominates all other pooling equilibria.

[16] Precisely, they will select also an s_p^{min} and for the lower boundary the first IC may be binding, but the upper boundary is the interesting one for our application.

[17] If Q is a pooling equilibrium, $\{s_L = s_H\}$ and $\{\hat{x}_L = \hat{x}_H\}$

2.4.1 Bank's expected payoff

Let π_L^Q (resp. π_H^Q) denote the low (high) type's ex-post *expected* payoff in any equilibrium Q, and p the probability mass of low-type banks. Then the expected payoff for a bank from ex-ante point of view (i.e. before learning his type) can be formulated as:

$$\mathbb{E}\left[\pi^Q\right] = p\pi_L^Q + (1-p)\pi_H^Q \tag{2.18}$$

Recall that (irrespectively how the threshold and in turn the probability of survival is calculated) the bank's payoff can be written generally for $\tau \in \{L, H\}$ as

$$\pi_\tau^Q = \frac{1}{2\eta}\left(\overline{\theta}_2 - \hat{\theta}_2^{\tau,Q}(s_\tau^Q)\right)\left(k - cs_\tau^Q\right)$$

where the variables $\hat{\theta}_2^{\tau,Q}$ and s_τ^Q are the equilibrium-Q values for type τ.

Pooling equilibrium [PE]: In any PE $s_\tau^Q = s_p$ for all τ, so for an arbitrary s_p,

$$\begin{aligned}\mathbb{E}\pi^{PE} &= p\pi^L(s_p) + (1-p)\pi^H(s_p) = \pi(\theta_1^{AV}, s_p)\\ &= \frac{1}{2\eta}\left(\overline{\theta}_2 - \hat{\theta}_2(\theta_1^{AV}, s_p)\right)(k - cs_p)\end{aligned}$$

The simple formula facilitates easy comparative statics: the ex-ante expected payoff in pooling is *not* a function of noise, and is decreasing in p [18].

Separating equilibrium [SE]: We focus only on the least-costly SE, which is well defined by exogenous parameters. Recall that in this equilibrium the separating action by the low-type is the optimum value according to the symmetric information benchmark $s_L^\star$, and she obtains her minimax profit, while the high-type sets the appropriate critical value $s_{H.2}^{cri}$ and obtains a profit which is (weakly) less than the benchmark value. The ex-ante expected payoff to the bank is therefore always (weakly) below the symmetric information value, and the welfare loss relative to this benchmark is concentrated on the high-type. Formally, we have:

$$\mathbb{E}\pi^{SE} = p\pi_L(s_L^\star) + (1-p)\pi_H(s_{H.2}^{cri}),$$

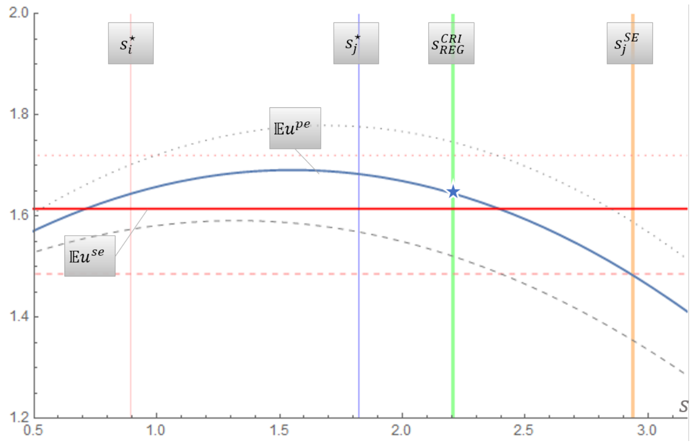
From ex-ante point of view the bank prefers pooling over separation if and only if

$$\mathbb{E}\pi^{PE}(s_p) > \mathbb{E}\pi^{SE} \tag{2.19}$$

This can be calculated analytically, however, is tedious due to the complicated formula for separating equilibrium expected payoffs, and we resort to numerical

[18] The derivative is $\frac{\partial}{\partial p} = -\frac{1}{2\eta}(k - cs_p)s_p\Delta\theta_1 < 0$

Figure 2.8: Illustration of banks' ex-ante expected payoff



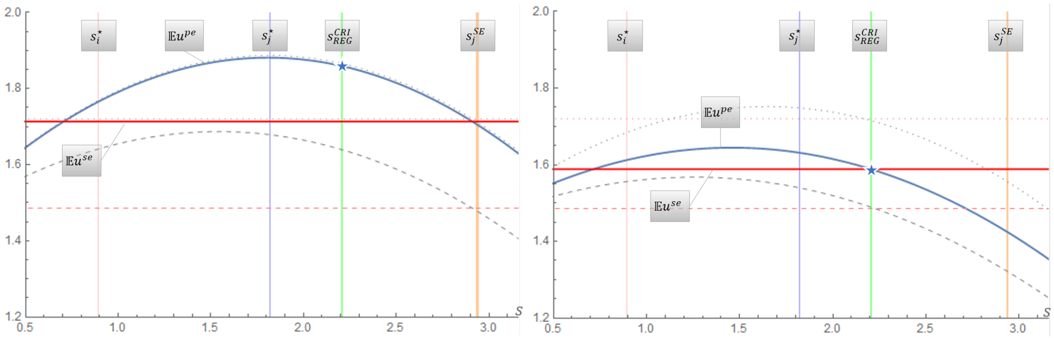
Expected payoff is larger in pooling then in separation from ex-ante point of view, and for the both types separately (dashed and dotted lines)

analysis. The ex-ante expected pooling profit is a concave function of s_p (blue curve in Figure 2.8). As long as the *optimal pooling*, derived analytically in 2.A.6.1, exceeds profits from separating equilibrium (red), there exists a compact interval of risk management actions at which pooling is ex-ante preferred to separation. We focus on the upper threshold of this region which we denote by notation s_p^{cri}. If this threshold exists, pooling is preferred by the banks from ex-ante point of view for all $s_p < s_p^{cri}$.

In the context of the welfare impact of regulations, this implies that pooling is preferred by banks whenever the regulatory requirement required to induce pooling is *not too large*. From the banks' point of view, the optimal regulatory threshold would be the one which is *just sufficient* to squeeze out incentives for inefficient signalling, but not too large yet to impose an extra burden for both types of banks in the form of a too restrictive regulation. Notice that as can be seen from the figure, similar conclusions follow not just from ex-ante point of view (when types are not known), but after the realization, for the two types separately. The high-type bank benefits from the reduction of costly signalling, why the low-type benefits from the cross-subsidy effect in the resulting pooling equilibrium.

The critical value s_p^{cri} strictly decreases in p. In particular, in the limit where $p \to 0$ (almost all banks are H-types), $s_p^{cri} = s_{H.2}^{cri}$, that is, the critical regulatory level equals the risk management level in a least-costly separating equilibrium. This means *every* pooling below this value is welfare-improving (Figure 2.9a). It might be useful to note that the least-costly separating equilibrium does *not* depend on

Figure 2.9: Illustration of banks' ex-ante expected payoff:
Limiting cases



(a) Banks' ex-ante payoff as $p \to 0$ (b) Banks' ex-ante payoff as $p \to p^\star$

the type distribution: maintaining separation is just as costly if there is only one single low-type bank as if there are many[19]. Consequently, being in a separating equilibrium is more 'inefficient', therefore more costly *overall*, when the mass of low-type banks is low (small p).

In the other limit, as $p \to 1$ the critical value s_p^{cri} converges to the low-type's minimax payoff, and pooling is never welfare-improving. The intuitive reason is that pooling is not particularly desirable for either types when very large number of low-type banks are present, as low-types have little to gain while high types have much to loose. Using continuity arguments it is possible to show the existence of a critical probability level $p^\star$ such that a (critical) regulatory threshold which just imposes pooling improves ex-ante payoff if and only if $p < p^\star$ (this situation is depicted in Figure 2.9b). Formally, this is the value of p which solves expression (2.19) with equality.

Finally, the critical value s_p^{cri} strictly decreases in the creditors' private noise σ, and this implies the existence of a critical noise σ^{cri} such that pooling is beneficial if and only if $\sigma < \sigma^{cri}$. Intuitively, pooling is more likely to be preferred by banks if the noise in creditors' private information is not too large, because in this case maintaining separation is relatively costly, just as discussed before in more details.

To sum up the intuition, we can conclude the followings: (i) ex ante, banks *may* (but do not necessarily) prefer pooling on not too large risk management actions. This prevents them from sending inefficiently high signals in a separating equilibrium, and boosts payoff due to cross-subsidy. (ii) Lower noise / higher precision increases the costs of inefficient separation for the high-type. This makes banks

[19] The difference compared to a hypothetical payoff in a pooling equilibrium, however, is not distribution-independent.

to like pooling even more, and to prefer it over separation at even higher (regulatory) levels. In addition, lower noise in private information pushes down the minimum threshold which is required to force out separating equilibria. (iii) With more high type banks in the population, costly signalling happens more often, therefore banks prefer pooling even more from ex-ante point of view.

2.4.2 Creditors' expected payoff

The creditor's payoff is t if she withdraws, 1 if the bank survives and she stays, and 0 in case of bankruptcy. The expected payoff reflects that (i) higher risk management action has a fundamental stabilizing role, so ceteris paribus creditors would always prefer higher signals; (ii) when comparing pooling with separation, the benefits of high-types' (excessive) risk management signals are weighed against low-type's lower action.

Separating equilibrium: In Appendix we derive creditors' ex-ante expected payoff analytically. Because in any separating equilibrium types are perfectly revealed, the ex-ante payoff is the probabilty-weighted average of the symmetric information benchmark payoffs using the prior probability distribution (p) as weights.

Lemma 4 *The (conditional) expected payoff to the creditors of a bank with known (or correctly deduced) type $\tau \in \{L, H\}$ who exert an arbitrary risk management action s is*

$$\mathbb{E}_1[u|\tau] = \frac{1}{2\eta}\left((\overline{\theta}_2 - t\underline{\theta}_2) - \sigma t(1-t) - (t - \theta_1^\tau s)(1-t)\right)$$

, where $\underline{\theta}_2 = \overline{\theta}_2 - 2\eta$.

Proof. See Appendix ■

We are particularly interested in the unique, *least-costly separating equilibrium*, in which the risk management signals send by low (high) type respectively are given by $s_L^\star$ and $s_{H.2}^{cri}$. The ex-ante expected payoff in the least-costly separating equilibrium is

$$\begin{aligned}\mathbb{E}_0 u &= p\mathbb{E}_1[u(s_L^\star)|L] + (1-p)\mathbb{E}_1[u(s_{H.2}^{cri})|H] \\ &= \frac{1}{2\eta}\left((\overline{\theta}_2 - t\underline{\theta}_2) - \sigma t(1-t) - (1-t)\left(t - p\theta_1^L s_L^\star - (1-p)\theta_1^H s_{H.2}^{cri}\right)\right)\end{aligned}$$

Ceteris paribus creditors prefer higher risk management action because of their stabilizing role. In a separating equilibrium the high-type sends higher signals which directly improves creditors' payoff, but the low-type's low action decreases creditors' payoff compared to pooling on some higher value.

Lower precision of creditors' private signal decreases payoff, as it increases strategic threshold, while keeping the fundamental threshold fixed. That increases the likelihood of making both types of errors (withdrawal if the bank eventually survives or stay while it fails). Lower precision has another indirect effect which points to the same direction: it decreases the separating signal level, which decreases payoff of the creditors.

Pooling equilibrium: Expected payoff in a pooling equilibrium is slightly more tedious to calculate because we have to take into account that the two types of banks are facing a mass of creditors who follow the same strategy in equilibrium, but the banks have different failure thresholds. Nevertheless, the calculations are fairly straightforward using the expressions derived under Pooling Equilibrium section.

Lemma 5 *In any feasible pooling equilibrium when banks pool at risk management signal s, creditors facing the type $\{L, H\}$ respectively obtain the following expected payoffs:*

$$\mathbb{E}_1[u|L] = \frac{1}{2\eta}\left((\bar{\theta}_2 - t\underline{\theta}_2) + (t^2(1+2\sigma) - t\theta_1^{AV}s - 2\sigma t) + t\sigma - \sigma(t - \frac{sp_H\Delta\theta_1}{1+2\sigma})^2 - \hat{\theta}_P^L\right)$$
$$\mathbb{E}_1[u|H] = \frac{1}{2\eta}\left((\bar{\theta}_2 - t\underline{\theta}_2) + (t^2(1+2\sigma) - t\theta_1^{AV}s - 2\sigma t) + t\sigma - \sigma(t + \frac{sp_L\Delta\theta_1}{1+2\sigma})^2 - \hat{\theta}_P^H\right)$$

The ex-ante expected payoff is therefore

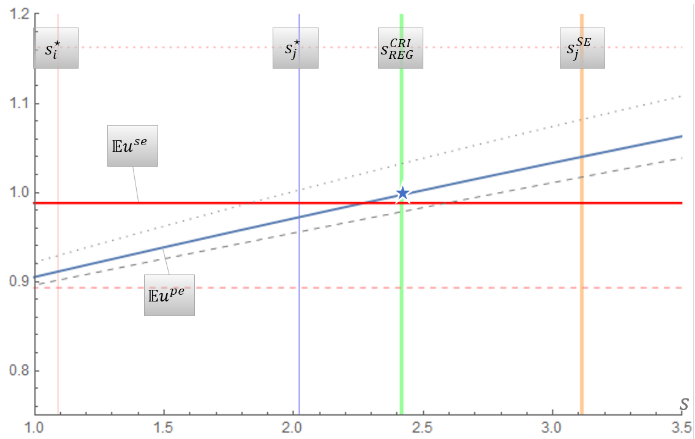
$$\mathbb{E}_0 u = \frac{1}{2\eta}\left((\bar{\theta}_2 - t\underline{\theta}_2) + t^2(1+2\sigma) + \theta_1^{AV}s(1-t) - \sigma t - \sigma p_L\left(t - \frac{sp_H\Delta\theta_1}{1+2\sigma}\right)^2 - \sigma p_H\left(t + \frac{sp_L\Delta\theta_1}{1+2\sigma}\right)^2 - t\right)$$

where $p_L = p$ and $p_H = 1-p$.

Proof. See Appendix. ■

It is possible to analytically determine the region where a pooling equilibrium is preferred by creditors from ex-ante point of view, but the calculations are tedious. When $p = 1$ (all types are low-types), the SE-payoff equals to the PE-payoff if s is set to $s_L^\star$. Pooling on every other value higher than that improves the creditors' payoff. In contrast, when $p = 0$ (all types are high-types), the SE-payoff equals the pooling payoff if pooling were set to the separating (high) level action. In this situation imposing pooling at any level $s < s_i^{cri}$ reduces the payoff to creditors (but leaves more profit to the banks). In every point between those extremes, there exists a pooling level s_p^0 which is just sufficient to make sure that $\mathbb{E}u^{SE} < \mathbb{E}u^{PE}$. This value is decreasing in p (with more low-type banks, a lower pooling level is sufficient

Figure 2.10: Illustration of creditors' payoff



to increase utility of creditors).

Creditors' payoff is increasing in precision in all equilibria. This is due to the decrease in both types of errors what creditors make due to the effective strategy. Higher precision means higher separating signal for the high type, as established in the 'separating equilibrium' section. That means, a higher s^p is required to 'compensate' creditors. Higher precision pushes s_p^0 upwards.

2.5 Empirical analysis

We illustrate our model with two simple empirical analyses. The first example builds on the recent introduction of Liquidity Coverage Ratio (LCR) in the US. In essence, the LCR places a quantitative lower bound on the amount of liquid assets which must be held by financial institutions at all time. In the analysis we exploit the fact that the rule only applies to 'large, internationally active' bank holding companies and show that - consistently with the idea of squeezing out separating equilibrium - the introduction of LCR regulation was followed by a larger decrease of volatility of cash ratios for holding companies which were subject to the newly introduced regulation, compared to those below the threshold of qualifying for regulation.

Our second example investigates changes in equity ratios around the introduction of the first generation of Basel capital regulations in 1988. Interestingly, the data shows a significant increase of the number of banks with high capital ratios after the introduction of the new regulatory regime. According to our theory, this is consistent with a quantitative regulatory requirement set too low, and therefore being unable to squeeze out separating equilibrium.

2.5.1 Liquidity Coverage Ratio

As part of its regulatory reform package in response to the financial crisis known as Basel III, the Basel Committee of Banking Supervision has put forward a series of measures concerning the liquidity risk framework of financial institutions. The agenda consists of two key elements: the Liquidity Coverage Ratio (LCR) requires banks to hold an adequate amount of highly liquid assets to cover outflows in a crisis scenario over a 30 days period, while the Net Stable Funding Ratio (NSFR) supplements this measure by ensuring a sustainable asset-liability maturity structure over a longer time horizon. In our analysis we focus on the former measure.

The Basel Committee announced the final version of LCR in January, 2013[20] and adopted a gradual approach for implementation, with the full version in effect from January 2019. In November 2013, US authorities proposed an LCR regulation largely consistent with the guidelines set forth by the Basel Committee. The final rule was adopted in September 2014, being in effect from January 2015, with a much shorter transition period than the Basel III proposal *"to preserve the strong liquidity positions many U.S. banking organizations have achieved since the recent financial crisis"*.

For our analysis, we have constructed a 'treatment' sample consisting of internationally active bank holding companies (BHC's) with Total Assets more than $50bn, and we look at changes of cash ratios (defined as Cash / Total Assets) from 2011 to 2016, based on FR-Y-9C filings. Our control sample consists of BHC's with Total Assets between $10bn and $40bn, which consistently reported throughout the whole sample period. Table 2.1 summarizes the treatment and control sample.

	treatment	control
nrBanks	32	73
avgCashRat	0.0142	0.0143
stDevCashRat	0.0067	0.0082
totalAsset	441 765 936	16 799 271

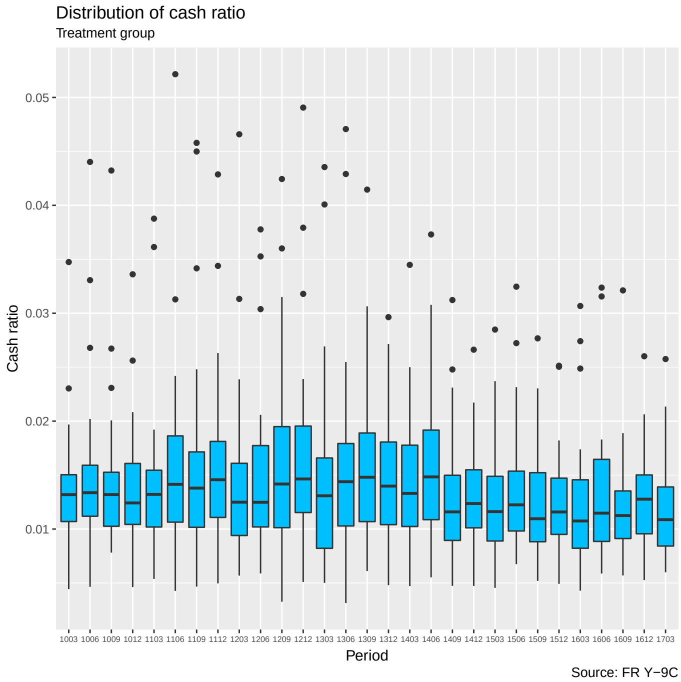
Table 2.1: Overview of treatment and control samples

Findings

We find that both the mean and the standard deviation of cash ratios decreased during the period for both the treatment and the control sample. However, following the announcement of Basel III LCR, and especially around the introduc-

[20]Basel III, B.C.B.S. "The Liquidity Coverage Ratio and liquidity risk monitoring tools." Bank for International Settlements (2013).

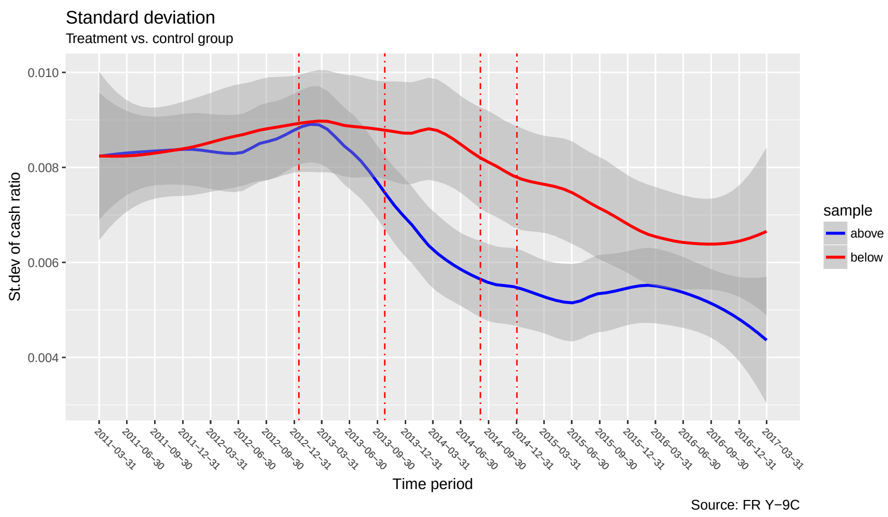
Figure 2.11: Distribution of cash-ratios of large BHC's



tion of the follow-up US regulation, the standard deviation of cash ratios in the treatment sample started to decrease significantly more sharply than that of the control sample (see Figure 2.12). This was mainly driven by the disappearance of larger values (i.e. BHC's with too high cash ratios, see Figure 2.11), which, in the context of our model, can be interpreted as an elimination of separating equilibrium. We plot the six BHC's with the biggest cash drops on average before/after the event date in Figure 2.13a. To further emphasize this finding, in Figure 2.13b we plot the difference of average standard deviation for the sample and treatment group before and after the date of announcement of US regulation. The increasing difference justifies the larger clustering of observations in the treatment sample relative to control as an effect of introducing liquidity regulation.

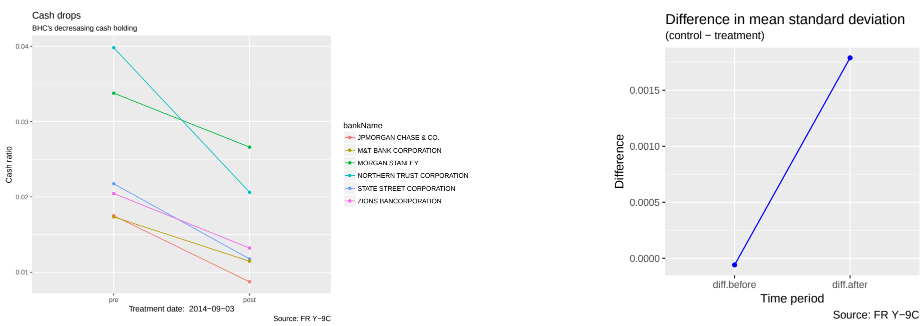
Finally, we perform a Kolmogorov-Smirnov test to mechanically compare the distributions of treatment and control samples. Before the announcement of LCR regulation, we cannot reject the null-hypothesis that the two distributions are the same (p-value: 0.79), while it can be rejected after the event (p-value: 0.01).

Figure 2.12: Standard deviation of cash ratios



Dispersion of cash-ratios decreased more sharply for BHC's above the regulation threshold than those of below the threshold.

Figure 2.13: Impact of LCR regulation

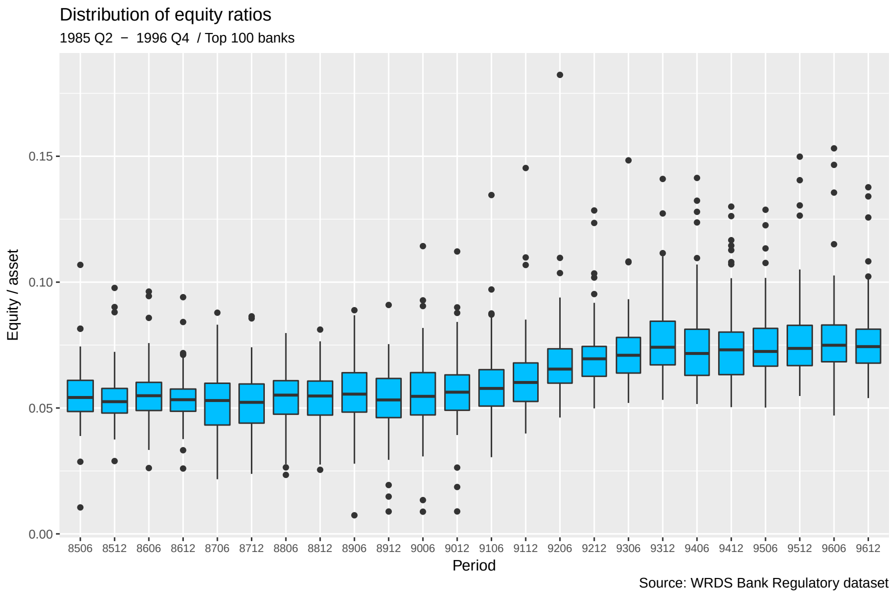


(a) (b)

2.5.2 Basel I Capital Regulation

Our second empirical analysis focuses on the introduction of minimum capital regulations under Basel I. The Basel Committee of Banking Supervision (BCBS) published the requirements in 1988 and required banks to maintain a minimum ratio of capital over total risk-weighted asset. Although we do not attempt reconstruct the nominator (weighted sum of various elements of banks' capital) or the denominator (risk-weighted assets), we believe that for our purposes the plain equity ratio defined

Figure 2.14: Distribution of equity ratios



as

$$er := \frac{\text{Total Equity}}{\text{Total Assets}}$$

will suffice. Figure 2.14 plots the distribution of equity ratios of the top 100 (by Total Asset) US-regulated bank in the period 1985-1996. Even without any formal statistical analysis, a significant structural change is recognizable around the introduction of Basel I (although announced in 1988, the capital requirements were binding from 1992). First, the bottom-end of the distributions notably shifted upwards, consistently with the regulatory intention behind the new set of rules.[21] Another visible characteristic however, which was certainly not an explicit regulatory aim, an increase in variance, and especially an increase of the number of 'outliers', i.e. institutions maintaining significantly larger equity ratios.

In the context of our model, this effect is consistent with a quantitative regulatory minimum which is not sufficiently large to squeeze out separating equilibrium. Indeed, as we have shown in Section 2.3.3, an insufficiently high regulatory minimum can preserve the incentives for signalling and even increase the minimum signal which is required to maintain separation.

[21]Note that this notable upward shift is missing from the time series of distributions of cash ratios. This is due to the fact that cash holdings have significantly increased during the financial crisis, but already were on a downward trajectory, and one of the reasons of the swift introduction of LCR in the US was to prevent the elevated levels to fall below pre-crisis levels.

2.6 Conclusions

In this paper we proposed a model which combines signalling and global games to understand the informational impact of risk management measures and the possible effect of regulation in banking. We established two main results: (i) absent regulation, banks have incentives to signal their quality and may engage in 'excess risk management', which is inefficient; (ii) a financial regulation can squeeze out inefficient separation and improve welfare by introducing a quantitative *minimum* of the given risk measure. Our results provide a novel perspective on understanding some consequences of financial regulation.

The model has testable empirical predictions: for example, in the context of liquidity holdings, in the absence of quantitative regulation in place, we expect high dispersion of liquidity ratios, which is consistently found in bank as well as mutual fund databases. Introducing liquidity regulation neutralizes this incentive, so we expect clustering of observations around the requirement. We find that changes in cash ratios following the recent introduction of LCR ratios in the US are consistent with this hypothesis. Changes in equity ratios, however, around the introduction of Basel I regulatory capital regime are rather consistent with an insufficiently large regulatory minimum, unable to squeeze out separation.

2.A Appendix A - Proofs

2.A.1 Proof of Lemma 1

We prove the Lemma in two steps. First, we focus on equilibria in monotone strategies, and show the existence of a unique threshold equilibrium in monotone strategies. The proof uses the contraction mapping theorem to show that the result holds for a large class of functional forms for $\mathcal{R}$. Then, we apply the procedure of iterative elimination of strictly dominated strategies (ISD) and show that this threshold equilibrium is the unique equilibrium which survives ISD, therefore, no other equilibria exist. We will assume throughout the proof that fundamental uncertainty is such that with some probability even the strongest banks can fail, while even the weakest bank can survive for any risk management actions which may be taken in any equilibrium.

Assumption 2 *We assume that there exists $\theta_{2\star}$ and $\theta_2^\star$ such that*

- *Lower dominance region: $\forall \theta_1 \in \{\theta_1^L, \theta_1^S\}$, and $\forall s \in \mathcal{S}^\star$, $\exists \theta_{2\star} > \underline{\theta}_2$ such that $\forall \theta_2 < \theta_{2\star}$:*
$$\mathcal{R}(\theta_1, \theta_2, s, 0) < 0$$
- *Upper dominance region: $\forall \theta_1 \in \{\theta_1^L, \theta_1^S\}$, and $\forall s \in \mathcal{S}^\star$, $\exists \theta_2^\star < \overline{\theta}_2$ such that $\forall \theta_2 > \theta_2^\star$:*
$$\mathcal{R}(\theta_1, \theta_2, s, 1) > 1$$

where $\mathcal{S}^\star$ is the set of equilibrium values of signal s.

Part 1: Equilibrium in monotone strategies:

Suppose that creditors run on the bank if and only if their private signal satisfies $x_i < \hat{x}$. Given this strategy, let us denote the aggregate size of the run for any realization of the fundamental $\theta_2 \in [\underline{\theta}_2, \overline{\theta}_2]$ by $\mathcal{A}(\hat{x}, \theta_2)$. Now we define the expected difference in utilities between actions WAIT and RUN for a creditor receiving signal x as

$$\begin{aligned} V(x, \hat{x}) :=& \mathbb{E}\left[u(\theta_1, \theta_2, s, \mathcal{A}(\hat{x}, \theta_2)) - t | x\right] \\ =& Prob[\mathcal{R}(\theta_1, \theta_2, s, \mathcal{A}) < 0 | x](0 - t) + Prob[\mathcal{R}(\theta_1, \theta_2, s, \mathcal{A}) > 0 | x](1 - t) \end{aligned} \tag{2.1}$$

Let us introduce the notation $\nu(x, \hat{x})$ for the posterior probability assessment of the bank's survival by a creditor who receives a private signal x, given that the equilibrium is that creditors run if and only if $x < \hat{x}$.

$$\nu(x, \hat{x}) := Prob[\mathcal{R}(\theta_1, \theta_2, s, \mathcal{A}(\hat{x}, \theta_2)) > 0|x]$$

With this notation, we can rewrite the difference utility function as

$$V(x, \hat{x}) = \nu(x, \hat{x}) - t \tag{2.2}$$

For every $\theta_2 \in [\underline{\theta}_2, \overline{\theta}_2]$ the aggregate size of a run is

$$\mathcal{A}(\hat{x}, \theta_2) = Prob[x_i \leq \hat{x}|\theta_2] = Prob(\theta_2 + \sigma\epsilon_i \leq \hat{x}|\theta_2) = Prob\left(\epsilon_i \leq \frac{\hat{x} - \theta_2}{\sigma}\right)$$

Given our assumption on the stochastic structure of the game, that ϵ is distributed uniformly, this implies

$$\mathcal{A}(\hat{x}, \theta_2) = \min\left(\max\left(\frac{\hat{x} - \theta_2 + \sigma}{2\sigma}; 0\right); 1\right) \tag{2.3}$$

As $\mathcal{A}(\hat{x}, \theta_2)$ monotone decreases in θ_2, and $\mathcal{R}$ strictly monotone decreases in $\mathcal{A}$, the function $\mathcal{R}$ monotonically and continuously increases in the fundamental θ_2. Furthermore, due to the existence of lower and upper dominance regions (Assumption 2), by the Intermediate Value Theorem, there exists a unique value of θ_2, denoted by $\hat{\theta}_2$, which is the implicit solution of the following equation:

$$\mathcal{R}(\theta_1, \hat{\theta}_2, s, \mathcal{A}(\hat{x}, \hat{\theta}_2)) = 0 \tag{2.4}$$

As a consequence, the bank fails if and only if $\theta_2 < \hat{\theta}_2(\hat{x})$ *(existence of fundamental threshold)*.

Because $\mathcal{A}(\hat{x}, \theta_2)$ is strictly monotone increasing in $\hat{x}$, the fundamental threshold $\hat{\theta}_2(\hat{x})$ is increasing in $\hat{x}$.

Given the existence of a unique $\hat{\theta}_2$, the posterior probability that a creditor receiving signal x and correctly anticipating equilibrium strategic threshold $\hat{x}$ attaches to the event of bank's survival is

$$\nu(x, \hat{x}) = Prob[\theta_2 > \hat{\theta}|x] = 1 - Pr[\theta_2 \leq \hat{\theta}_2|x] = Pr\left(\epsilon \leq \frac{x - \hat{\theta}_2}{\sigma}\right) = \frac{x - \hat{\theta}_2(\hat{x}) + \sigma}{2\sigma}$$

The function $\nu(x, \hat{x})$ is decreasing in $\hat{x}$, increasing in x, and is continuous in x. Therefore, it is possible to define a unique function $h(\hat{x})$, such that the solution of the equation $x = h(\hat{x})$ also solves

$$V(\hat{x}, \hat{x}) = 0$$

The function $h(\hat{x})$ summarizes best responses of the game, is continuous and increasing in $\hat{x}$, therefore its fixed points coincide with the monotone equilibria of the game.

The last step of the proof is to show the existence of a unique fixed point of the function $h(\hat{x})$. For simplicity we show this first to the main example functional form, then generalize. With $\mathcal{R} = \theta_1 s + \theta_2 - \mathcal{A}(\hat{x}, \theta_2)$, we can solve 2.4 explicitly, and we obtain

$$\hat{\theta}_2(\hat{x}) = \frac{\hat{x}}{1+2\sigma} - \frac{2\sigma}{1+2\sigma}\theta_1 s + \frac{\sigma}{1+2\sigma} \tag{2.5}$$

Recall the function $V(x, \hat{x})$:

$$V(x, \hat{x}) = \frac{x - \hat{\theta}_2 + \sigma}{2\sigma} - t \tag{2.6}$$

which implies that the function $h(\hat{x})$ can be written explicitly as

$$h(x) := 2\sigma t - \sigma + \hat{\theta}_2(\hat{x}) \tag{2.7}$$

Solving for the fixed point of h amounts to solving the equation $x = h(x)$, which gives the unique solution

$$\hat{x} = t - \theta_1 s + 2\sigma t - \sigma \tag{2.8}$$

Substituting back gives the result for equilibrium fundamental threshold

$$\hat{\theta}_2 = t - \theta_1 s \tag{2.9}$$

This concludes the *existence* proof for the *example functional form.*

The general proof: Now we generalize the last part of the proof to an arbitrary function $\mathcal{R}(\cdot)$ satisfying certain conditions. Notice that the critical part is to show the existence of a unique fixed point of function $h(x)$. We will demonstrate this by showing that the best response function is a contraction. A sufficient condition for contraction is that the derivative of a function is less than 1 everywhere on its domain. We are therefore looking for the derivative

$$\frac{\partial h(x)}{\partial x}$$

Notice that the generic functional form 2.7 is valid for an arbitrary $\mathcal{R}$, therefore

$$\frac{\partial h(\hat{x})}{\partial \hat{x}} = \frac{\partial \hat{\theta}_2(\hat{x})}{\partial \hat{x}} \tag{2.10}$$

The function $\hat{\theta}_2(\hat{x})$ is implicitly defined in Equation 2.4. First, suppose that $\mathcal{R}$ only depends on θ_2 directly and through the effect on $\mathcal{A}$, but not interacting with other variables. We can use Implicit Function Theorem to write

$$\frac{\partial \theta_2}{\partial x} = -\frac{\partial \mathcal{R}/\partial x}{\partial \mathcal{R}/\partial \hat{\theta}_2} = -\frac{\frac{\partial \mathcal{R}}{\partial \mathcal{A}}\frac{\partial \mathcal{A}}{\partial x}}{\frac{\partial \mathcal{R}}{\partial \theta_2} + \frac{\partial \mathcal{R}}{\partial \mathcal{A}}\frac{\partial \mathcal{A}}{\partial \theta_2}}$$

Because $\partial \mathcal{A}/\partial x = 1/2\sigma$ and $\partial \mathcal{A}/\partial \theta_2 = -1/2\sigma$ it simplifies a little bit:

$$\frac{\partial \theta_2}{\partial x} = -\frac{\partial \mathcal{R}/\partial x}{\partial \mathcal{R}/\partial \hat{\theta}_2} = -\frac{\frac{1}{1\sigma}\frac{\partial \mathcal{R}}{\partial \mathcal{A}}}{\frac{\partial \mathcal{R}}{\partial \theta_2} - \frac{1}{2\sigma}\frac{\partial \mathcal{R}}{\partial \mathcal{A}}}$$

The assumptions $\partial \mathcal{R}/\partial \mathcal{A} < 0$ and $\partial \mathcal{R}/\partial \theta_2 > 0$ are sufficient to guarantee that this expression is smaller than 1, and that is sufficient condition for contraction mapping.

We have established the existence of a threshold-equilibrium in monotone strategies for an arbitrary function of $\mathcal{R}$ satisfying the assumptions of our model.

Part 2: Uniqueness of equilibrium.

First, we maintain as an assumption that for all $s \in \mathcal{S}^\star$, there exists $\theta_2(s)$ such that $\mathcal{R}(\theta_1, \theta_2(s), s, 0) < 0$, where $\mathcal{S}^\star$ is the set of values of intervention s which may be taken in any equilibria. In this *lower dominance region* the bank defaults even with no runs ($\alpha = 0$). Denote this critical value by $H_0(s)$.[22] Under the *most optimistic beliefs*, a creditor whose posterior belief over the fundamental θ_2 places some positive weight on $\theta_2 < H_0(s)$ believes that the bank fails if and only if $\theta_2 < H_0(s)$. Because the conditional probability $Pr(\theta_2 < \xi | x_i)$ is decreasing in x_i (a lower signal increases the probability of a lower θ_2), it decreases the difference utility under these beliefs. Therefore, there exists a critical signal, denoted by $h_0(s)$ such that all creditors with signal $x_i < h_0(s)$ find it dominant to run on the bank. The critical signal $h_o(s)$ is determined implicitly by the indifference condition of that creditor:

$$Pr[\mathcal{R}(\theta_1, \theta_2, s, \mathcal{A}) > 0 | h_o(s)] = t$$

where

$$Pr[\mathcal{R}(\theta_1, \theta_2, s, \mathcal{A}) > 0 | h_o(s)] = Pr[\theta_2 > H_0(s) | h_0(s)] = \frac{h_0(s) - H_0(s) + \sigma}{2\sigma}$$

[22]For the example functional form of $\mathcal{R}$: $H_0(s) = -\theta_1 s$

therefore

$$h_0(s) = H_0(s) + 2\sigma t - \sigma$$

We have established that it is always strictly dominant to run whenever $x_i < h_0(s)$.

For $n \geq 0$ define the sequences $H_n(s)$ and $h_n(s)$ as follows:

$$\mathcal{R}(\theta_1, H_n(s), s, \mathcal{A}(h_{n-1}(s), H_n(s)) = 0 \tag{2.11}$$

and

$$\frac{h_n(s) - H_n(s) + \sigma}{2\sigma} = t \tag{2.12}$$

Equation 2.11 implicitly defines a sequence for the value of fundamental θ_2 which solves $\mathcal{R}(\cdot) = 0$, given that *only* creditors with $x_i < h_{n-1}(s)$ run on the bank. Because $\mathcal{R}$ is increasing in θ_2 and decreasing in $\mathcal{A}$, and $\mathcal{A}$ is increasing in $\hat{x}$ and decreasing in θ_2, we conclude that $H_n(s)$ is increasing in $h_{n-1}(s)$. The sequence $h_n(s)$ in Equation 2.12 defines the value of a private signal x, such that for a creditor it is strictly dominant to run, given he believes that the bank fails if and only if $\theta_2 < H_n(s)$. The sequence $h_n(s)$ is clearly monotonically increases in H_n. Because both sequences are bounded and monotonically increasing, they converge to some value $\underline{h}$ and $\underline{H}$. But $\underline{h}$ must be a fixed point of the function $h(\hat{x})$ defined previously.

What we have proven: even creditors *with the most optimistic beliefs* regarding other creditors' behaviour will find it iteratively dominant to run on the bank if $x_i < \hat{x}$.

Analogously, it is possible to construct iterative deletion of strictly dominated strategies from above: this will show that *even with the most pessimistic beliefs*, that means, RUN always whenever it is not strictly dominant to WAIT, it is never rationalizable to RUN if $x_i > \hat{x}$. The two parts together implies that the *unique, rationalizable action profile* for creditors is RUN if and only if $x_i < \hat{x}$.

2.A.2 Proof of Lemma 2

The objective function of the bank is to maximize expected profit, conditional on the equilibrium in the second stage. That is, the bank's program is

$$\begin{aligned} \max_s \quad & \pi(\theta_1, s) \\ \text{where} \quad & \pi(\theta_1, s) = Pr[\theta_2 > \hat{\theta}_2 | \theta_1, s]\,(k - cs) \end{aligned}$$

Let $\eta = \frac{\overline{\theta}-\underline{\theta}}{2}$ denote the "prior noise", and introduce $\rho(\cdot)$ to denote the probability of survival:

$$\rho(\theta_1, s) = Pr[\theta_2 > \hat{\theta}_2 | \theta_1, s] = \frac{1}{2\eta}\left(\overline{\theta}_2 - \hat{\theta}_2(\theta_1, s)\right)$$

Then we can write the expected profit for any given signal s as

$$\pi(\theta_1, s) = \int_{\hat{\theta}_2}^{\overline{\theta}_2} (k - c \cdot s)\, \mu(\cdot) d\theta_2 = \rho(\theta_1, s)(k - cs)$$

where $\mu(\cdot)$ denotes the prior on θ_2. The bank's optimal risk management action trades off cost of the signal with an increased probability of survival. The first-order condition

$$\frac{\partial \pi}{\partial s} = \frac{\partial \rho}{\partial s}(k - c \cdot s) - \rho c = \frac{1}{2\eta}\left(\theta_1(k - c \cdot s) - c\overline{\theta}_2 + c(t - \theta_1 s)\right) = 0$$

implies the optimal intervention $s^\star(\theta_1)$ and the associated optimal profit $\pi^\star(\theta_1)$ as

$$s^\star(\theta_1) = \frac{1}{2}\left(\frac{k}{c} - \frac{\overline{\theta}_2 - t}{\theta_1}\right)$$
$$\pi^\star(\theta_1) = \frac{\left(c\left(\overline{\theta}_2 - t\right) + \theta_1 k\right)^2}{4c\theta_1}$$

Next, we calculate the first derivatives with respect to θ_1 for both the optimal signal and the optimal profit. For the optimal signal:

$$\frac{\partial s^\star}{\partial \theta_1} = \frac{1}{2}\frac{\overline{\theta}_2 - t}{\theta_1^2}$$

This is always strictly positive whenever the problem satisfies the natural parametric assumptions $\overline{\theta}_2 > 1 > t$. For the optimal profit:

$$\frac{\partial \pi^\star}{\partial \theta_1} = \frac{k^2\theta_1^2 - [c(t - \overline{\theta}_2)]^2}{4c\theta_1^2} = \frac{1}{4}\left(\frac{k^2}{c} - \frac{c^2(t - \overline{\theta}_2)^2}{\theta_1^2}\right)$$

the derivative is strictly positive whenever the following condition holds.

$$\frac{k}{c} \geq \frac{\overline{\theta}_2 - t}{\theta_1}$$

Note that the condition trivially implies $s^\star \geq 0$. $\square$

2.A.3 Proof of Theorem 1

Separating equilibrium can be maintained if the following two incentive compatibility constraints are satisfied:

$$\pi^{eq}(s_L^\star, \theta_1^L) \geq \pi^{off}(s_H, \theta_1^L, \theta_1^H) \tag{IC_L}$$
$$\pi^{eq}(s_H, \theta_1^H) \geq \pi^{off}(s_{H.L}^\star, \theta_1^H, \theta_1^L) \tag{IC_H}$$

where the functions $\pi^{eq}(s,\theta)$ and $s_L^\star(\cdot)$ are defined in Lemma 1, while $\pi^{off}(s,\theta,\theta)$ and $s_{L.H}^\star$ are derived in equations 2.8 and 2.9 in the main text. Let us denote by $\rho_i(s)$ the probability that a bank of type $i \in \{L,H\}$ survives if they creditors believe that it is of type i with action s, while $\rho_{i.j}$ the probability that a bank of type $i \in \{L,H\}$ survives if they creditors believe that it is of type $-i$:

$$\rho_i(s) := Pr[\theta_2 > \hat{\theta}_2^i(s)] \quad \forall i \in \{L,H\}$$
$$\rho_{i.j}(s) := Pr[\theta_2 > \hat{\theta}_2^{i.j}(s)] \quad \forall i \in \{L,H\}$$

With this notation the incentive compatibility constraint for IC_L can be written as:

$$\rho_L(s_L^\star)(k - cs_L^\star) - \rho_{L.H}(s)(k - cs) \geq 0 \tag{2.13}$$

After substitution of the expressions for probabilities, we have

$$\frac{1}{2\eta}\left(\overline{\theta}_2 - \hat{\theta}_2^L(s_L^\star)\right)(k - cs_L^\star) - \frac{1}{2\eta}\left(\overline{\theta}_2 - \hat{\theta}_2^{L.H}(s)\right)(k - cs) \geq 0$$

Recall that $\hat{\theta}_2^L(s) = t - \theta_1^L s$, $s_L^\star = \frac{1}{2}\left(\frac{k}{c} - \frac{\overline{\theta}_2 - t}{\theta_1^L}\right)$ and $\hat{\theta}_2^{L.H}(s) = t - s\left(\frac{\theta_1^H + 2\sigma\theta_1^L}{1+2\sigma}\right) = t - s\Theta$, where for simplicity, we introduce $\Theta := \left(\frac{\theta_1^H + 2\sigma\theta_1^L}{1+2\sigma}\right)$. Then it follows that

$$\overline{\theta}_2 c(s - s_L^\star) - tc(s - s_L^\star) - c\theta_1^L[s_L^\star]^2 + c\Theta s^2 - k\left(s\Theta - \theta_1^L s_L^\star\right) \geq 0$$
$$c\Theta s^2 + \left(c(\overline{\theta}_2 - t) - k\Theta\right)s - \left(c(\overline{\theta}_2 - t) - k\theta_L\right)s_L^\star - c\theta_1^L[s_L^\star]^2 \geq 0$$

After substitution we have

$$\left(c(\overline{\theta}_2 - t) - k\theta_L\right)s_L^\star - c\theta_1^L[s_L^\star]^2 = \frac{c\theta_1^L}{4}\left(\frac{k}{c} - \frac{\overline{\theta}_2 - t}{\theta_1^L}\right)^2 \qquad \text{“C”}$$

$$\Theta^2 c^2\left(\frac{k}{c} - \frac{\overline{\theta}_2 - t}{\Theta}\right)^2 \qquad \text{“}B^2\text{”}$$

So the expression $\frac{B^2-4AC}{4A^2}$ takes a particularly simple form:

$$\frac{1}{4}\left(\left(\frac{k}{c}\right)^2\left[\frac{\Theta-\theta_1^L}{\Theta}\right]+\left(\overline{\theta}_2-t\right)^2\left[\frac{1}{\Theta^2}-\frac{1}{\Theta\theta_1^L}\right]\right)$$

Putting together the formula for solving a quadratic equation and substituting the expression for Θ we get

$$s_{L.2}^{cri}=\frac{1}{2}\left(\frac{k}{c}-\frac{(1+2\sigma)(\overline{\theta}_2-t)}{2\sigma\theta_1^L+\theta_1^H}\right)+\frac{1}{2}\sqrt{\frac{\Delta\theta_1}{\left(2\sigma\theta_1^L+\theta_1^H\right)}}\sqrt{\left(\frac{k^2}{c^2}-\frac{(1+2\sigma)}{\left(2\sigma\theta_1^L+\theta_1^H\right)}\frac{(\overline{\theta}_2-t)^2}{\theta_1^L}\right)}$$

With similar algebra it is possible to derive the solution of IC_H

$$s_{H.2}^{cri}=\frac{1}{2}\left(\frac{k}{c}-\frac{\overline{\theta}_2-t}{\theta_1^H}\right)+\frac{1}{2}\sqrt{\Delta\theta_1\left(\frac{k^2}{c^2}\frac{1}{(1+2\sigma)\theta_1^H}-\frac{(\overline{\theta}_2-t)^2}{\theta_1^H}\frac{1}{\theta_1^L+2\sigma\theta_1^H}\right)}$$

Both thresholds are decreasing functions of the creditors' noise σ.

Limit of large noise: We start with the limit of large noise where $\sigma\to\infty$.

$$\lim_{\sigma\to\infty}s_{L.2}^{cri}=\frac{1}{2}\left(\frac{k}{c}-\frac{(1+2\sigma)(\overline{\theta}_2-t)}{2\sigma\theta_1^L+\theta_1^H}\right)+0=s_L^\star$$

Explanation: all terms under the square-root trivially converge to 0. The second term within the brackets — it is easy to calculate the limit of the inverse:

$$\frac{2\sigma\theta_1^L+\theta_1^H}{(1+2\sigma)(\overline{\theta}_2-t)}=\frac{(1+2\sigma)\theta_1^L+\theta_1^H-\theta_1^L}{(1+2\sigma)(\overline{\theta}_2-t)}=\frac{\theta_1^L}{\overline{\theta}_2-t}+\frac{\Delta\theta_1}{(1+2\sigma)(\overline{\theta}_2-t)}\to\frac{\theta_1^L}{\overline{\theta}_2-t}$$

This implies the result. The limit for the high-threshold is trivial to calculate:

$$\lim_{\sigma\to\infty}s_{H.2}^{cri}=\frac{1}{2}\left(\frac{k}{c}-\frac{\overline{\theta}_2-t}{\theta_1^H}\right)=s_H^\star$$

Since $s_{L.2}^{cri}$ continuously and monotonically approaches $s_L^\star$ as $\sigma\to\infty$, and $s_L^\star<s_H^\star$, by the intermediate value theorem there exists a unique value of σ, denoted by $\overline{\sigma}$, such that

$$s_{L.2}^{cri}(\overline{\sigma})=s_H^\star$$

Whenever $\sigma>\overline{\sigma}$, the pair $\{s_L^\star;s_H^\star\}$ is incentive-compatible for the low (L) and high (H) types as well, and the first-best can be maintained as a separating equilibrium.

Limit of small noise: Now we turn to the analysis of the case of small noise, as $\sigma \to 0$. First, calculate the limit of critical signals as $\sigma \to 0$. After some algebraic manipulations it is possible to show that the critical incentive-compatible signals converge to the same expression:

$$\lim_{\sigma \to 0} s^{cri}_{L.2} = \lim_{\sigma \to 0} s^{cri}_{H.2} = \frac{1}{2}\left(\frac{k}{c} - \frac{\overline{\theta}_2 - t}{\theta_1^H}\right) + \frac{1}{2}\sqrt{\frac{\Delta\theta_1}{\theta_1^H}\left(\frac{k^2}{c^2} - \frac{(\overline{\theta}_2 - t)^2}{\theta_1^L \theta_1^H}\right)}$$

Conditions for existence of the equilibrium

A 'constructive' approach to prove the statement of the theorem by solving the equation $s_i^{cri} = s_j^{cri}$ is not possible due to the analytical complexity of the non/limiting case. Instead, we prove the theorem using the following steps, which are analytically easier to calculate:

1. Calculate the partial derivatives $\frac{\partial s_L^{cri}}{\partial \sigma}$ and $\frac{\partial s_H^{cri}}{\partial \sigma}$

2. Consider the value of the derivatives at $\sigma = 0$. If $(s_L^{cri})'(0) < (s_H^{cri})'(0)$ then s_H^{cri} approaches the limit faster, meaning that for sufficiently small σ, we must have $s_L^{cri} > s_H^{cri}$. Note that from the limiting cases at $\sigma \to \infty$ it is obvious that for large enough σ, $s_L^{cri} < s_H^{cri}$.

3. Analytically, we solve the equation $(s_L^{cri})'(0) = (s_H^{cri})'(0)$ for c. This gives a critical cost level $\hat{c}$ such that $s_L^{cri} > s_H^{cri}$ for sufficiently small σ

We omit the detailed calculations to save space. The solution for $\hat{c}$ is

$$\hat{c} = \sqrt{\frac{\Delta\theta_1 (k\theta_1^L)^2}{\left(3\theta_1^L + \theta_1^H\right)\left(\overline{\theta}_2 - t\right)}} \tag{2.14}$$

2.A.4 Proof of Lemma 3

We prove a more general version of Lemma 3, with an arbitrary N banks, from which the version in the main text will be trivial. Suppose the number of banks is $N \geq 2$, each with types θ_1^n. Without loss of generality we can determine the indexing of banks such that $n_i < n_j \Leftrightarrow \theta_1^{n_i} < \theta_i^{n_j}$. Let the prior distribution of types be $Pr[\theta_1 = \theta_1^n] = p_n$. It is useful to interpret $N = 1, 2...$ as *quality classes* and the probability p_n representing the mass of institutions belonging to this quality class. Suppose that a closed subset of institutions $\mathcal{N} := \underline{n} < n < \overline{n}$ are pooling on the same risk-management signal s_p. Let us define the conditional distribution of banks belonging to $\mathcal{N}$ as $\tilde{\mathcal{P}} := \{\tilde{p}_n\}_{\underline{n}}^{\overline{n}}$, it is straightforward that

$$\tilde{p}_n = \frac{p_n}{\sum_{n\in\mathcal{N}} p_n}$$

Conditional on observing s_p, $\tilde{p}_n$ represents creditors' posterior probability of the event that the bank is of type n. We define the (conditional) average type as

$$\overline{\theta}_1^n := \sum_{n\in\mathcal{N}} \tilde{p}_n \theta_1^n$$

Equation 1: we start with the creditors' indifference condition. A creditor is indifferent between actions WAIT and RUN if

$$Pr[Failure] \cdot \underbrace{(0-t)}_{\text{p/o wait-run}} + Pr[Survive] \cdot \underbrace{(1-t)}_{\text{p/o wait-run}} = 0$$

Because $Pr(Failure) = 1 - Pr(Survive)$, we can rewrite this equation as

$$Pr[Survive] = t$$

Using creditors' posterior probability, we can write

$$Pr[Survive] = \sum_{\underline{n}}^{\overline{n}} \tilde{p}_n \Phi\left(\frac{\hat{x} - \hat{\theta}_1^n}{\sigma}\right)$$

After substitution

$$\sum_{\underline{n}}^{\overline{n}} \tilde{p}_n \frac{\hat{x} - \hat{\theta}_1^n + \sigma}{2\sigma} = t$$

this can be rewritten as

$$\sum \tilde{p}_n \hat{x} - \sum \tilde{p}_n \hat{\theta}_1^n + \sum \tilde{p}_n \sigma = 2\sigma t$$
$$\hat{x} - \overline{\hat{\theta}_1^n} + \sigma = 2\sigma t$$
$$\hat{x} = 2\sigma t - \sigma + \overline{\hat{\theta}_1^n}$$

Equation 2: Given strategic threshold $\hat{x}$, the fundamental threshold solves $\mathcal{R} = \theta_1 s + \theta_2 - \alpha = 0$. After substituting α and rearranging the equation, we have for each $n \in \mathcal{N}$

$$\hat{\theta}_2^n = \frac{\hat{x} + \sigma - 2\sigma\theta_1^n s}{1 + 2\sigma}$$

We can calculate $\overline{\hat{\theta}_2}$ as

$$\overline{\hat{\theta}_2} = \frac{\hat{x} + \sigma - 2\sigma s\overline{\theta_1}}{1+2\sigma}$$

Substituting back to $\hat{x}$,

$$\hat{x} = 2\sigma t - \sigma + \frac{\hat{x} + \sigma - 2\sigma s\overline{\theta_1}}{1+2\sigma}$$
$$\hat{x} = \frac{1+2\sigma}{2\sigma}\left(2\sigma t - \sigma + \frac{\sigma}{1+2\sigma} - \frac{2\sigma}{1+2\sigma} s\overline{\theta_1}\right)$$

then back to $\hat{\theta}_2^n$:

$$\begin{aligned}\hat{\theta}_2^n &= t - \frac{1}{2} + \frac{1}{2(1+2\sigma)} - \frac{s\overline{\theta_1}}{1+2\sigma} + \frac{\sigma}{1+2\sigma} - \frac{2\sigma\theta_1^n s}{1+2\sigma} \\ &= t - s\left(\frac{1}{1+2\sigma}\overline{\theta_1} - \frac{2\sigma}{1+2\sigma}\theta_1^n\right)\end{aligned}$$

The formulas for $N = 2$ trivially follows. We note that **Corollary 1** also follows in the $N \geq 2$ general case, this can be seen with trivial algebra.

—

For the N=2 case which we discuss in the main text, the incentive-compatibility constraints can be solved analytically. The binding constraint will give the following upper boundary for the set of pooling equilibria which can be maintained:

$$\overline{s_p} := \frac{1}{2}\left(\frac{k}{c} - \frac{(1-2\sigma)(\overline{\theta}_2 - t)}{\overline{\theta_1} + 2\sigma\theta_1^L} + \sqrt{\frac{p_H\Delta\theta_1}{\overline{\theta_1} + 2\sigma\theta_1^L}\left(\frac{k^2}{c^2} - \frac{(1+2\sigma)(\overline{\theta}_2 - t)^2}{\theta_1^L(\overline{\theta_1} + 2\sigma\theta_1^L)}\right)}\right)$$

2.A.5 Proof of Theorem 2

As explained in the main text, critical regulation level is described by equation 2.14:

$$s_L^{cri} = s_H^{cri} \qquad \text{(2.14 revisited)}$$

The analytical solution to this equation is

$$\begin{aligned}s_P^{cri} &= \frac{2c(1+2\sigma)(\overline{\theta}_2 - t) + k\Delta\theta_1 - \sqrt{\left(2c(1+2\sigma)(\overline{\theta}_2 - t) + k\Delta\theta_1\right)^2 - 4kc(1+2\sigma)(\overline{\theta}_2 - t)\Delta\theta_1}}{2c\Delta\theta_1} \\ &= \frac{2c(1+2\sigma)(\overline{\theta}_2 - t) + k\Delta\theta_1 - \sqrt{\left(2c(1+2\sigma)(\overline{\theta}_2 - t)\right)^2 + (k\Delta\theta_1)^2}}{2c\Delta\theta_2} \\ &= \left(\frac{(1+2\sigma)(\overline{\theta}_2 - t)}{\Delta\theta_1} + \frac{k}{2c}\right) - \sqrt{\left[\frac{(1+2\sigma)(\overline{\theta}_2 - t)}{\Delta\theta_1}\right]^2 + \left[\frac{k}{2c}\right]^2}\end{aligned}$$

2.A.6 Proofs for Section 2.4

2.A.6.1 Bank-optimal pooling

Substituting to the ex-ante formula shows that total ex ante payoff is a decreasing function of the average of thresholds in any given equilibrium. We denote this by $\hat{\theta}_2^{AV}(s)$. The threshold $\hat{\theta}_2(s)$ is always a decreasing function of s (ceteris paribus a bank's survival is more likely with higher s). The derivative of the expected profit:

$$\frac{\partial \mathbb{E}\pi}{\partial s} = -k\frac{\partial \hat{\theta}_2^{AV}}{\partial s} - c\overline{\theta}_2 + cs\frac{\partial \hat{\theta}_2^{AV}}{\partial s} + c\hat{\theta}_2^{AV} = \frac{\partial \hat{\theta}_2^{AV}}{\partial s}(-k+cs) - c(\overline{\theta}_2 - \hat{\theta}_2^{AV}) \quad (2.15)$$

In any pooling equilibrium, from Corollary 2 we have $\hat{\theta}_2^{AV} = t - s\theta_1^{AV}$ so the partial derivative (2.15) is $-\theta_1^{AV}$ and the total expected payoff as a function of s is

$$\frac{\partial \mathbb{E}\pi}{\partial s} = \theta_1^{AV}(k-cs) - c(\overline{\theta}_2 - t + s\theta_1^{AV})$$

This allows us to calculate the *pooling* equilibrium which maximizes total bank welfare: this is pooling on the value of s which solves

$$\theta_1^{AV}(k-cs) = c(\overline{\theta}_2 - t + s\theta_1^{AV})$$

so the optimal pooling level is

$$s^{\star}_{pool} = -\frac{c(\overline{\theta}_2 - t) - \theta_1^{AV}k}{2c\theta_1^{AV}} = \frac{1}{2}\left(\frac{k}{c} - \frac{\overline{\theta}_2 - t}{\theta_1^{AV}}\right)$$

It is obvious that $s_i^{\star} < s^{\star}_{pool} < s_j^{\star}$, that is, the best pooling equilibrium is between the symmetric information benchmark intervention levels.

2.A.6.2 Creditors in separating equilibrium

The mass of creditors who run at θ_2 is $\alpha(\theta_2)$. We denote by $\theta_2^{\star}$ the value of θ_2 where $\alpha(\theta_2) = 0$ and by $\theta_{2\star}$ where $\alpha(\theta_2) = 1$. The total payoff to creditors is

$$\mathbb{E}u = \underbrace{\int_{\underline{\theta}_2}^{\theta_{2\star}} t\nu(\cdot)d\theta_2}_{\text{All withdraw}} + \underbrace{\int_{\theta_{2\star}}^{\theta_2^{\star}} t\alpha(\theta_2)\nu(\cdot)d\theta_2}_{\alpha \text{ withdraw}} + \underbrace{\int_{\hat{\theta}_2}^{\theta^{2\star}} 1-\alpha(\theta_2)\nu(\cdot)d\theta_2}_{1-\alpha \text{ stay and survive}} + \underbrace{\int_{\theta_2^{\star}}^{\overline{\theta}_2} 1\nu(\cdot)d\theta_2}_{\text{all stay and survive}}$$

where for example in the benchmark case

$$\theta_2^{\star} = t(1+2\sigma) - \theta_1 s$$
$$\theta_{2\star} = \hat{x} - \sigma = t(1+2\sigma) - \theta_1 s - 2\sigma$$

With straightforward algebra it is possible to calculate the followings:

$$\begin{aligned}\hat{\theta}_2 &= t - \theta_1 s \\ \hat{\theta}_2 - \theta_{2\star} &= 2\sigma(1-t) \\ \theta^{2\star} - \hat{\theta}_2 &= 2\sigma t \\ \alpha(\hat{\theta}_2) &= t\end{aligned}$$

The value of integrals follows:

$$\begin{aligned}\int_{\theta_{2\star}}^{\hat{\theta}_2} \alpha(\theta_2) d\theta_2 &= (1-t^2)\sigma \\ \int_{\hat{\theta}_2}^{\theta_2^\star} \alpha(\theta_2) d\theta_2 &= t^2\sigma \\ \int_{\theta_{2\star}}^{\theta_2^\star} \alpha(\theta_2) d\theta_2 &= \sigma\end{aligned}$$

so putting together we obtain:

$$Eu = \frac{1}{2\eta}\left(t(\theta_{2\star} - \underline{\theta}_2) + \sigma t - t^2\sigma + (\overline{\theta} - \hat{\theta}_2)\right)$$

We can calculate the derivative with respect to the risk-management action:

$$\frac{\partial Eu}{\partial s} = \frac{1}{2\eta}\left((1-t)\theta_1\right) > 0$$

this means creditors always prefers a higher action. After rewriting:

$$Eu = \frac{1}{2\eta}\left((\overline{\theta}_2 - t\underline{\theta}_2) - \sigma t(1-t) - (t - \theta_1 s)(1-t)\right)$$

All payoff difference is captured by the second term, the first term depends on exogenous parameters only. We can define the following measure of creditors' welfare, keeping only the endogenous variable (s) and the interesting parameters (σ, θ_1)

$$WR = (\theta_1 s - \sigma t)(1-t)$$

Creditors' payoff in the least-costly separating and in pooling equilibrium is

$$\begin{aligned}WR^{SE} &= \left(p\theta_1^i s_i^{FB} + (1-p)\theta_1^j s^{cri} - \sigma t\right)(1-t) \\ WR^{PE} &= \overline{\theta}_1 s^P (1-t)\end{aligned}$$

2.A.6.3 Banks in a Pooling equilibrium

Payoffs under pooling equilibrium. Given the fundamental thresholds we can characterize the bank's payoff. For any $i \in \{L, H\}$:

$$\Pi_i^P = \int_{\hat{\theta}_i^P}^{\overline{\theta}_2} k - cs^i \mu(\cdot) d\theta_2 = (k - cs^i)\left(\overline{\theta}_2 - \hat{\theta}_i^P\right)$$

Given the formula for the threshold, this can be written as

$$\Pi_L^P(s) = \Pi_L^{FI}(s) + (k - c \cdot s)\left(\frac{sp_H \Delta\theta_1}{1 + 2\sigma}\right)$$
$$\Pi_H^P(s) = \Pi_H^{FI}(s) - (k - c \cdot s)\left(\frac{sp_L \Delta\theta_1}{1 + 2\sigma}\right)$$

The difference between the two types' profit is

$$\Delta\Pi = \frac{2\sigma}{1 + 2\sigma}\left(\theta_1^H - \theta_1^L\right)(k - c \cdot s)s$$
$$\frac{\partial \Delta\Pi}{\partial s} = \frac{2\sigma}{1 + 2\sigma}\left(\theta_1^H - \theta_1^L\right)(k - 2c \cdot s)$$

The profit difference is increasing in s as long as $\frac{k}{2c} > s$ and decreasing thereafter. The total payoff given the prior type distribution is

$$\sum \Pi = (k - c \cdot s)\left(\overline{\theta}_2 - (p_L \hat{\theta}_2^L + p_H \hat{\theta}_2^H)\right)$$

we can rewrite the second term

$$(p_L \hat{\theta}_2^L + p_H \hat{\theta}_2^H) = t - \frac{s(p_L \theta_1^L) + p_H \theta_1^H)}{1 + 2\sigma} - \frac{2\sigma s}{1 + 2\sigma}\left(p_L \theta_1^L + p_H \theta_1^H\right) = t - \left(p_L \theta_1^L + p_H \theta_1^H\right)s$$

Implication: the total payoff to the bank in a pooling equilibrium equals the total profit which would occur with full information, for the given unique intervention level. That is, although agents are unable to distinguish the two types, the bank's ex-ante expected profit is exactly as if they were distinguishable. The only welfare loss (on the bank's side) stems from the fact that in a pooling equilibrium they are unable to set their first-best. This difference can be quantified

$$\Delta\Pi = \Pi(s^\star) - \Pi^s = (k - c \cdot s^\star)(\overline{\theta}_2 + s^\star \theta_1 - t) - (k - c \cdot s)(\overline{\theta}_2 + s\theta_1 - t)$$
$$= k\theta_1 \Delta s - c(\overline{\theta}_2 - t)\Delta s - c\theta_1(\Delta s)^2$$

2.A.6.4 Creditors in a Pooling equilibrium

Substituting pooling thresholds we obtain:

$$
\begin{aligned}
\alpha(\theta_2) &= \frac{\hat{x} - \theta_2 + \sigma}{2\sigma} = \frac{t(1+2\sigma) - \overline{\theta}_1 s - \theta_2}{2\sigma} \\
\theta_2^\star &= t(1+2\sigma) - \overline{\theta}_1 s \\
\theta_{2\star} &= \hat{x} - \sigma = t(1+2\sigma) - \overline{\theta}_1 s - 2\sigma \\
\hat{\theta}_2^i &= t - \theta_1^i s - \frac{sp_j \Delta\theta_1}{1+2\sigma}; \qquad \hat{\theta}_2^j = t - \theta_1^j s + \frac{sp_i \Delta\theta_1}{1+2\sigma}
\end{aligned}
$$

$$
\begin{aligned}
\hat{\theta}_2^i - \theta_{2\star} &= 2\sigma(1-t) + s(\overline{\theta}_1 - \theta_1^i) - \frac{sp_j \Delta\theta_1}{1+2\sigma} = 2\sigma(1-t) + sp_j \Delta\theta_1 \frac{2\sigma}{1+2\sigma} \\
\hat{\theta}_2^j - \theta_{2\star} &= 2\sigma(1-t) + s(\overline{\theta}_1 - \theta_1^j) + \frac{sp_i \Delta\theta_1}{1+2\sigma} = 2\sigma(1-t) - sp_i \Delta\theta_1 \frac{2\sigma}{1+2\sigma} \\
\theta_2^\star - \hat{\theta}_2^i &= 2\sigma t + s(\theta_1^i - \overline{\theta}_1) + \frac{sp_j \Delta\theta_1}{1+2\sigma} = 2\sigma t - sp_j \Delta\theta_1 \frac{2\sigma}{1+2\sigma} \\
\theta_2^\star - \hat{\theta}_2^j &= 2\sigma t + s(\theta_1^j - \overline{\theta}_1) - \frac{sp_i \Delta\theta_1}{1+2\sigma} = 2\sigma t + sp_i \Delta\theta_1 \frac{2\sigma}{1+2\sigma}
\end{aligned}
$$

$$
\alpha(\hat{\theta}_2^i) = t - \frac{sp_j \Delta\theta_1}{1+2\sigma} \quad \alpha(\hat{\theta}_2^j) = t + \frac{sp_i \Delta\theta_1}{1+2\sigma}
$$

$$
\begin{aligned}
\int_{\hat{\theta}_2^i}^{\theta_2^\star} \alpha(\theta_2) d\theta_2 &= \sigma \left(t - \frac{sp_j \Delta\theta_1}{1+2\sigma} \right)^2 \\
\int_{\theta_{2\star}}^{\hat{\theta}_2^i} \alpha(\theta_2) d\theta_2 &= \sigma - \sigma \left(t - \frac{sp_j \Delta\theta_1}{1+2\sigma} \right)^2 \\
\int_{\hat{\theta}_2^j}^{\theta_2^\star} \alpha(\theta_2) d\theta_2 &= \sigma \left(t + \frac{sp_i \Delta\theta_1}{1+2\sigma} \right)^2 \\
\int_{\theta_{2\star}}^{\hat{\theta}_2^j} \alpha(\theta_2) d\theta_2 &= \sigma - \sigma \left(t + \frac{sp_i \Delta\theta_1}{1+2\sigma} \right)^2
\end{aligned}
$$

Substituting to the welfare function:

$$
\begin{aligned}
\mathbb{E}u^i &= \frac{1}{2\eta} \left(t(\theta_{2\star} - \underline{\theta}_2) + t\sigma + (\theta_2^\star - \hat{\theta}_2) - \sigma(t - \frac{sp_j \Delta\theta_1}{1+2\sigma})^2 + (\overline{\theta}_2 - \theta_2^\star) \right) \\
&= \frac{1}{2\eta} \left((\overline{\theta}_2 - t\underline{\theta}_2) + (t^2(1+2\sigma) - t\overline{\theta}_1 s - 2\sigma t) + t\sigma - \sigma(t - \frac{sp_j \Delta\theta_1}{1+2\sigma})^2 - \hat{\theta}_2 \right) \\
\mathbb{E}u^j &= \frac{1}{2\eta} \left(t(\theta_{2\star} - \underline{\theta}) + t\sigma + (\theta_2^\star - \hat{\theta}_2) - \sigma(t - \frac{sp_j \Delta\theta_1}{1+2\sigma})^2 + (\overline{\theta}_2 - \theta_2^\star) \right) \\
&= \frac{1}{2\eta} \left((\overline{\theta}_2 - t\underline{\theta}_2) + (t^2(1+2\sigma) - t\overline{\theta}_2 s - 2\sigma t) + t\sigma - \sigma(t + \frac{sp_i \Delta\theta_1}{1+2\sigma})^2 - \hat{\theta}_2 \right)
\end{aligned}
$$

2.B Appendix B - Generalizations

2.B.1 Generalized functional forms

We leave the payoffs to the creditors and the bank unchanged, and generalize two components of the model: (1) The cost of action is a generic function $c(s)$ with the assumptions $\frac{\partial c(s)}{\partial s} > 0$ and $\frac{\partial^2 c(s)}{\partial s^2} > 0$ that is the cost function is increasing and strictly concave in the signal s. This implies the net payoff for the bank is

$$U(\theta, \alpha) = \begin{cases} k - c(s) & \text{if} \quad \mathcal{R}(\theta_1, \theta_2, s, \alpha) \geq 0 \\ 0 & \text{if} \quad \mathcal{R}(\theta_1, \theta_2, s, \alpha) < 0 \end{cases}$$

(2) The regime change function is a generic function $\mathcal{R}(\theta_1, \theta_2, s, \alpha)$ with assumptions

$$\frac{\partial \mathcal{R}}{\partial \theta_1} > 0; \qquad \frac{\partial \mathcal{R}}{\partial \theta_2} > 0; \qquad \frac{\partial \mathcal{R}}{\partial s} > 0; \qquad \frac{\partial \mathcal{R}}{\partial \alpha} < 0; \qquad \frac{\partial^2 \mathcal{R}}{\partial \theta_1 \partial \theta_2} = 0; \qquad \frac{\partial^2 \mathcal{R}}{\partial s \partial \theta_1} > 0;$$

where the last condition plays the role of a single crossing condition. All other components of the model remain unchanged.

Symmetric information benchmark: The indifference condition remain unchanged since the generalization does not alter the creditors' problem. The mass of agents who attack is:

$$\alpha = \frac{\hat{x} - \hat{\theta}_2 + \epsilon}{2\epsilon} = t$$

The failure condition solves - after substituting $\alpha = t$ -

$$\mathcal{R}(\theta_1, \theta_2, s, t) = 0$$

from which

$$\hat{\theta}_2 = f(t, \theta_1, s)$$
$$\hat{x} = f(t, \theta_1, s) + 2\sigma t - \sigma$$

We want to establish that the thresholds are decreasing in both θ_1 and s. This can be shown using *implifit function theorem* (IFT)

$$\frac{\partial \hat{\theta}_2}{\partial s} = -\frac{\partial \mathcal{R}/\partial s}{\partial \mathcal{R}/\partial \theta_2} = -\frac{[+]}{[+]} < 0$$
$$\frac{\partial \hat{\theta}_2}{\partial \theta_1} = -\frac{\partial \mathcal{R}/\partial s}{\partial \mathcal{R}/\partial \theta_2} = -\frac{[+]}{[+]} < 0$$

This generalizes the result that higher signal as well as higher type decreases threshold, therefore improves stability and increases probability of survival of the regime.

Optimal action: The expected profit for any risk-management choice s is:

$$\pi(\theta_1, s) = \rho\,(k - c(s))$$

Optimum risk management is determined by the FOC

$$\frac{\partial \pi(\theta_1, s)}{\partial s} = (k - c(s))\,\frac{\partial \rho}{\partial s} - \rho\frac{\partial c(s)}{\partial s} = 0$$

from the formula for $\rho = \frac{1}{2\eta}(\overline{\theta}_2 - \hat{\theta}_2)$

$$\frac{\partial \rho}{\partial s} = -\frac{1}{2\eta}\frac{\partial \hat{\theta}_2}{\partial s} > 0$$

so the FOC of optimality using IFT is (to simplify notation: $\partial_x \mathcal{R} = \frac{\partial \mathcal{R}}{\partial x}$)

$$\begin{aligned} -\frac{k}{2\eta}\frac{\partial \hat{\theta}_2}{\partial s} &= \frac{\partial c(s)}{\partial s} \\ \frac{k}{2\eta}\frac{\partial_s \mathcal{R}}{\partial_{\theta_2}\mathcal{R}} &= \frac{\partial c(s)}{\partial s} \end{aligned}$$

The question is how $s^\star$ changes with θ_1. For that we use IFT on the FOC. The second term doesn't change with θ_1. The first term

$$\partial\frac{\partial_s \mathcal{R}}{\partial_{\theta_2}\mathcal{R}}/\partial\theta_1 = \frac{\partial_{s\theta_1}\mathcal{R}\partial_{\theta_2}\mathcal{R} - \partial_{\theta_2\theta_1}\mathcal{R}\partial_s\mathcal{R}}{\partial^2_{\theta_2}\mathcal{R}}$$

$$\partial\frac{\partial_s \mathcal{R}}{\partial_{\theta_2}\mathcal{R}}/\partial s = \frac{\partial_{ss}\mathcal{R}\partial_{\theta_2}\mathcal{R} - \partial_{\theta_2 s}\mathcal{R}\partial_s\mathcal{R}}{\partial^2_{\theta_2}\mathcal{R}}$$

using the assumptions $\partial_{\theta_2\theta_1}\mathcal{R} = 0$ and $\partial_{ss}\mathcal{R} \geq 0$ and $\partial_{\theta_2}\mathcal{R} > 0$ and $\partial_{\theta_2 s}\mathcal{R} = 0$ this implies

$$\begin{aligned} \mathrm{sign}[\partial[RHS]/\partial\theta_1] &= \mathrm{sign}[\partial_{s\theta_1}\mathcal{R}] = [+] \\ \mathrm{sign}[\partial[RHS]/\partial s] &= \frac{[+]}{[+]} \\ \partial c(s)/\partial(s) &\geq 0 \end{aligned}$$

Another application of IFT implies for the case $\mathrm{sign}[\partial_{\theta_2 s}] > 0$

$$\frac{\partial s^\star}{\partial\theta_1} = -\frac{\partial_{\theta_1}FOC}{\partial_s FOC} = -\frac{+}{[+]-[+]} > 0$$

as long as $\mathcal{R}$ is not too convex in s compared to the cost function. Precisely, it must be the case that

$$\frac{\partial_{ss}\mathcal{R}\partial_{\theta_2}\mathcal{R} - \partial_{\theta_2 s}\mathcal{R}\partial_s\mathcal{R}}{\partial^2_{\theta_2}\mathcal{R}} = \frac{\partial_{ss}\mathcal{R}}{\partial_{\theta_2}} < \frac{\partial c'(s)}{\partial s}$$

2.B.2 Separating equilibrium

Summary: For any arbitrary function $\mathcal{R}(\boldsymbol{\theta}, s, \alpha)$ we derive a first-order approximation of failure thresholds off-the-equilibrium path, i.e. when agents act as if the type is j while the actual type is i. Alternatively, this threshold can also be seen as a decomposition of the difference between equilibrium thresholds for type i and j into two components - a direct effect through regime change function, and an indirect effect which is purely belief-based. The main result is that as the noise becomes more precise ($\sigma \to 0$), the indirect effect dominates the direct effect, pushing the off-equilibrium threshold towards the other type's threshold. Intuitively, this increases the potential gains for the low type from mimicking the high type.

We know that

$$\alpha(\hat{x}, \theta_2) = \frac{\hat{x} - \theta_2 + \sigma}{2\sigma}$$
$$\hat{x} = \hat{\theta}_2 + 2\sigma t - \sigma$$

The relevant thresholds on- and off-equilibrium path respectively solve

$$\begin{aligned}
\mathcal{R}(\theta_1^i, \theta_2, s^i, \alpha(\hat{x}^i, \theta_2)) &= 0 \qquad [\hat{\theta}_2^i] \\
\mathcal{R}(\theta_1^i, \theta_2, s^j, \alpha(\hat{x}^j, \theta_2)) &= 0 \qquad [\hat{\theta}_2^{i.j}] \\
\mathcal{R}(\theta_1^j, \theta_2, s^j, \alpha(\hat{x}^j, \theta_2)) &= 0 \qquad [\hat{\theta}_2^j] \\
\mathcal{R}(\theta_1^j, \theta_2, s^i, \alpha(\hat{x}^i, \theta_2)) &= 0 \qquad [\hat{\theta}_2^{j.i}]
\end{aligned}$$

Using implicit function theorem (IFT) we can calculate the derivative of the full-information fundamental threshold with respect to θ_1

$$\frac{\partial \hat{\theta}_2}{\partial \theta_1} = -\frac{\partial_{\theta_1}\mathcal{R}}{\partial_{\theta_2}\mathcal{R}} = \frac{\frac{\partial \mathcal{R}}{\partial \theta_1} + \frac{\partial \mathcal{R}}{\partial \alpha}\frac{\partial \alpha}{\partial \hat{x}}\frac{\partial \hat{x}}{\partial \hat{\theta}_2}\frac{\partial \hat{\theta}_2}{\partial \theta_1}}{\frac{\partial \mathcal{R}}{\partial \theta_2} + \frac{\partial \mathcal{R}}{\partial \alpha}\frac{\partial \alpha}{\partial \theta_2}}$$

This derivative can be used to approximate the change when type changes from θ_1^i to θ_1^j in 'full information' model, keeping signal s fixed. This defines a natural decomposition

$$\frac{\partial\hat{\theta}_2}{\partial\theta_1} = \underbrace{\frac{\frac{\partial\mathcal{R}}{\partial\theta_1}}{\frac{\partial\mathcal{R}}{\partial\theta_2} + \frac{\partial\mathcal{R}}{\partial\alpha}\frac{\partial\alpha}{\partial\theta_2}}}_{\text{Direct effect}} + \underbrace{\frac{\frac{\partial\mathcal{R}}{\partial\alpha}\frac{\partial\alpha}{\partial\hat{x}}\frac{\partial\hat{x}}{\partial\hat{\theta}_2}\frac{\partial\hat{\theta}_2}{\partial\theta_1}}{\frac{\partial\mathcal{R}}{\partial\theta_2} + \frac{\partial\mathcal{R}}{\partial\alpha}\frac{\partial\alpha}{\partial\theta_2}}}_{\text{Indirect effect}}$$

The difference between $\hat{\theta}_2^i$ and $\hat{\theta}_2^{i.j}$ is only through the second term, while the difference between $\hat{\theta}_2^{i.j}$ and $\hat{\theta}_2^j$ is only through the first term, which allows us to write

$$\hat{\theta}_2^{i.j} = \hat{\theta}_2^j + \frac{\frac{\partial\mathcal{R}}{\partial\theta_1}}{\frac{\partial\mathcal{R}}{\partial\theta_2} + \frac{\partial\mathcal{R}}{\partial\alpha}\frac{\partial\alpha}{\partial\theta_2}}\Delta\theta_1 = \hat{\theta}_2^j + \frac{\frac{\partial\mathcal{R}}{\partial\theta_1}}{\frac{\partial\mathcal{R}}{\partial\theta_2} - \frac{1}{2\sigma}\frac{\partial\mathcal{R}}{\partial\alpha}}\Delta\theta_1$$

$$\hat{\theta}_2^{i.j} = \hat{\theta}_2^i - \frac{\frac{\partial\mathcal{R}}{\partial\alpha}\frac{\partial\alpha}{\partial\hat{x}}\frac{\partial\hat{x}}{\partial\hat{\theta}_2}\frac{\partial\hat{\theta}_2}{\partial\theta_1}}{\frac{\partial\mathcal{R}}{\partial\theta_2} + \frac{\partial\mathcal{R}}{\partial\alpha}\frac{\partial\alpha}{\partial\theta_2}}\Delta\theta_1 = \hat{\theta}_2^i - \frac{\frac{1}{2\sigma}\frac{\partial\mathcal{R}}{\partial\alpha}\frac{\partial\hat{x}}{\partial\hat{\theta}_2}\frac{\partial\hat{\theta}_2}{\partial\theta_1}}{\frac{\partial\mathcal{R}}{\partial\theta_2} - \frac{1}{2\sigma}\frac{\partial\mathcal{R}}{\partial\alpha}}\Delta\theta_1$$

From here we can immediately generalize two results of the main text

1. As $\sigma \to 0$ the direct effect goes to zero, thereby $\hat{\theta}_2^{i.j} \to \hat{\theta}_2^j$
2. Due to assumptions on the derivatives, both effect terms are always positive, guaranteeing $\hat{\theta}_2^i \leq \hat{\theta}_2^{i.j} \leq \hat{\theta}_2^j$ (i and j interchangeable, $=$ only in limiting cases).

Example: Apply to example $\mathcal{R}$ the derivatives give precise solutions because of linearity.

$$\frac{\hat{\theta}_2}{\hat{\theta}_1} = \frac{s + \frac{1}{2\sigma}(s)}{1 + \frac{1}{2\sigma}} = s$$

which implies

$$\hat{\theta}_2^j - \hat{\theta}_2^i = s\Delta\theta_1$$

which is correct because from the formula $t - \theta_1^j s - t + \theta_1^i s = s\Delta\theta_1$. From the decomposition

$$\hat{\theta}_2^{i.j} - \hat{\theta}_2^j = \frac{s}{1 + \frac{1}{2\sigma}}\Delta\theta_1 = \frac{2\sigma s}{1 + 2\sigma}\Delta\theta_1$$

$$\hat{\theta}_2^i - \hat{\theta}_2^{i.j} = \frac{\frac{1}{2\sigma}s}{1 + \frac{1}{2\sigma}}\Delta\theta_1 = \frac{s}{1 + 2\sigma}\Delta\theta_1$$

which is exactly what we have in the main text.

2.B.3 Pooling equilibrium

Equation 1 is unchanged:

$$\hat{x}^P = 2\sigma t - \sigma + p_i \hat{\theta}_2^i + p_j \hat{\theta}_2^j$$

Equation 2 can only be expressed implicitly:

$$\mathcal{R}(\theta_1, \theta_2, s, \alpha(\hat{x}, \theta_2)) = 0$$

Write the Taylor-approximation of the difference between $\hat{\theta}_2^i$ (full-information threshold) and $\hat{\theta}_2^{iP}$ (pooling threshold).

$$\frac{\partial \hat{\theta}_2}{\partial \hat{x}} = -\frac{\frac{\partial \mathcal{R}}{\partial \hat{x}}}{\frac{\partial \mathcal{R}}{\partial \theta_2}} = -\frac{\frac{\partial \mathcal{R}}{\partial \alpha}\frac{\partial \alpha}{\partial \hat{x}}}{\frac{\partial \mathcal{R}}{\partial \theta_2} + \frac{\partial \mathcal{R}}{\partial \alpha}\frac{\partial \alpha}{\partial \theta_2}} = -\frac{\frac{1}{2\sigma}\frac{\partial \mathcal{R}}{\partial \alpha}}{\frac{\partial \mathcal{R}}{\partial \theta_2} - \frac{1}{2\sigma}\frac{\partial \mathcal{R}}{\partial \alpha}} > 0$$

$$\hat{\theta}_2^{iP} = \hat{\theta_2}^i + \frac{\partial \hat{\theta}_2}{\partial \hat{x}} \Delta \hat{x}$$

where

$$\Delta \hat{x} = \hat{x}^P - \hat{x}^i = (1-p)\Delta \hat{\theta}_2 = -p_j s \Delta \theta_1 = -s(\overline{\theta}_1 - \theta_1^i)$$

The derivative goes to zero as $\frac{1}{\sigma} \to 0$. That implies, low precision pushes pooling thresholds (respectively, profits) towards the full-information thresholds. This is consistent with the 'smoothing out differences' intuition. The results are symmetric for type j. First define

$$\Delta \hat{x}^j = \hat{x}^{jP} - \hat{x}^j = p_i \hat{\theta}_2^i - p_i \hat{\theta}_2^j = -p\Delta \hat{\theta}_2 = ps\Delta\theta_1 = -s(\overline{\theta}_1 - \theta_1^j)$$

then approximate the average thresholds

$$\overline{\theta}_2^P = p_i \theta_2^{iP} + p_j \theta_2^{jP} \cong \overline{\hat{\theta}_2} + \frac{\partial \hat{\theta}_2}{\partial \hat{x}} \overline{\Delta \hat{x}} = t - s\overline{\theta}_1 + \frac{\partial \hat{\theta}_2}{\partial \hat{x}}(-s)(\overline{\theta}_1 - \overline{\theta}_1) = t - s\overline{\theta}_1$$

This is again the same result as for the specific case, but now proven generally!

Example: Using this result on the example $\mathcal{R} = \theta_2 + \theta_1 s - \alpha = 0$ we have

$$\frac{\partial \hat{\theta}_2}{\partial \hat{x}} = \frac{1}{2\sigma + 1}$$
$$\hat{\theta}_2^{iP} = t - \frac{s\overline{\theta}_1}{1+2\sigma} - \frac{2\sigma}{1+2\sigma} s\theta_1^i$$

2.B.4 Regulation

The incentive compatibility constraints for an arbitrary level of (minimum) pooling

$$\pi_{i.j}(s) \leq \pi_i(s_p) \qquad [IC_i]$$
$$\pi_j(s) \geq \pi_{j.i}(s_p) \qquad [IC_j]$$

These inequalities say that s must not be profitable for the low-type even if it is believed to be of high type, but must be profitable for the high-type, if otherwise is believed to be of low-type.

All 4 functions in the IC's reach zero at $s^{lim} = c^{-1}(k)$. Suppose our conjecture is correct. This is equivalent with the existence of some $s_p < s$ such that both inequalities are satisfied with equality at the same critical s, which just solves IC_i and just fail to solve IC_j. This critical situation, if exist, therefore characterized equivalently by a pair $\{s_p; s\}$ where $s_p < s$ which solves both inequality with equality. Consider all possible pairs of $\{s_p; s\}$, not just those which solves the IC's, and define the following profit-differences

$$\Delta\pi_i(s) = \pi_j(s) - \pi_{i.j}(s) := \varphi(s)$$
$$\Delta\pi_j(s) = \pi_{j.i}(s) - \pi_i(s) := \varphi(s)$$

Viewed as a function of an arbitrary s, these are both *the same* concave functions with zeroes at $\{0, s^{lim}\}$, which we denoted by $\varphi(s)$. If a critical pair $\{s_p; s\}$ exists, it must satisfy $\varphi(s_p) = \varphi(s)$, call this the 'critical condition'. In addition, if a candidate critical pair $\{s_p, s\}$ satisfies the critical condition and at least one IC, say $[IC_i]$, then automatically satisfies the other one. This follows from the definition of 'critical condition'.

We know that both $s_i^{cri}(s_p) \to s^{lim}$ from below, and $s_i^{cri}(s_p)$ increases in s_p and its image is a compact interval $[s_{i.2}^{cri}, s^{lim}]$. It is obvious to see that the solution of the critical condition (that is an explicit expression of pairs s_p, s viewed as a function), $s(s_p)$ is continuously decreasing in s_p with the image $[s^{max}, s^{lim}]$ where s^{max} is a value which maximizes $\varphi(s)$, the difference between profits. A sufficient condition for the existence of a critical s_p, s is that $s^{max} < s_{i.2}^{cri}$, for which, we only have to prove that $s_j^\star < s^{max}$. This is tivial (although some more formality would be nice here) as long as $\pi(s = 0) > \pi(s = s^{lim}) = 0$

Chapter 3

Current Account and Overdraft Pricing in Retail Banking

3.1 Introduction

In this research we investigate equilibrium pricing of personal current accounts (checking accounts) and overdrafts in retail banking. In some markets, especially in the US and the UK, the predominant personal current account (PCA) scheme is the so-called 'free-if-in-credit' (FIIC) pricing, sometimes loosely referred to as 'free banking'. Under this price schedule, banks charge zero monthly or regular fee for the access to the account and for basic services. However, accounts are usually bundled together with an overdraft-service — essentially a short-term borrowing facility — which allows the customer to go into debit in her account, incurring extensive charges in the form of service fees or interest payments. These charges are sizeable enough to make overdrafts one of the most expensive forms of short-term customer credit. The purpose of this research is to better understand the equilibrium consequences of some underlying market frictions, such as customer naiveté and information asymmetry, on prices, profits, and on the market structure.

Many observers point out that FIIC-pricing is consistent with a cross-subsidy across business lines as well as across various groups of customers, with significant potential welfare consequences. According to the common narrative, hidden and expensive overdraft fees exploit poorer households — more likely in need of extra liquidity — to support more sophisticated, wealthier clients, and to generate industry rents. This view has generated a massive policy debate on the potential role and shape of regulatory intervention.

Two structural characteristics of retail banking have been identified as the root cause of this *exploitative innovation*[1]: first, the overdraft facility, being closely linked to payment services which is the primary use of a current account, can be seen as an add-on product, and — at least to the extent as the aftermarket is technologically or otherwise linked to the base product — banks have market power and can exert monopoly pricing on the market for overdrafts. Second, this market power is exacerbated by customers' behavioural biases, such as their limited ability to comprehend contractual terms ex ante (Gabaix and Laibson [2006]), limited attention to track their own account usage (Grubb [2014]), or the lack of switching behaviour due to perceived or genuine switching costs.

In turn, FIIC pricing leads to two main policy concerns. First, although on the surface 'free-banking' appears to be the result of strong competition, it may mask non-competitive distortions and market power, leading to inefficient outcomes. Second, as FIIC pricing essentially amounts to below-marginal-cost base-good and above-marginal-cost add-on prices, to the extent different groups of customers differ in their use of the two services, it leads to potential distributional concerns and welfare effects. To be specific, if overdraft users tend to be poorer households more likely in financial trouble and in the need of short-term financing, FIIC-pricing implies a wealth-transfer 'from the poor to the rich', leading to an overall reduction in social welfare.

In this paper we describe a model of overdraft pricing which captures some relevant characteristics of the retail banking sector to illustrate how FIIC-pricing can emerge as an equilibrium phenomenon. We define *FIIC equilibrium* as a market equilibrium in which a base product (current account) is priced at its lower bound at least by some players, while connected services — specifically, overdrafts — are sold to customers at a mark-up (above marginal cost). In our benchmark model we consider a two-stage duopoly model, where two banks compete for a continuum of customers by selling two products, a *'personal current account'* (PCA) in the first, and a connected *overdraft*-facility in the second stage.

Deviating from much of the theoretical literature, we assume that competition for customers is possible even in the second stage. This captures the idea that customers' lock-in is not perfect in a financial services context. Indeed, the nature of an existing relationship to one's bank is distinctively different from a typical industrial aftermarket situation - the most cited example being the market for printers and cartridges - where technological linkages tie customers strongly to the primary supplier, while high initial investment costs prevent switching to another provider. Strictly speaking there is no such technological reason to link payment services (the

[1]The term 'exploitative innovation' is based on Heidhues et al. [2016a]

primary market) to the provision of short-term customer credit (the aftermarket), yet the market shows similar characteristics: lack of switching behaviour, and 'exploitative' aftermarket prices.

Traditionally, there has been two oft-cited sources of banks' market power over existing customers. First, naive customers' lack of attention to other deals after establishing the first relationship may lead to banks' exploitation of those customers. Second, relationship banks obtain relevant information regarding their customers' creditworthiness, which may lead to adverse selection on the market for customer credit. Both issues are at the centre of recent regulatory innovations: requiring insider banks to disclose information on request may moderate the adverse selection problem, while automatized switching services and price-comparison websites potentially decrease psychological and material costs associated with low switching behaviour, which is closely related to the concept of customer naiveté. Regulatory interventions also induced changes in banks' behaviour: instead of 'shrouding' the information on fees and other contractual terms which potentially impose additional costs, they tend to heavily advertise up-front, thereby compete much more strongly on overdraft terms as well.

We begin our analysis with a model of banking competition featuring customer naiveté. In the first stage, banks compete for a group of naive and sophisticated customers. In the second stage, they charge different overdraft prices for own customers and for customers of the other bank. Banks have some market power over customers (this is captured by a Hotelling-model) in the first stage, but the lock-in is imperfect in the second-stage, so overdraft fees emerge endogenously in a Bertrand-competition. Naive customers are sticky after their initial choice of bank, and won't consider the possibility of switching when an overdraft facility is required. This creates market power, which distorts standard competitive pricing outcome.

Our results demonstrate that, despite the possibility of competition in the second-stage, the presence of naive customers turns to be an important source of market power and economic profit. Specifically, we show that for an arbitrary low number of naive customers, prices deviate from marginal-costs, and the unique Nash-equilibrium of the Bertrand-game on the overdraft market is a mixed-strategy Nash-equilibrium, where both insider and outsider banks earn positive profit. This makes overdraft a profitable business, and induces competition in the first stage (on the market for PCA) to expand market share. When primary markets are sufficiently competitive, there exist a symmetric 'FIIC-equilibrium' for a significant subset of the parameter space - specifically, FIIC-pricing can prevail even with relatively modest number of naive customers, and on highly competitive markets.

In the rest of the paper we extend the baseline setup with adverse selection, and show how the combination of adverse selection and customer naiveté, being the two most important frictions on banking markets, affect equilibrium pricing, profits, and customer behaviour. The presence of adverse selection makes it more difficult for the outsider to enjoy the benefits on the aftermarket. As a consequence, the profits will be tilted towards the bank's role as an insider. This increases the incentives to obtain more customers on the primary market, so it decreases first-period PCA prices even further — making it even more likely that an FIIC-pricing prevails as an equilibrium. In our model differences in the mass of myopic customers is not the only possible explanation anymore for the observed differences across countries: 'free banking' may or may not develop depending on the severity of adverse selection, or the extent of primary market competition as well.

Our approach is a novelty in the theoretical banking literature from a crucial aspect, which we emphasize here again. Although various models of add-on pricing have been proposed recently to explain exploitative overdraft pricing practices (Gabaix and Laibson [2006], Armstrong and Vickers [2012], Heidhues et al. [2016a]), these models usually stipulate ex-post monopoly power on the aftermarket, in our case, on the market for overdrafts. Although that assumption is a good approximation of the short-term behaviour and first-time overdrafting — in which case the main behavioural bias to consider is that inattentive customers, being unaware of their potential overdraft usage[2], indeed face a de-facto monopolist seller — it cannot provide an explanation for the long-term lack of switching behaviour and the persistence of exploitation of a certain group of customers. Furthermore, recent observed pricing schemes by major UK banks seem to be consistent with the presence of a (limited) aftermarket competition.

We believe that the new modelling approach is useful to properly address these recent changes in retail banking. Due to increased pressure from customer protection groups and policy makers, it is hard to argue anymore that overdraft fees are 'shrouded', as the literature usually assumes, and not transparent at the time of contracting — in contrast, in the UK, overdraft conditions seem to be one of the most heavily advertised selling points of PCA's offered to new customers. Comparison websites, easy switching services make the implicit assumption of full monopoly power on the aftermarket less-and-less tenable. At the same time, this seem to have little effect (for now) on the exploitative charges on those who are permanently in overdraft. Our model takes a first step towards a more robust explanation of the large-scale existence of these schemes even with competitive aftermarkets.

[2] Models of bill-shock regulation: Grubb [2014]

The dominant theoretical explanation usually invoked to explain FIIC-style pricing is the framework by Gabaix and Laibson [2006], which combines aftermarket pricing of an add-on product with a homogeneous base-good, in the presence of naive customers[3]. In a typical equilibrium of this modelling tradition a bank is facing a perfectly inelastic demand function on the overdraft market, whenever a specific group of customers decides to participate and consume the overdraft facility. Naive customers always 'decide to use' the add-on, while sophisticated customers substitute away if they find it too expensive. The bank must decide whether it wants to serve only naive, or all types of customers, and then set the appropriate monopoly price which just preserves the applicable participation constraints of the customers. Aftermarket monopoly profits are then competed away in the base-good market, if possible[4]. If there is a sufficiently large number of sophisticated customers, the bank's optimal decision is to serve all customers, and the add-on price will be relatively modest (bounded by the participation constraint of sophisticated customers). In this equilibrium, there is no cross-subsidy across the groups of customers, and there is no FIIC-pricing[5]. With relatively large number of myopes, however, a bank decides to serve only myopic customers, leading to excessive add-on price, low base-price, and possibly severe cross-subsidy. As sophisticated customers must exert effort which is more costly than producing the overdraft service to substitute away, the equilibrium is inefficient. If the base-price hits the lower bound, aftermarket profits cannot be eliminated by competition on the base-good market, and this will be the source of monopoly profit for banks. This equilibrium, referred to as 'shrouding equilibrium' in Gabaix and Laibson [2006][6] is loosely identified with FIIC-banking in the relevant banking literature.

We believe there are two shortcomings of this theory which makes it problematic to apply directly to the retail banking sector, and which calls for additional research work. First, the theory would predict FIIC-pricing whenever the mass of naive customers (which is routinely identified with financial illiteracy) is sufficiently large. However, there is no evidence that UK and US customers are significantly different in this respect from their European peers. While FIIC with overdraft is common and standard package in the UK, it is not particularly widespread in countries

[3]The GL2006 framework is directly applied to the UK retail banking sector by Armstrong and Vickers [2012], and extended among others in Heidhues et al. [2016a], Heidhues et al. [2016b].

[4]The early literature on aftermarket pricing (Shapiro [1994]) recognizes that monopoly profits are distributed back to the customers if prices can be decreased sufficiently on a competitive base-good market. However, a lower bound on the price can prevent competing away these profits (Heidhues et al. [2016a]). This lower bound can arise endogenously in certain markets (Miao [2010]).

[5]However, above-marginal-cost add-on and below-marginal-cost base good price still prevails.

[6]In this equilibrium firms are incentivized to hide information regarding the add-on prices and conditions in the first stage, to make the myopic population as large as possible

like France or Italy. Second, identifying shrouding equilibrium with FIIC-banking is not straightforward. It is true, that the shrouding equilibrium is disproportionally costly for naive customers and might lead to positive profits. However, these two frictions do not typically occur within the same parameter range. The base-good price decreases continuously and linearly with the mass of naive customers, so 'FIIC' and the associated positive profits would be observed only with very large number of myopes. The relative 'exploitation', however, is particularly concerning when the number of naive customers is moderate (close to the critical threshold), because redistributing profits in the form of lowered base-good price does not ease yet the extra burden caused by the excessive add-on price.

3.1.1 Related literature

Behavioural theories for overdrafts: The concept of shrouding equilibrium with naive customers put forward in Gabaix and Laibson [2006] has formed the basis of several theoretical and empirical contributions. In two strongly related recent papers, Heidhues et al. [2016a] and Heidhues et al. [2016b] develop new insights which are directly applicable to financial markets. First, they show that a binding price floor in the market for the base-product leads to positive economic profit in a shrouding equilibrium. Then, they demonstrate how these profits, emerging in a shrouding equilibrium, can lead to various welfare-reducing market practices: in Heidhues et al. [2016a], firms choose between investing into value-enhancing or 'exploitative' innovations, and the paper shows that the incentives for the latter are stronger, as — in contrast to the former one — it raises other participants' incentives to maintain the shrouding equilibrium. The authors note that binding price floor leading to positive profits is likely to hold for consumer financial products. In a similar vein, Heidhues et al. [2016b] argues that socially wasteful products are more likely to survive on the market, as in this case it is more likely that shrouding prevails as a unique equilibrium and guarantees the positivity of profits, which then cannot be competed away due to binding price floors. They interpret these findings in the context of financial markets as an explanation for (i) why banks invest effort to develop complex pricing practices (like overdraft) which can maintain exploitation, and (ii) why seemingly inferior (expensive, active) mutual funds can survive. Our research loosely links to these papers by showing that the additional profit which can be obtained on the overdraft market is limited by the presence of adverse selection.

Grubb [2014] follows a distinct, but related modelling approach and emphasizes another behavioural bias, customers' inability to closely track their 'usage' of the base product (i.e. spending from your account), which leads to a surprise ('bill-shock') overdraft usage. Instead of this surprise-effect, our naive customers proxy rather customers' sticky behaviour on the long term, despite potentially significant benefits from switching[7].

The paper is also related to recent research on the possibility of price discrimination based on customer naiveté, in the context of financial services (Kosfeld and Schüwer [2017]), or more generally (Heidhues and Kőszegi [2017]). The first of these, Kosfeld and Schüwer [2017] shows that educating customers may have unintended welfare consequences if naivete-based discrimination is possible on the aftermarket. Heidhues and Kőszegi [2017] also focuses on the welfare aspects in a more general settings, and provide conditions where naiveté-based discrimination negatively affects welfare. In our model we exclude naiveté-based price discrimination, but allow for price discrimination based on other payoff-relevant and observable qualities.

There are other approaches to explain exploitation of customers' bounded rationality. Carlin [2009] describes a model of oligopoly banking where complexity arises as part of an equilibrium pricing structure, as it can be a source of market power. The endogenous complexity choices of banks determine the mass of uninformed players on the market, and all banks share the demand of uninformed players, independently of the price. The unique equilibrium is in mixed strategies, so the model also predicts price dispersion. However, the model is static and does not capture either the relationship-banking nature or the cross-subsidy characteristics of FIIC accounts, which is at the center of our research agenda.

To my knowledge, the only theoretical model which specifically addresses the overdraft fees (or more generally, contingent charges) in the retail banking sector is Armstrong and Vickers [2012], building on a simplified version of Gabaix and Laibson [2006]. This version allows the authors to transparently focus on possible regulatory interventions, such as price caps, overdraft warnings or restrictions of negative balances altogether. We focus rather on a different type of equilibrium, where fundamental assumptions of the model are altered.

Empirically, Alan et al. [2018] provides direct evidence of the exploitation of naive customers from Turkey using an overdraft market experiment: by randomizing messages which affect consumers' attention in various ways they demonstrate unawareness of prices and underestimation of future usage, and that firms indeed respond to this behaviour with shrouding. Adams [2017] confirms on US data that overdraft prices consist a significant part of banks' revenue, which tend to be larger in low-income regions. He finds total expenditure to be higher for "poorer, younger

[7]The Competition and Market Authority in the UK estimates *('Retail banking market investigation', Final report, 9 August 2016)* the potential annual gains from switching in the UK retail banking sector to be more than £*4bn*.

or less educated" populations — a finding which seems to be consistent with the theory, if those groups can be characterized as relatively more naive.[8]

Adverse selection in Banking: The adverse selection component of our model largely follows the ideas presented in the seminal paper by Sharpe [1990] which introduces the concept of relationship banking. This paper illustrates how information which is created during customer-bank relationship can generate ex-post monopoly power and lead to an endogenous emergence of switching barriers. Von Thadden [2004] points out a mistake in the original paper and proves the existence of a unique mixed-strategy equilibrium with partial informational lock-in where switching occurs with positive probability. Rajan [1992] makes a similar point regarding the importance of relationships, but his work is based on moral hazard instead of adverse selection, and also emphasizes possible benefits of control by insider bank. Our benchmark adverse selection model in Section 4 can be considered as a much simplified, backbone version of Sharpe's model which still delivers the same equilibrium structure and characteristics.

Price dispersion: Our work predicts price dispersion both as a result of adverse selection and customer naiveté. As mentioned before, the former is a well-known result in the literature. The latter, to my knowledge, has not been explicitly addressed in a theoretical contribution before. However, the common behavioural notion of customer naiveté / myopia is a close cousin of some earlier research ideas with boundedly rational customers. To start with, customers who do not switch despite a cheaper price available on the market are reminiscent to the 'uninformed' customers who do not search in Varian [1980], which also initiates price dispersion. Burdett and Judd [1983] shows that this can happen even with identical and rational agents and homogeneous search costs, with nonsequential search or with noisy sequential search. Narasimhan [1988] studies brand-loyalty in a duopoly framework and shows that the presence of loyal customers initiates price dispersion, due to similar reasons, as loyal customers won't consider buying a competitor's product.

The rest of the paper is structured as follows. In Section 2 we introduce our model in its most complete form. Section 3 derives equilibrium for a reduced setup where customer naiveté without adverse selection is addressed. Section 4 describes benchmark results for the model version with adverse selection but without customer naiveté, while Section 5 derives equilibrium of the full model. Section 6 is a discussion of the results and possible policy consequences, while Section 7 concludes.

[8] Other notable empirical studies on overdraft fees and customer naiveté are Stango and Zinman [2009], Stango and Zinman [2014], Morgan et al. [2012], Melzer and Morgan [2015], Williams [2016].

3.2 Model setup

Consider the following two-stage competition game representing a retail banking market. Two banks ($j \in \{1, 2\}$) are offering a uniform personal current account (PCA) in stage one, and a liquidity facility (overdraft) in stage two. A unit measure of customers with a fixed demand for one product in each stage are located uniformly over the interval $[0, 1]$, while the two banks are located at the two opposite ends of the interval. Let ℓ_j denote the location of Bank j, that is, $\ell_1 = 0$ and $\ell_2 = 1$. In the first stage only, a customer i located at $\gamma_i \in [0, 1]$ incurs a transportation cost $\tau \times d_{ij}$ where τ is exogenous constant, while $d_{ij} = |l_j - \gamma_i|$ is customer i's distance from Bank j. In the second stage, products are homogeneous, and there is no transportation cost.

At $t = 1$ each bank j simultaneously announces a fee $p_j \geq 0$ for the current account, which is observable for all customers. Following the price announcement, customers choose exactly one bank, which we then refer to as their *insider bank*. In period $t = 2$ all customers are hit by a liquidity shock, and want to consume the overdraft service. In this period they can decide whether to switch to the other bank, so their action space is {"STAY", "SWITCH"}. As noted at t=2 the consumers do not incur a transport cost in selecting their bank. This assumption implies that the overdraft service provided – the lending of money to cover a liquidity need – is homogeneous across the banks.

Customers ex-ante differ in two aspects: (i) sophistication and (ii) profitability (riskiness). Sophisticated customers (type S, fraction $1 - \alpha$) are fully rational, and in both stages choose the bank with lower expected total outlay, including transportation costs. If the expected payment is equal, they choose randomly in the first stage, or remain with the insider in the second stage. Naive customers (type N, fraction α) fail to predict their future demand for the overdraft facility, therefore they base their decision only on the observable first-period prices. Furthermore, in the second stage, they do not consider the possibility of switching and always stay with their insider bank.[9]

A proportion of β of customers yields low-profit to the bank (type L), while a fraction $(1 - \beta)$ is highly profitable (type H). The information on profitability (but not on naiveté) is observed by the insider bank during the first stage of the customer-banking relationship, and is contractable in the second period. Profitability is captured concisely by parameters $r_L < r_H$, representing exogenous revenues generated from the relationship in the overdraft-stage. Formally, let the customers'

[9] One possible justification is that without planning they become involuntary overdraft users.

type space be $\Theta := \{S, N\} \times \{L, H\}$. The joint probability mass function of a customer type $\theta \in \Theta$ is completely characterized by parameters α and β, which are treated as exogenous throughout the analysis. We assume that the two types (naiveté and profitability) are independently distributed.

Customers' valuation of the account is ν, while on the overdraft service is ν'. We will assume that these valuations are sufficiently high so that all customers decide to consume in all equilibria which we focus on in the main text.

Bank j can condition the price of the overdraft service for its insider customers on profitability, but not on sophistication. Therefore, overdraft is offered at a price $\phi_{in}^{L,(j)}$ and $\phi_{in}^{H,(j)}$ for type L (type H) insider customers of Bank j respectively. Furthermore, Bank j offers the service for customers of the other bank *(outsider customers)* at an on-demand price $\phi_{out}^{(j)}$. Other than that, it is not possible to price-discriminate based on the location of the customer. Overdraft fees are exogenously capped at $\overline{\phi}$. Marginal cost of opening and maintaining a PCA is normalized to 0, while overdraft is offered at a marginal cost c_{od}.

We solve the game for *Perfect Bayesian Equilibrium.* Let $\boldsymbol{p} \in \mathbb{R}_+^2$ denote the PCA price vector, $\boldsymbol{\phi} \in \mathbb{R}_+^6$ the overdraft fee vector, while $a_1 \in \{1, 2\}$ and $a_2 \in \{\text{"STAY"}, \text{"SWITCH"}\}$ denote customers' decisions in the first and the second stage respectively. Then

Definition 1 *A Perfect Bayesian Equilibrium of the overdraft-pricing game consists of*

1. *A first-period PCA price offer by the two banks:* $\boldsymbol{p}^\star := \{p_1^\star; p_2^\star\}$*;*
2. *Customers' decision over which bank to choose in the first stage:* $a_1^\star(\theta, \gamma, \boldsymbol{p}) : \Theta \times [0, 1] \times \mathbb{R}_+^2 \to \{1, 2\}$*;*
3. *A second-period overdraft fee by the two banks:* $\{\phi_{in}^{L,j\star}; \phi_{in}^{H,j\star}, \phi_{out}^{j\star}\}$ *for all* $j \in \{1, 2\}$
4. *Customers' decision whether to switch:* $a_2^\star(\theta, a_1(\theta, \gamma, \boldsymbol{p}), \boldsymbol{\phi}) : \Theta \times \{1, 2\} \times \mathbb{R}^6 \to \{\text{"STAY"}, \text{"SWITCH"}\}$.

where decisions are sequentially rational:
(i) each bank maximizes profit at each stage, and
(ii) customers' decisions a_1 *and* a_2 *are optimal given their appropriate subjective beliefs regarding equilibrium prices.*

Notice that the full model is a 4-stage strategic-form game, but it is useful to interpret it as a game which unfolds in two *periods*: in period 1 banks offer current accounts and customers engage with exactly one of them, while in period 2 banks offer an add-on liquidity service and customers decide whether to switch. This game is solvable by backward induction: first, we determine 2nd-period equilibrium taking the prices and customer-bank relationships from period 1 as given. Then, we consider deviations in the first stage, and analyse customers' reaction to such deviations. We find PBE if no such profitable deviation exists.

It is useful to introduce some further notation to simplify algebra. First, let us define *net cost* from serving a customer in the overdraft-stage as:

$$\begin{aligned} c^L &= c_{od} - r_L \\ c^H &= c_{od} - r_H \\ c^{LH} &= c_{od} - \beta r_L + (1-\beta) r_H \end{aligned}$$

The new variable c conveniently captures all exogenous components of the profit from a *unit mass* of customers who are respectively low, high or both (reflecting population probabilities) types in the second stage.

In the following, we first derive two special cases: (i) the case without adverse selection ($\beta = 0$), and (ii) the case without customer myopia ($\alpha = 0$).

3.3 Overdraft with customer naiveté

We start by deriving the equilibrium for the special case where $\beta = 0$. Clearly, in this case there is no information asymmetry between insider and outsider, so there is no adverse selection problem. This version illustrates how the presence of naive customers with an add-on product distorts standard competitive pricing results.[10] As with $\beta = 0$ we have only one (high) type of customers, we can drop profitability indices and simplify notation: $\{\phi_{in}^{(j)}, \phi_{out}^{(j)}\}$ denotes overdraft fees by Bank j for insider and outsider respectively, and c denotes net costs.

3.3.1 Second-stage equilibrium

Suppose that Bank j starts with a mass γ_j of customers[11], and within these customers, the percentage of naive types is α_j. Notice first that the second-period game is separable into two distinct components: banks compete for the *insider customers* of Bank j through the choice of $\phi_{in}^{(j)}$ and $\phi_{out}^{(-j)}$, for all $j \in \{1, 2\}$. Given this observation, we can formulate the subgame from an arbitrary bank's perspective as follows: (i) two banks jointly announce an overdraft fee $\{\phi_{in}^{(j)}, \phi_{out}^{(-j)}\}$, and (ii) all

[10]Note that without naive customers the pure-strategy Nash-equilibrium of the pricing game is trivially a competitive equilibrium.

[11]Note that these are not necessarily the customers 'closest' to Bank j.

sophisticated customers of j, that is, $\gamma_j(1-\alpha_j)$ mass of customers switch if and only if $\phi_{out}^{(-j)} < \phi_{in}^{j}$. Since indices are clearly pinned down when looking at the problem from the perspective of any of the two banks, we can omit for further analysis.

Charging the maximum value $\overline{\phi}$ is always a feasible strategy for both insider and outsider banks. Whenever the outsider charges $\overline{\phi}$, it loses the competition for sophisticated customers with certainty, and makes zero profit on overdrafts. Therefore, in any equilibrium, outsider must only be willing to offer a fee which leads to a nonnegative profit from overdraft business, that is,

$$\phi_{out} \geq c$$

The insider bank is bounded by a similar incentive compatibility constraint. Charging $\overline{\phi}$ is always a feasible strategy, and even if at this fee it loses the competition for sophisticated customers with certainty, it obtains the following profit[12]:

$$\pi_{in}^{0} := \alpha(\overline{\phi} - c)\gamma$$

This quantity can be regarded as insider's *minimax payoff*: in any proposed equilibrium, its profit from overdrafts must be *at least* π_{in}^{0}. Now suppose that the equilibrium is such that for a sufficiently low offer ϕ insider wins the price competition with probability 1. Even in this case, the offer ϕ must satisfy the inequality

$$(\phi - c)\,\gamma \geq \alpha(\overline{\phi} - c)\gamma$$

Let ϕ_{in}' denote the value of ϕ which solves the corresponding equation, and we refer to this as the *incentive-compatible overdraft fee*[13]. By rearranging we obtain:

$$\phi_{in}' := \alpha\overline{\phi} + (1-\alpha)c \tag{3.1}$$

According to this expression, the incentive-compatible overdraft fee is a weighted average of the maximum fee and the break-even fee where the weights are the mass of naive (resp. sophisticated) customers. We proceed with a formal proof that there is no pure-strategy Nash-Equilibrium.

Lemma 1 *If $\alpha > 0$, no Pure-strategy Nash-equilibrium exists.*

Proof. At stage 2 we have Bertrand-competition. Insider's offer must be such that $\phi_{in} \in [\phi_{in}', \overline{\phi}]$, while outsider's offer is $\phi_{out} \in [c, \overline{\phi}]$. Given $\alpha > 0$, we have $\phi_{in}' > c$.

[12]We don't know yet whether this is a binding constraint, but we will see that later.
[13]We don't know yet whether it is played in equilibrium by any of the players.

Then outsider's best response to any offer ϕ_{in} is $\phi_{in} - \epsilon$ with some $\epsilon > 0$. Insider's best response to any offer $\phi_{out} \in [\phi'_{in}, \overline{\phi}]$ is ϕ_{out}, and to any offer $\phi_{out} < \phi'_{in}$ is $\overline{\phi}$. The mapping has obviously no fixed point, so there is no PSNE. ■

Notice that the Lemma is true for every $\alpha > 0$ but fails for $\alpha = 0$. For $\alpha = 0$ we have $\phi'_{in} = c$ and its easy to see that playing ϕ'_{in} by both players is the unique PSNE of the game, which is also the competitive (zero-profit) outcome. This verifies our previous claim that without customer naiveté the unique equilibrium is the competitive outcome. It is the presence of customer naiveté which initiates price dispersion in the overdraft market.

As there is no pure-strategy Nash-equilibrium, we look for a mixed-strategy equilibrium ('MSNE'). Suppose both outsider and insider mix according to CDFs F_{out} and F_{in}, with support $[\underline{F}_{out}, \overline{F}_{out}]$ and $[\underline{F}_{in}, \overline{F}_{in}]$ respectively. The following lemma establishes boundaries for the distributions.

Lemma 2 *The supports of the CDFs F_{out} and F_{in} must satisfy*

1. $\underline{F}_{out} = \underline{F}_{in} = \phi'_{in}$
2. $\overline{F}_{in} = \overline{\phi}$

Proof. See Appendix 3.A.1 ■

The expected payoff from any action which is played in equilibrium must be equal over the range of equilibrium actions (*'indifference condition'*). The MSNE is established in two steps: first, we apply the indifference condition to outsider's strategies to derive F_{in}, then, apply it to insider's strategies to derive F_{out}.

Lemma 3 *The insider bank is mixing according to a continuous distribution $F_{in}(\phi)$ with support $[\phi'_{in}, \overline{\phi}]$, and is placing a probability mass α on $\overline{\phi}$, where*

$$F_{in}(\phi) = 1 - \alpha \frac{\overline{\phi} - c}{\phi - c} \tag{3.2}$$

Proof. First, we establish that there is no probability mass by the insider on ϕ'_{in}. Suppose $Pr[\phi_{in} = \phi'_{in}] > 0$. Then outsider can charge $\phi_{out} = \phi_{in} - \epsilon$ for some $\epsilon > 0$ and win all sophisticated customers with probability 1. By continuity, he would make strictly larger profit than by winning at ϕ'_{in} with some probability strictly less than 1. Contradiction to equilibrium.

Because of that, by playing the lower boundary ϕ'_{in} the outsider wins, all sophisticated customers switch with probability 1, and the bank's profit is

$$\pi^0_{out} = (1-\alpha)\left(\phi'_{in} - c\right)\gamma$$

For any higher bid $\phi > \phi'_{in}$ it must be that

$$\pi_{out}(\phi) = \underbrace{Prob(\phi < \phi_{in})}_{\text{O wins}} * \underbrace{(1-\alpha)\pi(\phi)\gamma}_{\text{profit}|\text{win}} + \underbrace{Prob(\phi \geq \phi_{in})}_{\text{I wins}} * 0 = \pi^0_{out}$$

where the function $\pi(\phi) := \phi - c$ denotes the profit from serving a unit measure of customers with overdraft. The indifference condition in equilibrium is therefore:

$$(1 - F_{in}(\phi))\,(1-\alpha)\,(\phi - c)\,\gamma = (1-\alpha)\left(\phi'_{in} - c\right)\gamma$$

After substituting the value of ϕ'_{in} from Equation (3.1), we obtain insider's CDF:

$$F_{in}(\phi) = \frac{\phi - \alpha\overline{\phi} + (1-\alpha)(-c)}{\phi - c} = 1 - \alpha\frac{\overline{\phi} - c}{\phi - c}$$

This CDF satisfies $F_{in}(\phi'_{in}) = 0$ and $F_{in}(\overline{\phi}) = 1 - \alpha$, which implies that insider is mixing over $[\phi'_{in}, \overline{\phi}]$ and is placing a probability mass of α on $\overline{\phi}$. ■

Lemma 4 uses insider's indifference property to derive outsider's fee dispersion:

Lemma 4 *The outsider bank is mixing according to a continuous distribution $F_{out}(\phi)$ with support $[\phi'_{in}, \overline{\phi}]$ where*

$$F_{out}(\phi) = \frac{1}{1-\alpha} - \frac{\alpha}{1-\alpha}\frac{\overline{\phi} - c}{\phi - c} \tag{3.3}$$

Proof. We know from the proof of Lemma 2 that there is no probability mass on $\overline{\phi}$ by the outsider, implying

$$\pi^0_{in} = \alpha\left(\overline{\phi} - c\right)\gamma$$

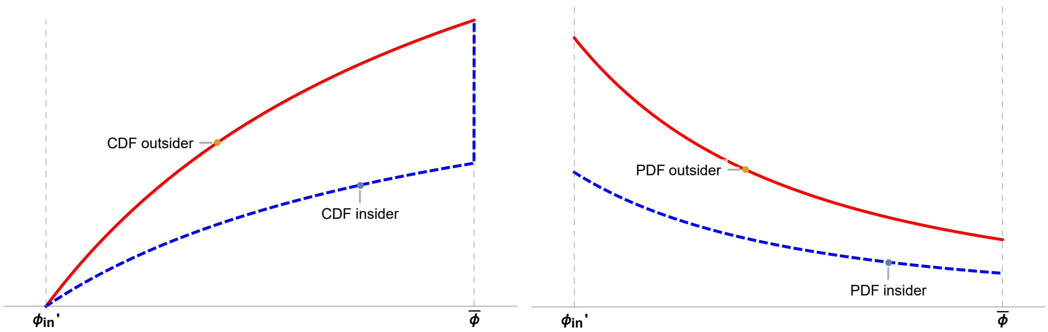
The indifference property implies

$$\underbrace{Prob(\phi \leq \phi_{out})}_{\text{I wins}} * \underbrace{(\phi - c)\,\gamma}_{\text{profit from all}} + \underbrace{Prob(\phi > \phi_{out})}_{\text{O wins}} * \underbrace{\alpha(\phi - c)\gamma}_{\text{profit from myopes}} = \pi^0_{in}$$

Leading to the following equality:

$$(1 - F_{out}(\phi))\,(\phi - c)\,\gamma + F_{out}(\phi)\alpha\,(\phi - c)\,\gamma = \alpha\left(\overline{\phi} - c\right)\gamma$$

Figure 3.1: Overdraft fee dispersion
Customer naiveté



This defines the CDF for *outsider*:

$$F_{out}(\phi) = \frac{1}{1-\alpha} - \frac{\alpha}{1-\alpha}\frac{\overline{\phi}-c}{\phi-c}$$

This CDF satisfies $F_{out}(\phi'_{in}) = 0$ and $F_{out}(\overline{\phi}) = 1$ ■

From the two Lemmas, immediately follows the characterization of the unique MSNE, which we summarize in Theorem 1 and illustrate in Figure 3.1.

Theorem 1 *The unique mixed-strategy Nash equilibrium of the overdraft pricing game with $\beta = 0$ is as follows: both insider and outsider mix between $[\phi'_{in}, \overline{\phi}]$, outsider according to F_{out} as defined in Equation 3.3, and insider according to F_{in} as defined in equation 3.2. Insider is placing a positive mass of α on $\overline{\phi}$.*

3.3.2 Equilibrium characterization

3.3.2.1 Overdraft fee dispersion

We can make probabilistic statements regarding expected fee and switching behaviour in equilibrium. For easy readability all proofs are in appendix. First, we establish the probability that outsider wins.

Lemma 5 *The probability that outsider wins is a linear function of the mass of naive customers (α)*

$$Prob[\phi_{out} < \phi_{in}] = \frac{1+\alpha}{2}$$

Proof. See Appendix 3.A.2 ■

As $\alpha \to 0$, there is no 'naive distortion' effect, and competition guarantees that outsider wins with probability 1/2. As the mass of naive customers increases, the insider places more-and-more emphasis on 'exploitation' (i.e. $\phi = \overline{\phi}$) which increases the probability of winning by the outsider. As $\alpha \to 1$ the probability mass on $\overline{\phi}$ goes to 1, and outsider wins with probability 1. It is useful to mention that the total probability that outsider wins is the sum of two probabilities: when the banks play mixture ($\frac{1-\alpha}{2}$) and when the insider plays the mass-point and outsider always wins (α). The appendix also shows formally that *conditional on both mixing*, the outsider wins with probability 1/2.

Next, we calculate expected overdraft fees conditional on switching or staying. This is an interesting characterization on its own, and also an important ingredient for later calculations.

Lemma 6 *Conditional on switching, the (sophisticated) customers are expected to pay an overdraft fee of*

$$\mathbb{E}[\phi_{out}|\phi_{out} < \phi_{in}] = c + \frac{2\alpha}{1+\alpha}\left(\overline{\phi} - c\right)$$

Conditional on remaining with insider, the (sophisticated) customers are expected to pay

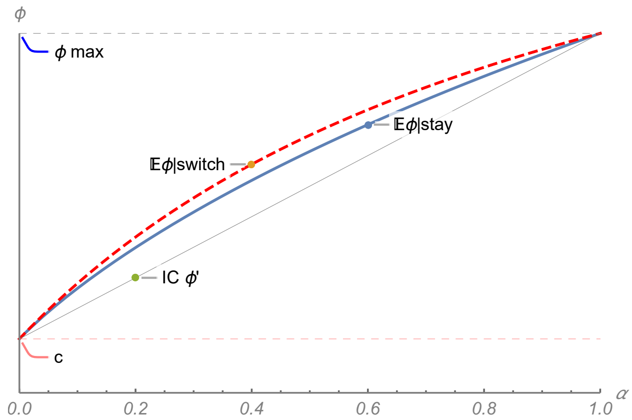
$$\mathbb{E}[\phi_{in}|\phi_{in} \leq \phi_{out}] = c + \frac{2\alpha}{1-\alpha}\left(1 + \frac{\alpha \ln[\alpha]}{1-\alpha}\right)\left(\overline{\phi} - c\right)$$

Proof. See Appendix 3.A.3 ■

The lemma is illustrated in Figure 3.2. Blue (solid) line is expected overdraft fee *conditional on staying*, as a function of α, while red (dashed) line is expected fee conditional on switching. Sophisticated customers are facing with an actual realization from the distribution, not this expected value. As $\alpha \to 0$, overdraft fees converge to the competitive outcome $\phi_{out} = \phi_{in} = c$[14] and both insider and outsider obtain zero-profit from overdraft business. With $\alpha > 0$, banks randomize, while insider is placing more-and-more weight on larger fees, increasing the expected fee conditional on winning. This allows the outsider bank also to (probabilistically) raise its offers. It is straightforward to interpret the expressions for expected fees in Lemma 6 as a mark-up pricing formula: in the presence of naive customers, banks can increase overdraft fees in expectation relative to the break-even, competitive fee (c). The mark-up is proportional to the difference between the maximum price and the break-even price, is increasing in α, and is always higher for the outsider

[14]Both CDF's converge to a mass-point on c

Figure 3.2: Conditional expected overdraft fees
Customer naiveté



customer[15].

3.3.2.2 Second period profit

From the indifference conditions, we know that bank j's expected profit from its role as an insider and as an outsider is

$$\pi_{in}^{(j)} = \gamma_j \alpha_j \left(\overline{\phi} - c\right)$$
$$\pi_{out}^{(j)} = \gamma_{-j}(1-\alpha_{-j}) \left(\phi_{in}^{\prime -j} - c\right)$$

Notice that the 'outsider' profit comes entirely from sophisticated customers who switch, while insider profit consists of two components: sophisticated customers who stay in equilibrium, and naive customers who always stay by assumption. In the Appendix we derive a decomposition of the two latter terms. The following Lemma provides results for the overall bank profit, calculated as an appropriately weighted sum of the components defined above.

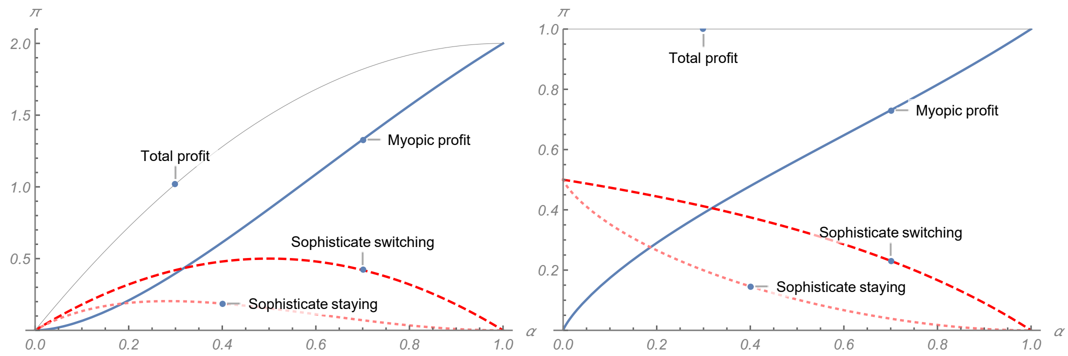
Lemma 7 *For arbitrary values of γ_j and α_j, Bank j's profit is*

$$\pi^{(j)} = \left[\alpha - \frac{(\alpha - \alpha_j\gamma_j)^2}{1-\gamma_j}\right] \left(\overline{\phi} - c\right) \tag{3.4}$$

In two specific cases we have simpler expressions. Whenever the naive customers are evenly distributed among the two banks in the second stage, that is, $\alpha_1 = \alpha_2 = \alpha$,

[15] This can be proven by showing that their difference is decreasing and is 0 as $\alpha \to 1$.

Figure 3.3: Expected profits as a function of α
Customer naiveté



(a) Decomposition of bank's profit (b) % decomposition of bank's profit

banks obtain the following overall profit in the second stage:

$$\pi^{(j)} = \alpha(1-\alpha) + \gamma_j \alpha^2$$

Whenever also the market share's are equal, that is $\gamma_1 = \gamma_2 = 1/2$, *the banks obtain the following second-period profit*

$$\pi^{(j)} = \frac{1}{2}\alpha\,(2-\alpha)\left(\overline{\phi} - c\right)$$

This symmetric equilibrium profit can be decomposed as

$$\begin{aligned}
\mathbb{E}\pi^{myop} &= \alpha^2\,(1-\ln[\alpha])\left(\overline{\phi}-c\right)\\
\mathbb{E}\pi^{soph,switch} &= (1-\alpha)\,\alpha\left(\overline{\phi}-c\right)\\
\mathbb{E}\pi^{soph,stay} &= \alpha\,(1-\alpha+\alpha\ln[\alpha])\left(\overline{\phi}-c\right)
\end{aligned}$$

Profits in the most general case are increasing in market share whenever $2\alpha_j > \gamma_j\alpha_{-j}$. *When the naive customers are evenly distributed, profits always increase in market share, in the maximum fee* ($\overline{\phi}$), *and decrease in net costs* (c).

Proof. See Appendix 3.A.4 ■

Figure 3.3a illustrates sources of expected profits as a function of α from myopic (blue,solid) and from sophisticated customers who switch (red,dashed) and stay (pink, dotted), while Figure 3.3b shows the percentage distribution of those profits.

3.3.3 First-stage equilibrium

3.3.3.1 Equilibrium first-period prices

Naive customers: As naive customers only take into account the first-period announced price vector $\boldsymbol{p}$, they follow a threshold strategy: a naive customer chooses Bank 1 if and only if she is located at $\gamma < \gamma^N$, where γ^N is determined by the indifference condition:

$$p_1 + \tau\gamma^N = p_2 + \tau(1 - \gamma^N)$$

The value of the threshold is:

$$\hat{\gamma}^N = \frac{1}{2} + \frac{\Delta p}{2\tau} \tag{3.5}$$

Sophisticated customers: Sophisticated customers predict second-period equilibrium overdraft fees to calculate expected payment. As the equilibrium expected fee conditional of choosing Bank j is the function of the mass of myopic customers within Bank j, the following is always an equilibrium:

Lemma 8 *For any PCA price vector $\boldsymbol{p}$ such that $\hat{\gamma}^N(\boldsymbol{p}) \in (0,1)$ there exist an equilibrium of the induced 3-stage subgame where (i) both sophisticated and naive customers follow a common threshold defined in Equation* (3.5) *— now denoted as $\hat{\gamma}$ —, so that only customers with $\gamma_i < \hat{\gamma}$ choose Bank 1, and (ii) second-period overdraft fees are determined according to Proposition 1, with $\alpha_1 = \alpha_2 = \alpha$.*

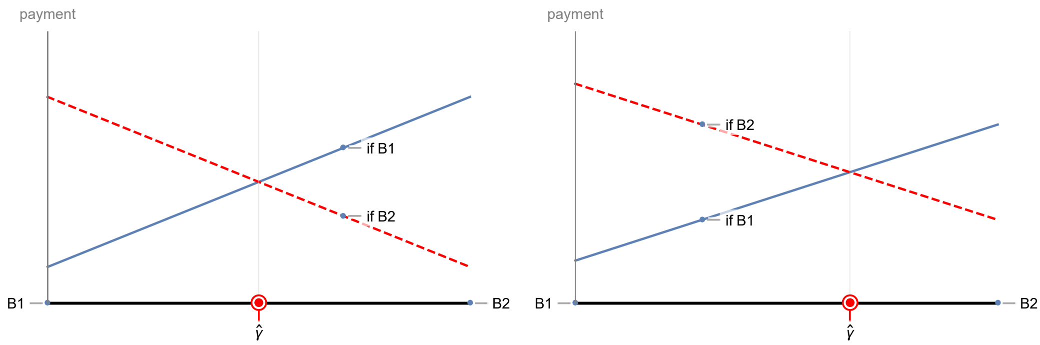
Proof. Whenever $\alpha_1 = \alpha_2 = \alpha$, according to Lemma 6, the expected second-period overdraft fee for sophisticated customers if choosing Bank 1 or Bank 2 is exactly equal, as it only depends on the fraction of naive customers within the bank. Therefore, if they predict that in equilibrium $\alpha_1 = \alpha_2$, they will base their decision only on first-period prices. Consequently, their decision will be identical to that of naive customers, following a threshold strategy with $\hat{\gamma} = \hat{\gamma}^N$, which justifies the belief that $\alpha_1 = \alpha_2$. ■

As a consequence, in this equilibrium, decreasing price always increases market share. This is illustrated in Figure 3.4, where blue (solid) line represents expected payment for sophisticated customers when choosing Bank 1, while red (dashed) is expected payment from choosing Bank 2.

Bank 1's total (ex-ante expected) profit, using Lemma 7, is

$$\Pi_1 = \hat{\gamma}p_1 + \pi^1_{t=2} = \hat{\gamma}(p_1,p_2)p_1 + \left(\alpha(1-\alpha) + \hat{\gamma}(p_1,p_2)\alpha^2\right)\left(\overline{\phi} - c\right) \tag{3.6}$$

Figure 3.4: Equilibrium market share
Customer naiveté



(a) ...with $p_1 = p_2$ (b) ...with $p_1 < p_2$

For any given p_2, the optimal choice of p_1 is given by the first-order condition:

$$\frac{\partial \Pi_1}{\partial p_1} = \frac{\partial \hat{\gamma}}{\partial p_1} p_1 + \hat{\gamma} + \alpha^2 \left(\overline{\phi} - c\right) \frac{\partial \hat{\gamma}}{\partial p_1} = 0$$

Note that

$$\frac{\partial \hat{\gamma}}{\partial p_1} = -\frac{1}{2\tau}$$

therefore:

$$\frac{-p_1}{2\tau} + \left(\frac{1}{2} + \frac{p_2 - p_1}{2\tau}\right) - \frac{\alpha^2 \left(\overline{\phi} - c\right)}{2\tau} = 0$$
$$p_1 = \frac{1}{2}\left(p_2 + \tau - \alpha^2 \left(\overline{\phi} - c\right)\right)$$

Solving for symmetric equilibrium:

$$\frac{p}{t} = \frac{1}{2} + \frac{p}{2t} - \frac{\alpha^2 \left(\overline{\phi} - c\right)}{2t}$$
$$p = \tau - \alpha^2 \left(\overline{\phi} - c\right)$$

Whenever $\tau < \alpha^2(\overline{\phi} - c)$ this is negative. In that case, within the feasible range of parameters, decreasing first-period price would always be a profitable deviation, therefore first-period prices hit the lower bound. Rearranging this expression for α, we obtain that FIIC prevails whenever the mass of naive customers exceed a threshold value, and this threshold *converges to zero* as the first-period competition parameter (τ) converges to perfect competition. Theorem 2 formalizes the result, which we analyse further in the next section.

Theorem 2 *Symmetric equilibrium in the first stage:*

- *Suppose* $\tau \leq \alpha^2 \left(\overline{\phi} - c\right)$. *Then decreasing prices is always a profitable deviation, so the only symmetric equilibrium is* $p_1 = p_2 = 0$

- *Suppose* $\tau > \alpha^2 \left(\overline{\phi} - c\right)$. *Then there exist a unique symmetric-price equilibrium, defined by*

$$p^\star = \tau - \alpha^2 \left(\overline{\phi} - c\right) \tag{3.7}$$

3.3.3.2 Equilibrium profits

Equilibrium profits can be computed using Equation (3.6).

Figure 3.5: Equilibrium profit and PCA price
Customer naiveté

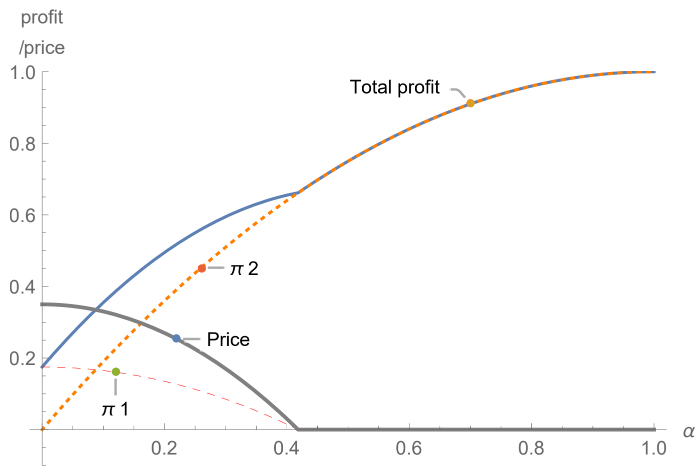


Figure 3.6 illustrates that — in contrast to the Gabaix and Laibson [2006]-tradition — strictly positive profits emerge for any mass of naive customers. The distorted Bertrand-competition in the second stage of our game guarantees that prices are jointly determined in equilibrium so that banks can enjoy the benefits of their market power. In turn, the relative market power in first-stage (τ) and in the second stage (α) then determines whether prices are positive, or hit the lower bound: the more competitive are banks in the first-period, the more likely that a zero-price equilibrium develops, as a result of their competition for market share. In particular, as $\tau \to 0$, so does this threshold α — at the limit, there is zero-price for all possible values of α. However, the equilibrium profits remain positive everywhere, not only when the price bound is hit.

3.4 Overdraft with adverse selection

In this section we switch off customer naiveté, set $\alpha = 0$ (all customers are sophisticated), and introduce the adverse selection problem $\beta > 0$. To simplify analysis, we derive equilibrium by assuming that the insider bank follows a pure-strategy against low-types by charging the maximum (supremum) fee which is offered to the other type in equilibrium. Formally, let

$$\phi_{in}^{L} = \sup\{\phi : F_{in}^{H}(\phi) < 1\}$$

where $F_{in}^{H}(\phi)$ denotes the CDF of the fees insider offers to high-type customers. This assumption intuitively places a simple and plausible restriction on the equilibrium structure, namely that any outcome must be such that low-types are being charged a higher fee than high-types[16]. We introduce the notation $\pi^{L}(\phi) := \phi - c^{L}$, $\pi^{H}(\phi) := \phi - c^{H}$ and $\pi^{LH}(\phi) = \phi - c^{LH}$ for the 2nd-period profit of serving a unit measure of customers of type Low/High/Mixed (according to population probabilities) at an overdraft fee ϕ.

3.4.1 Second-stage equilibrium

The subgame has obviously no pure-strategy Nash-Equilibrium. The first Lemma establishes boundaries for the distributions of the mixed strategies F_{in}^{H} and F_{out}:

Lemma 9 *Both outsider and insider mix over the interval* $[c^{HL}, c^{L}]$. *Outsider's profit must be zero in equilibrium.*
Proof. *See Appendix 3.B.1* ■

Intuitively, the Lemma establishes that because all L-types switch whenever outsider tries to undercut insider by offering a lower fee, outsider's offer cannot be below c^{LH} (participation constraint). On the other hand, the fees cannot be above c^{L}, because outsider cannot make profit in equilibrium.

Outsider's indifference condition: At any fee $\phi < \phi_{in}^{L}$ outsider obtains all low-types, therefore the indifference condition is:

$$\begin{aligned} Pr\left[\phi < \phi_{in}^{H}\right]\left(\beta\pi^{L}(\phi) + (1-\beta)\pi^{H}(\phi)\right) + Pr\left[\phi > \phi_{in}^{H}\right]\beta\pi^{L}(\phi) = 0 \\ \left(1 - F_{in}^{H}(\phi)\right)\left(\beta\pi^{L}(\phi) + (1-\beta)\pi^{H}(\phi)\right) + F_{in}^{H}(\phi)\beta\pi^{L}(\phi) = 0 \end{aligned}$$

[16] The more general proof of the next section proves that the assumption of degenerate distribution for the low-type is without loss of generality, as it arises as a limiting case. For this section, this is assumed for simplicity.

This can be solved for the CDF of insider-H's price dispersion:

$$F_{in}^{H}(\phi) = \frac{1}{1-\beta} - \frac{\beta}{1-\beta}\frac{\Delta c}{\phi - c^H}$$

where $\Delta c = c^L - c^H$. This expression satisfies

$$F_{in}^{H}(c^L) = 1 \qquad \text{and} \qquad F_{in}^{H}(c^{LH}) = 0$$

So the insider is mixing continuously between c^{LH} and c^L, without a mass-point.
Insider's indifference condition: The insider can make a positive profit by charging c^{LH} to the high-types, and winning the competition with probability 1:

$$\pi_{in}^{0} = (1-\beta)\left(c^{LH} - c^H\right) = (1-\beta)\beta\Delta c$$

The indifference condition at any given offer ϕ

$$\begin{aligned} Pr\left[\phi < \phi_{out}\right] \cdot \left((1-\beta)\pi^H(\phi)\right) + Pr\left[\phi > \phi_{out}\right] \cdot 0 &= \pi_{in}^{0} \\ (1-F_{out}) \cdot \left((1-\beta)\left(\phi - c^H\right)\right) + F_{out} \cdot 0 &= (1-\beta)\left(c^{LH} - c^H\right) \end{aligned}$$

This implies the mixed strategies of the outsider in equilibrium:

$$F_{out}(\phi) = \frac{\phi - c^{LH}}{\phi - c^H} = 1 - \frac{c^{LH} - c^H}{\phi - c^H} = 1 - \beta\frac{\Delta c}{\phi - c^H}$$

This CDF trivially satisfies $F_{out}(c^{LH}) = 0$. We can also calculate that

$$F_{out}(c^L) = 1 - \beta$$

which implies that the outsider's mixed strategy is given by $F_{out}(\phi)$ over $[c^{LH}, c^L)$, and a mass point of β on c^L. It might be useful to rewrite F_{out} as[17]

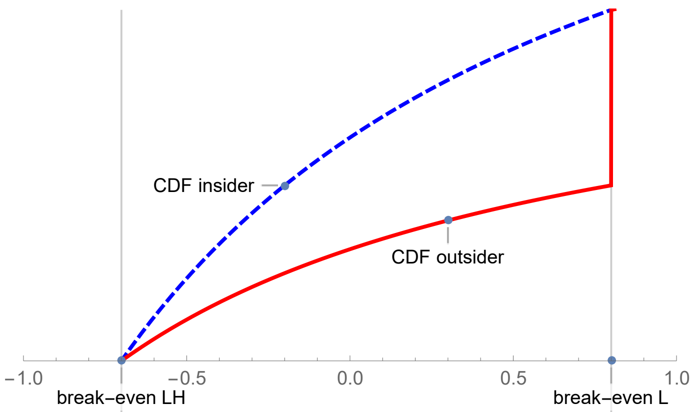
$$F_{out}(\phi) = (1-\beta)F_{in}(\phi)$$

The results are summarized in Theorem 3, which is illustrated in Figure 3.6.

Theorem 3 *The unique mixed-strategy Nash equilibrium of the overdraft-stage of the game with $\alpha = 0$ and $\beta > 0$ is as follows: both insider and outsider mix between*

[17] $\pi^{LH}(c^{LH}) = 0$ by definition. The relationship between F_{in} and F_{out} is obvious from the formulas as well. This simple relationship is reminiscent to von Thadden (2004).

Figure 3.6: Overdraft fee dispersion
Adverse selection



$\left[c^{LH}, c^{L}\right]$, outsider according to F_{out} and insider according to F_{in} as defined below.

$$F^{H}_{in}(\phi) = \frac{1}{1-\beta} - \frac{\beta}{1-\beta}\frac{\Delta c}{\phi - c^{H}} \tag{3.8}$$

$$F_{out}(\phi) = 1 - \beta\frac{\Delta c}{\phi - c^{H}} \tag{3.9}$$

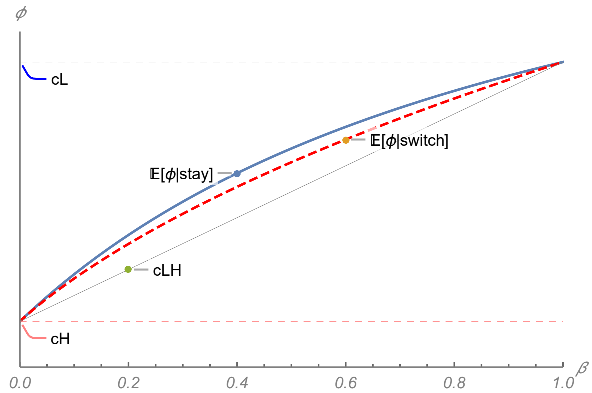
Outsider is placing a strictly positive probability mass of β on c_L (the upper boundary of the distribution), while insider charges c^L to the low-types with probability 1.

3.4.2 Equilibrium characterization

3.4.2.1 Fee dispersion

We calculate expected fees conditional on switching versus staying with the insider bank. A careful examination of the CDF's and the payoff structure reveals similarities with the case of customer naiveté, which is reflected in the following expressions. The probability of winning by insider is a linearly *increasing* function of the mass of *low-type* customers. In the limits, the CDF's become degenerate, and bank's charge the respective fees: as $\beta \to 0$, and banks charge c^L, while they charge c^H as $\beta \to 1$ (see Figure 3.7). The profit is zero in both extreme cases. The formulas below already foreshadows an interesting but intuitive result which is formally established in the next step: possible profit due to adverse selection is the highest when the uncertainty regarding the underlying types is the highest - that is, when $\beta = 1/2$.

Figure 3.7: Conditional expected overdraft fees
Adverse selection



Lemma 10 *The probability that insider (outsider) wins is*

$$Prob[\phi_{in}^H < \phi_{out}] = \frac{1+\beta}{2} \qquad \text{and} \qquad Prob[\phi_{out} < \phi_{in}^H] = \frac{1-\beta}{2}$$

The expected fees conditional on winning are

$$\begin{aligned}
\mathbb{E}[\phi_{in}^H|\phi_{in}^H < \phi_{out}] &= c^H + \frac{2\beta}{1+\beta}\Delta c \\
\mathbb{E}[\phi_{out}|\phi_{out} < \phi_{in}^H] &= c^H + \frac{2\beta}{1-\beta}\Delta c + \frac{2\beta^2}{(1-\beta)^2}\ln[\beta]\Delta c
\end{aligned}$$

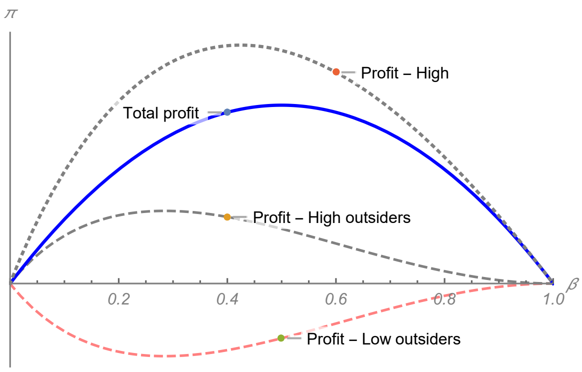
Proof. See Appendix 3.B.2 ■

3.4.2.2 Banks' profits

The outsider makes zero profit: First, verify that in equilibrium the outsider obtains zero profit. In any equilibrium profit comes from two parts which might also be negative: (1) low-types who switch (with certainty) and pay ϕ_{out}, that is $\mathbb{E}[\phi_{out}]$ (the unconditional expected fee) in expectation (2) High-types who switch whenever outsider wins the bid, and pay ϕ_{out}, that is $Pr[\phi_{out} < \phi_{in}]\mathbb{E}[\phi_{out}|\phi_{out} < \phi_{in}]$ in expectation. Denoting π_{out}^L and π_{out}^H the profits from the bank's role as an outsider from low/high types respectively, we can write

$$\pi_{out}^L = \beta\left(\underbrace{(1-\beta)c^H + \beta c^L - \beta\ln[\beta]\Delta c}_{\text{overdraft fee}} - \underbrace{c^L}_{\text{cost}}\right)$$

Figure 3.8: Expected profit
Adverse selection



$$\pi_{out}^{H} = (1-\beta)\left(\frac{1-\beta}{2}\right)\left(\underbrace{c^H + \frac{2\beta}{1-\beta}\Delta c + \frac{2\beta^2}{(1-\beta)^2}\ln[\beta]\Delta c}_{\text{overdraft fee}} - \underbrace{c^H}_{\text{cost}}\right)$$

After straightforward algebra we obtain

$$\pi_{out} := \pi_{out}^{L} + \pi_{out}^{H} = 0$$

Intuitively, the expected loss from low-type customers who switch regardless of pricing conditions (formally because insider is charging a very high fee) is exactly offset by the expected profit on high-type customers who switch because the outsider is winning the price competition. The outsider makes zero profit, therefore all profit in equilibrium comes from its 'role' as an insider bank.

The insider makes positive profit: We have proven that all positive profit (if it exist) comes from the insider, and it therefore must come from high-types when the bank wins the competition. This happens when they mix and the insider wins, and also whenever the outsider plays the probability mass.

$$\begin{aligned}\pi_{in}^{H} &= (1-\beta)\, Pr[\phi_{in}^{H} < \phi_{out}] \times \left(\mathbb{E}[\phi_{in}^{H}|\phi_{in}^{H} < \phi_{out}] - c^H\right) \\ &= (1-\beta)\left(c^H + \frac{2\beta}{1+\beta}\Delta c - c^H\right)\end{aligned}$$

This expression simplifies to the following:

$$\pi_{in}^{H} = \beta(1-\beta)\Delta c$$

Expected profits are illustrated in Figure 3.8. We can also verify that $\pi_{in}^{H} = \pi_{in}^{0}$.

3.5 Overdraft with naiveté and adverse selection

Now we turn to the main model and analyse the interaction between customer naiveté and adverse selection. In this case we pursue an equilibrium in the following generic form: insider sets a price for low-types and high-types according to distributions F_{in}^{L} and F_{in}^{H}, while outsider's price dispersion is given by F_{out}.[18] The model is solved backwards: we start with the equilibrium for the second-stage, then iterate back to the first-stage.

3.5.1 Second-stage equilibrium

Recall that with adverse selection but without customer naiveté in equilibrium insider follows a pure-strategy against low-types by charging c^L with probability 1. As part of the proof for the main theorem of this section, we show first that this simple equilibrium structure cannot survive when naive customers are also present in the economy and $\alpha < \beta$. Then, we demonstrate the existence of a specific type of equilibrium where for $\alpha < \beta$ insider randomizes independently for high-types and low-types over some non-overlapping intervals $[\phi_{in}^{''}, \hat{\phi}]$ and $[\hat{\phi}, \overline{\phi}]$, while outsider randomizes according to a piecewise-defined, continuous distribution over the union of those intervals. The analytical values of $\phi_{in}^{''}$ and $\hat{\phi}$ will be exactly determined. The proof consists of two main steps: first we show that any equilibrium must satisfy this structure; then, we derive overdraft fee dispersion taking the structure as given. In the main text below we give an overview of steps and intuition, while rigorous proof and analytical calculations are relegated to Appendix. The proposed equilibrium is depicted in Figure 3.9.

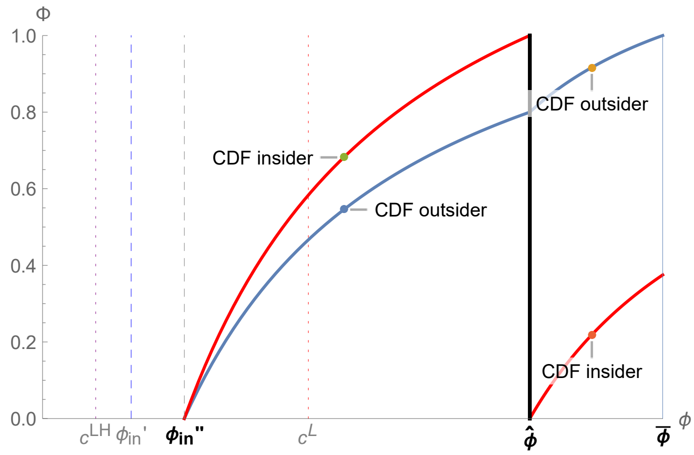
For this section, let us introduce the following notation: $\pi(\theta, \phi, \rho)$ denotes profit from serving customers of type $\theta \in \{L, H, LH\}$ with overdrafts, when overdraft fee is ϕ, and the bank wins with probability ρ. Subscripts $_{in}$ and $_{out}$ refer to profit for insider and outsider respectively. $\underline{\pi}(\theta)$ stands for a minimax payoff according to an alternative strategy.

First, notice that insider bank can always revert to the strategy of serving naive customers only, and charging the fee cap $\overline{\phi}$. This defines two candidate 'minimax' payoffs for insider, for low and for high types independently. It will be clear later that as they are facing with the same outsider distribution F_{out}, only one of these can be binding.

Suppose first that the minimax payoff from low-types, $\underline{\pi}_{in}(L)$ is binding,

[18] The only assumption we make is that insider sets the prices indifferently for low-types and high-types.

Figure 3.9: Overdraft price dispersion
Naiveté and adverse selection



which must be the case if insider places a positive probability mass at $\overline{\phi}$ in F_{in}^L. This pins down the *upper piece* of outsider's piecewise-defined CDF (denoted F_{out}^L[19]) through insider's indifference condition against the low-type customers, that is, insider is willing to offer an arbitrary ϕ to low-types only if its expected profit by charging ϕ equals to its minimax profit:

$$\pi_{in}\left(L, \phi, Pr[\phi \leq \phi_{out}]\right) = \underline{\pi}_{in}\left(L\right)$$

Suppose now that we know the cutoff value of the distributions, $\hat{\phi}$. Then it is possible to formulate the profit for *insider* from serving *high-type* customers. By playing $\hat{\phi}$, insider wins with probability $Pr[\phi_{out} \geq \hat{\phi}]$, which we denote by $\hat{\rho}_{out}$ for simplicity:

$$\hat{\rho}_{out} := 1 - F_{out}^L(\hat{\phi})$$

By playing $\hat{\phi}$, insider's payoff from serving *high*-types is $\pi_{in}(H, \hat{\phi}, \hat{\rho}_{out})$. Due to insider's indifference property, this must be equal to its expected payoff at the lower boundary $\underline{F}$. As by playing $\underline{F}$ insider wins with probability 1, $\underline{F}$ is the value of ϕ which solves

$$\pi_{in}(H, \underline{F}, 1) = \pi_{in}(H, \hat{\phi}, \hat{\rho}_{out})$$

Notice that this equality defines the lower boundary of the support as a function of $\hat{\phi}$. By charging $\underline{F}(\hat{\phi})$ *outsider wins* the competition with probability 1 and obtains both (sophisticated) types, leading to profit $\pi_{out}\left(LH, \underline{F}(\hat{\phi}), 1\right)$. Due to the indifference

[19]Note that although we use parallel notation, F_{in}^L and F_{in}^H denote two different distributions, while F_{out}^L and F_{out}^H is one piecewise-defined CDF!

condition for outsider, this must be equal to its profit when charging $\overline{\phi} - \epsilon$ (with $\epsilon \to 0$). In the latter case outsider wins the low-type customers with the probability that *insider* is placing on $\overline{\phi}$. Let this probability be $\overline{\rho}_{in}$, which can also be written as a function of $\hat{\phi}$, and the associated profit is $\pi_{out}(L, \overline{\phi}, \overline{\rho}_{in}(\hat{\phi}))$.

The next step is to determine the *function* $\overline{\rho}_{in}(\hat{\phi})$. We know that for an arbitrary $\hat{\phi}$ insider's CDF against low-types must reach zero in equilibrium. Therefore, outsider's indifference condition at the two boundaries of the support of F_{in}^L is

$$\pi_{out}(L, \hat{\phi}, 1) = \pi_{out}(L, \overline{\phi}, \overline{\rho}_{in}(\hat{\phi}))$$

This equation allows us to express insider's probability mass at $\overline{\phi}$ as a function of $\hat{\phi}$, so we have an analytical expression for $\overline{\rho}_{in}(\hat{\phi})$.

The final step of the proof combines the lower and upper-part of outsider's indifference condition to solve for the (unique) threshold value $\hat{\phi}$. At the equilibrium value of $\hat{\phi}$ outsider must be indifferent between charging $\underline{F}(\hat{\phi})$ and win the low-types with probability 1, or charging $\overline{\phi} - \epsilon$ and win the low-types with probability $\overline{\rho}_{in}(\hat{\phi})$. That is

$$\pi_{out}(LH, \underline{F}(\hat{\phi}), 1) = \pi_{out}(L, \overline{\phi}, \overline{\rho}_{in}(\hat{\phi}))$$

The Appendix derives the analytical value of $\hat{\phi}$:

$$\hat{\phi}^{\star} = c^L + \frac{\alpha}{\beta}\left(\overline{\phi} - c^L\right)$$

The result is very intuitive, and shows immediately that the equilibrium described here emerges for $\alpha \in (0, \beta)$ only. As $\alpha \to 0$, $\hat{\phi}^{\star} \to c^L$ and we get back the solution for the 'adverse selection only' benchmark case. Whenever $\alpha = \beta$, $\hat{\phi}^{\star} = \overline{\phi}$ and the insider's mixture over low-types becomes degenerate. For every value of $\alpha > \beta$ the fee cap is binding for the low-types, leading to a different type of equilibrium, which is more straightforward to derive (see Appendix for details).

The rest of the proof is straightforward. First, we can use the equilibrium threshold value $\hat{\phi}^{\star}$ to pin down insider's equilibrium profit: at $\hat{\phi}^{\star}$ insider wins the high-types with probability $\left(1 - F_{out}(\hat{\phi}^{\star})\right)$, and this defines its new, modified equilibrium payoff $\pi_{in}^{H,\star}$. We show in Appendix formally that $\pi_{in}^{H,\star} > \underline{\pi}_{in}^{H}$ whenever $\alpha < \beta$, verifying the claim that only the minimax payoff for the low-type is binding (which is indeed binding by construction of the equilibrium). As the equilibrium payoff is pinned down, we can write insider's indifference condition against high-types:

$$\pi_{in}(H, \phi, Pr[\phi < \phi_{out}]) = \pi_{in}^{H,\star}$$

This equation defines the functional form for $F_{out}(\phi)$ on the interval $[\underline{F}, \hat{\phi}^{\star}]$.

Insider's distributions are derived using outsider's indifference condition: at every ϕ, outsider must be indifferent between playing ϕ or its alternative payoff, which is pinned down by the mass-point by insider on $\overline{\phi}$. That leads to the two independent indifference conditions in a straightforward way. Theorem 4 fully characterizes second-stage equilibrium.

Theorem 4 *The equilibrium of the second-stage overdraft pricing game is as follows:*

- *If $\alpha > \beta$, insider charges a fee for low-types according to a degenerate distribution and places all probability mass at $\overline{\phi}$. For high-types, insider mixes over $[\phi'_{in}, \overline{\phi}]$ according to $F^H_{in}(\phi)$, while outsider mixes over the same interval according to $F_{out}(\phi)$. Insider places a positive probability mass $\rho^H_{in} := \frac{\alpha-\beta}{1-\beta}$ on $\overline{\phi}$, where*

$$F^H_{in} = \frac{1}{1-\beta} - \frac{\alpha}{1-\beta}\frac{\overline{\phi}-c^H}{\phi-c^H}$$
$$F_{out} = \frac{1}{1-\alpha} - \frac{\alpha}{1-\alpha}\frac{\overline{\phi}-c^H}{\phi-c^H}$$

- *If $\alpha \leq \beta$, insider mixes over $[\phi''_{in}, \hat{\phi}]$ according to $F^H_{in}(\phi)$ for high-types, and over $[\hat{\phi}, \overline{\phi}]$ according to $F^L_{in}(\phi)$ for low-types. Outsider mixes over $[\phi''_{in}, \overline{\phi}]$ according to F_{out}, without mass-point. Insider places a positive mass ρ_{in} on $\overline{\phi}$, where*

$$F^H_{in}(\phi) = \frac{1}{1-\beta} - \frac{\alpha(\overline{\phi}-c^L)+\beta\Delta c}{(1-\beta)(\phi-c^H)}$$
$$F^L_{in}(\phi) = 1 - \frac{\alpha}{\beta}\frac{\overline{\phi}-c^L}{\phi-c^L}$$

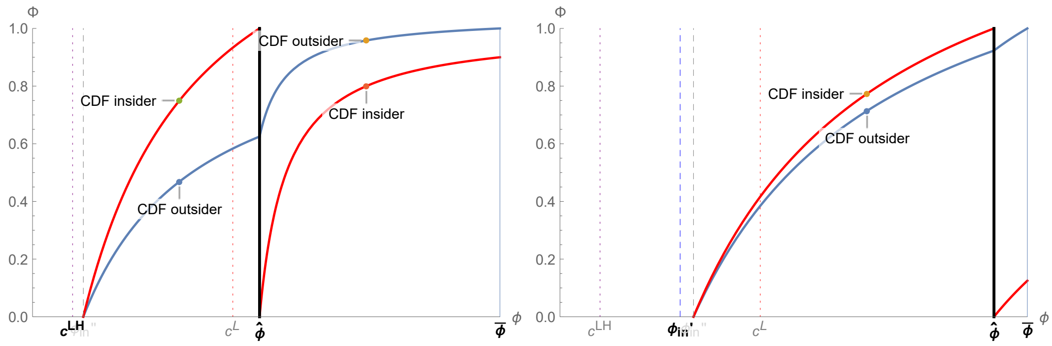
F_{out} is piecewise defined as

$$F_{out}(\phi) = \begin{cases} \frac{1}{1-\alpha} - \frac{\alpha(\overline{\phi}-c^H)+(\beta-\alpha)\Delta c}{(1-\alpha)(\phi-c^H)} & \textit{if } \phi \leq \hat{\phi} \\ \frac{1}{1-\alpha} - \frac{\alpha}{1-\alpha}\frac{\overline{\phi}-c^L}{\phi-c^L} & \textit{if } \phi > \hat{\phi} \end{cases}$$

Proof. Appendix, which also defines analytically the boundaries ϕ''_{in} and $\hat{\phi}$. ■

The following figure illustrates how the proposed equilibrium approaches simpler models as limiting cases. The left figure in 3.10 depicts the limit as $\alpha \to 0$ (adverse selection only). The lower-bound of the distributions approaches c^{LH}, while

Figure 3.10: Overdraft price dispersion - limits
Naiveté and adverse selection



(a) $\alpha \to 0$ (b) $\alpha \to \beta$

the upper bound of F^H_{in} approaches c^L. At the same time, the distributions on the upper range (for $\phi \in (c^L, \overline{\phi})$ converge to a mass-point on c^L for both insider and outsider. This is exactly the distribution we derived in Section 3.

The right figure illustrates that as $\alpha \to \beta$, the upper part of the distributions 'disappear' and F^L_{in} becomes degenerate. The lower part of the distributions then will be similar to Section 2, where $\beta = 0$ and $\alpha > 0$, which is clearly a specific case of $\beta < \alpha$.

3.5.2 Equilibrium characterization

Switching probabilities and expected equilibrium overdraft fees are characterized [20] in Appendix. Here in the main text we only focus on bank's profits from their role as insider and outsider, as this is sufficient to characterize the symmetric first-stage equilibrium what we introduce below.

Banks' second-period equilibrium profits can be decomposed into a sum of the profit from their role as outsider and insider bank. The insider profit is further decomposed into profits from high-types and from low-types. That is if Bank j have market share $\gamma^{(j)}$, its profit is

$$\pi^j = \gamma^{(j)}\pi^{L(j)}_{in} + \gamma^{(j)}\pi^{H(j)}_{in} + \gamma^{(-j)}\pi^j_{out} \tag{3.10}$$

Notice that the superscript of the market share for bank's outsider role is $(-j)$, as it obtains the other bank's customers as outsider. The components can be obtained

[20]We provide for this section only a partial characterization, as in some cases the analytical solutions are difficult.

directly from the respective indifference conditions. From Appendix, we have for the case of $\alpha \leq \beta$

$$\begin{aligned}
\pi_{in}^{L,j} &= \alpha^{(j)}\beta^{(j)}\left(\overline{\phi} - c^L\right) \\
\pi_{in}^{H,j} &= (1-\beta^{(j)})\left(\beta^{(j)}\left(c^L - c^H\right) + \alpha^{(j)}\left(\overline{\phi} - c^L\right)\right) \\
\pi_{out}^{j} &= \alpha^{(-j)}\left(1-\alpha^{(-j)}\right)\left(\overline{\phi} - c^L\right)
\end{aligned}$$

Whenever $\alpha_j = \alpha_{-j} = \alpha$ and $\beta^{(j)} = \beta^{(-j)}$ for an arbitrary γ_j (and noting that $\forall j\ \gamma_j = 1-\gamma_{-j}$), this simplifies to

$$\pi = \gamma(1-\beta)\beta\Delta c - \alpha^2(1-\gamma)(\overline{\phi} - c^L) + \alpha(\overline{\phi} - c^L)$$

With symmetric $\gamma = \frac{1}{2}$ and α we obtain

$$\pi_{\alpha\leq\beta}^{(2)} = \frac{1}{2}\left(\alpha(2-\alpha)(\overline{\phi} - c^L) + \beta(1-\beta)\Delta c\right) \tag{3.11}$$

which in the special case of $\alpha = 0$ equals the parallel expression in Section 3.

For the case of $\alpha > \beta$, the appropriate profits are:

$$\begin{aligned}
\pi_{in}^{L,j} &= \alpha^{(j)}\beta^{(j)}\left(\overline{\phi} - c^L\right) \\
\pi_{in}^{H,j} &= \alpha^{(j)}(1-\beta^{(j)})\left(\overline{\phi} - c^H\right) \\
\pi_{out}^{j} &= (1-\alpha_{-j})\left(\phi'_{in} - c^{LH}\right) \\
&= (1-\alpha_{-j})\left(\alpha_{-j}\overline{\phi} + (1-\alpha_{-j})c^H - c^{LH}\right)
\end{aligned}$$

Whenever $\alpha_j = \alpha_{-j} = \alpha$ for an arbibrary γ_j (and noting that $\forall j\ \gamma_j = 1-\gamma_{-j}$), this simplifies to

$$\pi = \alpha\gamma_j\left(\overline{\phi} - c^H - \beta\Delta c\right) + (1-\alpha)(1-\gamma_j)\left(\alpha(\overline{\phi} - c^H) - \beta\Delta c\right)$$

With symmetric $\gamma = \frac{1}{2}$ and α we obtain

$$\pi_{\beta<\alpha}^{(2)} = \frac{1}{2}\left(\alpha(2-\alpha)\left(\overline{\phi} - c^H\right) - \beta\Delta c\right) \tag{3.12}$$

which in case of $\beta = 0$ equals the parallel expression in Section 2.

3.5.3 First-stage equilibrium

Naive customers' decision is simply the threshold value calculated from first-period prices:

$$\hat{\gamma}^N = \frac{1}{2} + \frac{\Delta p}{2\tau}$$

Sophisticated customers first correctly predict the equilibrium value of *their* threshold strategy, that's $\hat{\gamma}^S$. In turn, this will pin down the parameters required to calculate equilibrium overdraft fees, that is (potentially), α_j and β_j, and γ_j. This allows to calculate customers' expected payment conditional on that everyone follows the equilibrium strategy. Every *candidate* $\hat{\gamma}^S$ generates some expected payment — equilibrium is found where the sophisticated customer located at $\hat{\gamma}^S$ is indeed indifferent between choosing Bank 1 and Bank 2. This raises the possibility of multiple equilibria as well. In the rest of the chapter we focus on a relatively simple symmetric equilibrium, which circumvents the need for calculating overdraft fees. More detailed analysis of the first-stage, including the case of multiple equilibria, is left for future work.

3.5.3.1 Symmetric equilibrium

Establishing the equilibrium consists of two steps. First, we will show that for any, not necessarily equal[21] values of p_1 and p_2, following the exact same behaviour by sophisticated and naive customers in the first stage is still an equilibrium. The argument is similar to the one in the previous section with naiveté. By assumption, low-types and high-types make the same first-period choice[22], therefore, for any strategy of sophisticated customers, the fraction of low and high types will be the same within the two banks. In addition, if sophisticated customers follow the same strategy as naive customers, the mass of naive customers will also be the same. As the two bank has the same 'structural' parameters, and (overdraft-) pricing is scale-free in the sense that it is independent of the market share, given this predicted equilibrium, sophisticated customers predict the same equilibrium expected overdraft fee payment. Therefore, they will base their decision solely on the first-period payout, and would follow the same strategy as naive customers, justifying this as an equilibrium action.

Next, taking this behaviour as given, we look at the banks' *total profit* function, and calculate the first-order condition for optimum PCA-price. The first-order

[21] such that $\hat{\gamma}^N \in (0, 1)$ is maintained

[22] customers do not know ex-ante whether they will be perceived low or high-types by the bank, that would require them to know the bank's behavioural scoring system, which is unlikely

condition captures an intuitive trade-off: as first-period prices decrease, a bank can increase its market share, but obtains lower profit from each individual customers. This leads to a first-period price which is unique in this symmetric equilibrium.

We restrict attention to symmetric equilibrium with $p_1 = p_2$. The Bank's profit is given by Equation (3.10), and below we derive the fixed point of the best-response mapping to the two cases, $\alpha \leq \beta$ and $\alpha > \beta$. The overall profit-functions for both period - as a sum of first-period profit from PCA-deals and second period profit on overdrafts - in the case of a symmetric equilibrium can be constructed using (3.11) and (3.12) for example from Bank 1's point of view as follows:

$$\pi_{\alpha\leq\beta} = \gamma_1 (p_1, p_2)\, p_1 + \pi^{(2)}_{\alpha\leq\beta}$$
$$\pi_{\alpha>\beta} = \gamma_1 (p_1, p_2)\, p_1 + \pi^{(2)}_{\alpha>\beta}$$

We substitute $\gamma_1 = \gamma^N$ and calculate the best response by Bank 1 to any p_2 as a solution of the following first-order conditions:

$$\frac{\partial \pi_{\alpha\leq\beta}}{\partial p_1} = \frac{1}{2\tau}\left(\beta^2\Delta c - \beta\Delta c - 2p_1 + p_2 - \alpha^2(\overline{\phi} - c^L) + \tau\right) = 0$$
$$\frac{\partial \pi_{\alpha>\beta}}{\partial p_1} = \frac{1}{2\tau}\left(2\alpha\beta\Delta c - \beta\Delta c - 2p_1 + p_2 - \alpha^2(\overline{\phi} - c^H) + \tau\right) = 0$$

this gives us the best-response function:

$$\tilde{p}_{1,\alpha\leq\beta}(p_2, \cdot) = \frac{1}{2}\left(-\beta\Delta c + \beta^2\Delta c + p_2 - \alpha^2\left(\overline{\phi} - c^L\right) + \tau\right)$$
$$\tilde{p}_{1,\alpha>\beta}(p_2, \cdot) = \frac{1}{2}\left(-\beta\Delta c + 2\alpha\beta\Delta c + p_2 - \alpha^2\left(\overline{\phi} - c^H\right) + \tau\right)$$

The symmetric equilibrium we are after here is given by the fixed-point equations:

$$p = \frac{1}{2}\left(-\beta\Delta c + 2\alpha\beta\Delta c + p - \alpha^2\left(\overline{\phi} - c^L\right) + \tau\right)$$
$$p = \frac{1}{2}\left(-\beta\Delta c + \beta^2\Delta c + p - \alpha^2\left(\overline{\phi} - c^H\right) + \tau\right)$$

which gives the symmetric equilibrium solution:

$$p^\star_{\alpha\leq\beta} = \beta^2\Delta c - \beta\Delta c + \tau - \alpha^2\left(\overline{\phi} - c^L\right) \tag{3.13}$$
$$p^\star_{\alpha>\beta} = 2\alpha\beta\Delta c - \beta\Delta c + \tau - \alpha^2\left(\overline{\phi} - c^H\right) \tag{3.14}$$

Whenever $\alpha = \beta$ the two predicted PCA prices coincide, so $p^\star$ is continuous at this point of the parameter space.

Theorem 5 *There exist a symmetric equilibrium in the first stage, characterized as follows:*

- *Whenever* $\alpha \leq \beta$*, the symmetric equilibrium is 'FIIC' with* $p^\star = 0$ *whenever* $\tau \leq \alpha^2(\overline{\phi} - c^L) - \beta(1-\beta)\Delta c$*. Otherwise there exist a positive price given by Equation* (3.13)*.*
- *Whenever* $\alpha > \beta$*, the symmetric equilibrium is 'FIIC' with* $p^\star = 0$ *whenever* $\tau \leq \alpha^2(\overline{\phi} - c^H) - \beta(1-2\alpha)\Delta c$*. Otherwise there exist a positive price given by Equation* (3.14)*.*

The difference between the equilibrium fee in (3.14) and the case of customer naiveté only is exactly $\beta\Delta c(2\alpha - 1)$, which is zero if $\beta = 0$ or if $\Delta c = 0$ — that is, if there is no adverse selection. Furthermore, the difference is negative whenever $\alpha < \frac{1}{2}$. On the other hand, the difference between 'naiveté only' PCA-price and Equation (3.13) is always negative. Recall that relatively low α values are typically the parameter regions where PCA prices may turn positive as a result of first-stage market power (relatively high value of τ), as the potential to exploit naive customers is extremely limited. This is inherent in all Gabaix-Laibson-style models, but also verified by our approach. The comparison above implies the following corollary:

Corollary 3 *The presence of adverse selection typically decreases first-period prices and makes FIIC-pricing more likely. Specifically, whenever* $\tau > 0$*, any adverse selection (which requires* $\beta > 0$ *and* $\Delta c > 0$*) decreases any positive PCA prices whenever* $\alpha < \beta$*, and also if* $\beta < \alpha < \frac{1}{2}$*.*

Figure 3.11 provides an illustration of how the presence of information asymmetry and adverse selection affects equilibrium outcome. We choose the parameters so that the *expected (net) cost* of the overdraft business is exactly equal (that is, $\beta c^L + (1-\beta)c^H$ is constant). The left-side figure depicts an already known situation with customer naiveté only ($\beta = 0$). The presence of naiveté raises profits over the whole interval $\alpha \in (0, 1)$, imposing FIIC for a significant part of the parameter space, except for lower values of α. In contrast, the right-hand figure depicts a situation where $\beta = 0.5$. Despite the expected cost of overdraft business being equal, information asymmetry tilts the sources of profits towards insider role, leading to further increase of PCA prices even for low values of α.

Figure 3.11: Bank profit and PCA price
Without / with adverse selection

Positive PCA prices without information asymmetry (left-hand figure) are competed away as adverse selection makes insider role of a bank more prominent compared to its outsider role.

3.6 Discussion

In this section we revisit the most important results of the preceding analysis and emphasize the contribution from economic theory point of view, as well as our message to banking research in the context of overdraft markets.

We started the research by changing an important component of the economic literature on aftermarket-pricing, namely that the seller of the primary good has monopoly power on the aftermarket. Instead, we take an almost opposite view by modelling a homogeneous add-on good with Bertrand-competition on the aftermarket. We argued that this is a more accurate description of retail banking, as there are no conventional technological constraints linking the aftermarket-good to the base-market good, like for example in the case of a printer and its cartridges. In the meantime, we certainly acknowledge that there is market power in retail banking: however, its source is rather attributable to behavioural biases (customer naiveté), or informational frictions (adverse selection). We therefore constructed a stylized model which is able to capture the interaction of these two effects in an otherwise competitive aftermarket environment.

We believe that this approach is more suitable to address the potential impact of recent policy developments in the Banking sector. Due to an increasing pressure from customer protection groups, policy makers recently started to look more seriously at some prevalent retail banking practices, such as FIIC-pricing. As a result of this scrutiny, new measures - such as the 2nd Payment Services Directive in the EU, or the Open Banking Programme in the UK - have been implemented to address the potentially welfare-reducing consequences. The most important elements of these changes seem to target customers' behavioural biases, like naiveté — possibly exacerbated by exploitative practices such as shrouding —, and banks' information monopoly over customers, which gives rise to adverse selection and prevents potential competitors to enter the market and improve efficiency. Professor Alasdair Smith from the UK's Competition and Market Authority summarizes this point neatly as[23]

> *The Open Banking programme (...) tackle the central problem in retail banking competition. That central problem is not "free banking"; it's the fact that bank customers don't have the information or tools they need to get the best deal from their banks.*

The implemented measures among others make it easier for customers to shop around using price comparison websites, switch to a competitor bank using switching services, and force banks to share customer account information on request to alleviate the potential adverse selection problem. As a result, banks' seem to have started to advertise heavily their overdraft conditions (instead of shrouding this information in advance), and introduced temporarily discounted overdraft fees as a selling-point.[24]

We started the analysis with a simple duopoly model where in the second stage naive customers can be exploited, while there is Bertrand-competition for sophisticated customers. We showed that there exist no pure-strategy Nash-equilibrium of the overdraft-pricing stage, and in the mixed-strategy equilibrium both insider and outsider bank randomizes its overdraft fees. We have demonstrated that the mixed-strategy equilibrium has intuitive limits: as $\alpha \to 0$ the equilibrium converges to a competitive (zero-profit) equilibrium, while with $\alpha \to 1$ it converges to a maximum-exploitation monopoly pricing. An interesting property of the mixed-strategy equilibrium is that insider's price dispersion allows the outsider bank as well to increase their prices, and obtain positive profit. When a bank decides whether to decrease prices it faces the following trade-off: it increases its market-share, therefore it obtains more profit in the first-period and from its role as insider, but it will obtain less profit from its outsider role. The equilibrium PCA

[23] Financial Times, FEBRUARY 21, 2017, `https://www.ft.com/content/72de5dc0-f79c-11e6-bd4e-68d53499ed71`

[24] Recall that the first, 'authentic' interpretation of price dispersion of homogeneous goods in economics was 'the model of sales', by Varian (1980).

price trades off these two forces. As the insider role is profitable, the equilibrium price will always be below the price which would arise on the market for PCA without the aftermarket (the latter is positive as banks have some market power derived from product differentiation in the first-stage), but it is limited by the presence of outsider profits, and therefore positive second-period profits are not competed away in the first-stage, even with positive prices. Therefore, in contrast to the literature, in this model the source of the positivity of bank's profits is not the binding lower boundary constraint, but a direct consequence of endogenous price determination with customer naiveté.

The presence of adverse selection changes this picture in an important way. As our analysis with adverse selection only (Section 4) points out, the information monopoly gives rise to an equilibrium with similar price dispersion, but this price dispersion only lets the insider to profit from its information advantage - outsider would get zero profit in equilibrium. This is a restatement of known results from the domain of relationship-banking literature (see for example, von Thadden (2004)). [25]

This intuition carries through to the full model where we combine customer naiveté and adverse selection. The appearance of adverse selection prevents the outsider to enjoy the extra profits from naiveté. As a result, the source of profit will be tilted towards insider customers, so the bank would compete more strongly for them in the first stage. This effect pushes the PCA prices downward, making FIIC-equilibrium even more likely. In particular, with the same mass of naive customers, low adverse-selection might lead to positive prices, while high adverse selection to zero-bound pricing and FIIC equilibrium. Adverse selection influences exploitation-ability in a nontrivial way.

Another aspect of this interaction links to the question of why overdraft markets develop in the first place.[26] The following high-level argument is motivated by Heidhues et al. [2016a], who study the incentives to 'innovate' pricing practices designed to exploit naive customers on the market — the overdraft segment itself being a prime example. Without much formality, and admitting that this argument may be subject to criticism, we say that there is 'incentive to innovate' and enter into a new market, if the new equilibrium (after sufficient convergence including follow-up entries) promises a higher equilibrium payoff, than the current status quo. Let this status quo be a banking market where overdrafts do not exist. Does

[25] Our version of the relationship-banking model presented here is essentially an extremely parsimonious version of the classic relationship-banking problem with adverse selection.

[26] This is not addressed formally in the model, but the discussion here can be a basis of a potentially important extension.

Figure 3.12: Bank profit and PCA price
Without / with adverse selection (2)

Market power (larger τ) in the first period makes overdrafts less profitable compared to an equilibrium where only PCA's are offered, without overdrafts.

the banking sector has incentives to make an investment to introduce this new product and move to a new equilibrium?[27] Figure 3.12 illustrates the effect of first-period PCA-price competition. Without adverse selection (left figure) — for an *arbitrary* level of α — overdrafts generate additional profits, therefore the answer is positive: there is always 'incentive to innovate' and introduce the overdraft business segment into the market. The presence of adverse selection (right figure) makes profits much less responsive to the presence of naive customers, especially for low α values. Although the profits obtained in the new equilibrium are still positive, the extra profit compared to the case where banks *only* sell PCA's and exploit product differentiation shrinks, and for sufficiently high τ, it even converges to zero. This means there is less incentives to innovate. Whenever the banks have large first-period market power (a more concentrated, oligopolistic retail banking market), equilibrium PCA prices tend to be larger, so it is less likely that the equilibrium is FIIC. The presence of adverse selection can bring this to the extreme, when banks have no incentives to innovate and introduce overdrafts at the first place.

To conclude, more concentrated, less competitive markets with higher product differentiation in Europe compared to the large and competitive banking sector in the US and the UK might be the the reason why FIIC is much more prevalent in the latter countries. In addition, higher adverse selection (more uncertainty re-

[27]Even if there is a generic demand for a certain type of product - like an overdraft-style liquidity facility -, it is not immediate, and not costless to introduce a new banking product unknown to customers, and start fulfilling the potential demand.

garding the borrowers' creditworthiness) makes FIIC more-likely, which suggest that FIIC might be more prevalent in countries with relatively higher income inequality and lower social safety net. These potential explanations which are suggested by our model are invariant to the level of customer naiveté, so FIIC can prevail or not even with the exact same number of naive customers - a puzzle which motivated our research at the beginning.

3.7 Conclusions

In this paper we proposed a novel way of thinking about bank's overdraft prices and aftermarket pricing in general. Even with competitive aftermarkets, behavioural frictions such as customer naiveté - the presence of customers who do not switch, despite a cheaper product is available on the market, and there are no other barriers to switch -, or informational frictions such as adverse selection can generate an equilibrium which is reminiscent to pricing schemes observed in retail banking. Specifically, the model predicts free-if-in-credit pricing ('free banking') for a large subset of parameters. In contrast to alternative models, FIIC arises with moderate number of naive customers, and as long as the market is subject to more severe adverse selection, even with small number of naive customers. Higher competition on the primary market also raises the possibility of the emergence of FIIC-pricing. These are novel findings in the theoretical literature, and help to explain better the observed differences in retail banking markets around the world. Our modelling framework is also better suited to analyse the potential impact of certain regulatory interactions, such as price-comparison websites, 'open banking', or switching services. Most of this analysis, and further comparative exercise on cross-country differences is left for future research work.

3.A Appendix A - Proofs, Customer myopia

3.A.1 Proof of Lemma 2

The Lemma is established through a series of claims.

Claim 1 *Insider will never bid any* $\phi_{in} < \phi'_{in}$*, so* $\underline{F}_{in} \geq \phi'_{in}$*. Furthermore,* $Pr[\phi_{in} = \phi'_{in}] = 0$.

Proof. (i) Charging $\overline{\phi}$ is always a feasible action for insider, and even if she wins at some $\phi_{in} < \phi'_{in}$ with probability 1 and loses at $\overline{\phi}$ with probability 1, the latter still gives higher profit by the definition of ϕ'_{in}. (ii) Suppose there is a mass point at ϕ'_{in} by the insider, that is, $Pr[\phi_{in} = \phi'_{in}] > 0$. This can only be equilibrium if $\underline{F}_{out} \leq \phi'_{in}$, otherwise insider would have incentives to increase the price. In addition, insider must win at ϕ'_{in} with certainty (otherwise would find it better to charge $\overline{\phi}$ by the definition of ϕ'_{in}). This implies outsider loses at $\phi_{out} = \phi'_{in}$ with a strictly positive probability. In that case, outsider is better off by charging $\phi'_{in} - \epsilon$ with probability 1, winning with certainty, and obtaining a profit of $\alpha\left(\phi'_{in} - c\right)\gamma - \epsilon$. Insider would lose at ϕ'_{in}. Contradiction to equilibrium. ■

Claim 2 *Outsider will never bid below* ϕ'_{in}*, so* $\underline{F}_{out} \geq \phi'_{in}$

Proof. Suppose the bid is $\phi < \phi'_{in}$. Because of claim 1, he wins with certainty, but than he would be better off by bidding $\frac{\phi'_{in}+\phi}{2}$. Contradiction. ■

Claim 3 *Whenever* $\phi'_{in} > c$ *outsider makes positive profit in equilibrium.*

Proof. For any bid $\phi_{out} \in (c, \phi'_{in})$ outsider would win with certainty and make positive profit. As this is a feasible deviation, there must be positive profit in equilibrium. ■

Claim 4 *Outsider never places positive mass on any* $\phi \geq \overline{F}_{in}$*. In particular,* $Pr[\phi_{out} = \overline{F}_{in}] = 0$.

Proof. In this region he would lose with certainty, implying zero profit and contradicting Claim 3. ■

Claim 5 $\overline{F}_{in} = \overline{\phi}$ *and insider's profit is* $\alpha\left(\overline{\phi} - c\right)\gamma$

Proof. Claim 4 implies that insider loses the bid with probability 1 at $\overline{F}_{in}$. Therefore, her profit when playing $\overline{F}_{in}$ can be at most $\alpha\left(\overline{F}_{in} - c\right)\gamma$. Because the minimax payoff is $\alpha\left(\overline{\phi} - c\right)\gamma$, and the profit is increasing in ϕ, it follows immediately that $\overline{F}_{in} = \overline{\phi}$ and the profit throughout the mixture is $\alpha\left(\overline{\phi} - c\right)\gamma$. ■

Claim 6 *Outsider's lower boundary must be exactly* $F_{out} = \phi'_{in}$

Proof. Suppose $\underline{F}_{out} > \phi'_{in}$. Then there exist a strategy for insider to bid $\underline{F}_{out}$, win the competition with probability 1, and make profit $(\underline{F}_{out} - c)\,\gamma > \pi^0_{in}$. This contradicts to Claim 5. ■

Claim 7 *Insider's lower boundary must be exactly* $\underline{F}_{in} = \phi'_{in}$

Proof. Suppose $\underline{F}_{in} > \phi'_{in}$. Then outsider could win all customers at $\phi_{out} = \underline{F}_{in} - \epsilon$, and would never bid anything below. Insider would then find it profitable to undercut this by bidding $\underline{F}_{in} - \epsilon$. ■

Claim 8 *The constant profit to outsider over the mixture is* $\pi_{out} = (\phi'_{in} - c)\,\gamma$.

Proof. At ϕ'_{in} outsider wins with probability 1, because there is no mass by insider. ■

3.A.2 Proof of Lemma 5

We want to calculate $Prob(\phi_{out} < \phi_{in})$. Start with the continuous part of the distributions where both players mix, ignoring the mass points. With the joint CDF f_{io} it is possible to write formally

$$Pr[\phi_{out} < \phi_{in}] = \int_{\phi'_{in}}^{\overline{\phi}} \int_{\phi'_{in}}^{\phi_{in}} f_{io} d\phi_{out} d\phi_{in}$$

From earlier lemmas we can calculate the two PDF's:

$$f_{in}(\phi) = F'_{in}(\phi) = \alpha \cdot \frac{\overline{\phi} - c}{(\phi - c)^2}$$
$$f_{out}(\phi) = F'_{out}(\phi) = \frac{\alpha}{1-\alpha} \cdot \frac{\overline{\phi} - c}{(\phi - c)^2}$$

The joint distribution because of the independence assumption is:

$$f_{io}(\phi_{in}, \phi_{out}) = f_{in}(\phi_{in}) \cdot f_{out}(\phi_{out}) = \frac{a^2}{1-a} \cdot \frac{\left(\overline{\phi} - c\right)^2}{(\phi_{out} - c)^2(\phi_{in} - c)^2}$$

The internal integral, with respect to ϕ_{out} is

$$\int f_{io} d\phi_{out} = -\frac{a^2}{1-a} \cdot \frac{(\overline{\phi} - c)^2}{(\phi_{out} - c)(\phi_{in} - c)^2}$$

$$\int_{\phi'_{in}}^{\phi_{in}} f_{oi} d\phi_{out} = \frac{a^2}{1-a} \cdot \frac{(\overline{\phi} - c)^2}{(\phi_{in} - c)^2} \cdot \left(\frac{1}{(\phi'_{in} - c)} - \frac{1}{(\phi_{in} - c)} \right)$$

$$= \frac{\alpha}{1-\alpha} \frac{\overline{\phi} - c}{(\phi_{in} - c)^2} - \frac{\alpha^2}{1-\alpha} \frac{(\overline{\phi} - c)^2}{(\phi_{in} - c)^3}$$

After integrating each components we get:

$$P\tilde{r}ob[\phi_{out} < \phi_{in}] = \frac{\alpha}{1-\alpha}(\overline{\phi} - c) \left[\frac{-1}{\phi_{in} - c} \right]_{\phi'_{in}}^{\overline{\phi}} - \frac{\alpha^2}{1-\alpha}(\overline{\phi} - c)^2 \left[\frac{-1}{2(\phi_{in} - c)^2} \right]_{\phi'_{in}}^{\overline{\phi}}$$

After substitutions of the integral boundaries, we get the formula for the probability:

$$P\tilde{r}ob[\phi_{out} < \phi_{in}] = \frac{1-a}{2}$$

This probability only considers the mass over the continuous-part of the two distributions, so it gives the probability mass of winning when insider plays mixture. In addition, outsider wins with certainty whenever insider plays $\overline{\phi}$. Together with the mass-point we obtain

$$Prob[\phi_{out} < \phi_{in}] = \frac{1-\alpha}{2} + \alpha = \frac{1+\alpha}{2}$$

$$Prob[\phi_{in} < \phi_{out}] = 1 - \frac{1+\alpha}{2} = \frac{1-\alpha}{2}$$

3.A.3 Proof of Lemma 6

We calculate the expected fee offered by insider resp. outsider *conditional on winning the price competition*, that is, $\mathbb{E}[\phi_{out}|\phi_{out} < \phi_{in}]$ and $\mathbb{E}[\phi_{in}|\phi_{in} \leq \phi_{out}]$. Because there are mass-points by the insider placed on $\overline{\phi}$, we have

$$\mathbb{E}[\phi_{out}|\phi_{out} < \phi_{in}] = \frac{1}{\frac{1+\alpha}{2}} \left(\int_{\phi'_{in}}^{\overline{\phi}} \int_{\phi'_{in}}^{\phi_{in}} \phi_{out} f_{io} d\phi_{out} d\phi_{in} + \alpha \int_{\phi'_{in}}^{\overline{\phi}} \phi_{out} f_{out} d\phi_{out} \right) \tag{3.15}$$

$$\mathbb{E}[\phi_{in}|\phi_{in} \leq \phi_{out}] = \frac{1}{\frac{1-\alpha}{2}} \left(\int_{\phi'_{in}}^{\overline{\phi}} \int_{\phi'_{in}}^{\phi_{out}} \phi_{in} f_{io} d\phi_{in} d\phi_{out} \right) \tag{3.16}$$

Because the joint distribution is symmetric in the two variables, the double-integral components are symmetric and equal. For example,

$$\int_{\phi'_{in}}^{\overline{\phi}} \int_{\phi'_{in}}^{\phi_{in}} \phi_{out} f_{io} d\phi_{out} d\phi_{in} = \alpha\overline{\phi} + \frac{1}{2}\left(\frac{2\alpha^2 \ln[\alpha](\overline{\phi} - c)}{1-\alpha} - (3\alpha - 1)(c)\right)$$

The details of this calculation are included below:

1. The internal integral

$$\begin{aligned}
\int_{\phi'_{in}}^{\phi_{in}} \phi_{out} f_{\phi_{out},\phi_{in}} d\phi_{out} &= \frac{\alpha^2}{1-\alpha}\frac{\left(\overline{\phi} - c\right)^2}{(\phi_{in} - c)^2}\int_{\phi'_{in}}^{\phi_{in}} \frac{\phi_{out}}{(\phi_{out} - c)^2} d\phi_{out} \\
&= \frac{\alpha^2}{1-\alpha}\frac{\left(\overline{\phi} - c\right)^2}{(\phi_{in} - c)^2}\left[\frac{-c}{\phi_{out} - c} + \ln(\phi_{out} - c)\right]_{\phi'_{in}}^{\phi_{in}} \\
&= \frac{\alpha^2}{1-\alpha}\frac{\left(\overline{\phi} - c\right)^2}{(\phi_{in} - c)^2}\left(\frac{-c}{\phi_{in} - c} + \ln(\phi_{in} - c) + \frac{c}{\alpha(\overline{\phi} - c)} - \ln(\alpha(\overline{\phi} - c))\right)
\end{aligned}$$

2. The full integral: required "ingredients":

$$\int \frac{1}{(x-c)^3} dx = -\frac{1}{2\,(x-c)^2} \Rightarrow \left[\frac{1}{2\left(\alpha(\overline{\phi} - c)\right)^2} - \frac{1}{2\left(\overline{\phi} - c\right)^2}\right] = \left[\frac{1-\alpha^2}{2\alpha^2}\frac{1}{\left(\overline{\phi} - c\right)^2}\right]$$

$$\begin{aligned}
\int \frac{\ln(x-c)}{(x-c)^2} dx = -\frac{\ln(x-c)+1}{x-c} &\Rightarrow \left[\frac{\ln(\alpha(\overline{\phi} - c)) + 1}{\alpha(\overline{\phi} - c)} - \frac{ln(\overline{\phi} - c) + 1}{\overline{\phi} - c}\right] \\
&= \frac{\ln(\overline{\phi} - c)}{\overline{\phi} - c}\left(\frac{1-\alpha}{\alpha}\right) + \frac{1}{\overline{\phi} - c}\left(\frac{1-\alpha}{\alpha}\right) + \frac{\ln\alpha}{\alpha(\overline{\phi} - c)} \\
&= \left[\frac{1-\alpha}{\alpha(\overline{\phi} - c)}\left(\ln(\overline{\phi} - c) + 1\right) + \frac{\ln\alpha}{\alpha(\overline{\phi} - c)}\right]
\end{aligned}$$

$$\int \frac{1}{(x-c)^2} dx = \frac{-1}{x-c} \Rightarrow \left[\frac{-1}{\overline{\phi} - c} + \frac{1}{\alpha(\overline{\phi} - c)}\right] = \left[\frac{1-\alpha}{\alpha(\overline{\phi} - c)}\right]$$

Therefore

$$\begin{aligned}
\int_{\phi'_{in}}^{\overline{\phi}} [...] d\phi_{in} = \frac{\alpha^2\left(\overline{\phi} - c\right)^2}{1-\alpha}&\left(-c\left[\frac{1-\alpha^2}{2\alpha^2}\frac{1}{\left(\overline{\phi} - c\right)^2}\right] + \left[\frac{1-\alpha}{\alpha(\overline{\phi} - c)}\left(\ln(\overline{\phi} - c) + 1\right) + \frac{\ln\alpha}{\alpha(\overline{\phi} - c)}\right]\right. \\
&\left. + \left(\left(\frac{c}{\alpha(\overline{\phi} - c)}\right) - \ln\alpha - \ln(\overline{\phi} - c)\right)\left[\frac{1-\alpha}{\alpha(\overline{\phi} - c)}\right]\right)
\end{aligned}$$

This simplifies significantly:

$$\begin{aligned}\int_{\phi'_{in}}^{\overline{\phi}} [...]d\phi_{in} &= \frac{-(1+\alpha)c}{2} + \alpha(\overline{\phi}-c)\big(\ln(\overline{\phi}-c)+1\big) + \frac{\alpha\ln\alpha(\overline{\phi}-c)}{1-\alpha} \\ &\quad + c - \alpha(\overline{\phi}-c)\big(\ln[\alpha]+\ln(\overline{\phi}-c)\big) \\ &= \alpha\overline{\phi} + \alpha(\overline{\phi}-c)\left(-\frac{\ln[\alpha]}{1-\alpha} - \ln[\alpha]\right) + c - \frac{(1+\alpha)c}{2} - \alpha c \\ &= \alpha\overline{\phi} + \alpha(\overline{\phi}-c)\left(\frac{\alpha\ln[\alpha]}{1-\alpha}\right) + c\left(1-\alpha-\frac{1+\alpha}{2}\right) \\ &= \alpha\overline{\phi} + \frac{\alpha^2\ln[\alpha]}{1-\alpha}(\overline{\phi}-c) + c\left(\frac{1-3\alpha}{2}\right)\end{aligned}$$

The calculation of the unconditional expected value is relatively straightforward, so we omit details:

$$\int_{\phi'_{in}}^{\overline{\phi}} \phi_{out} f_{out} d\phi_{out} = c - \frac{\alpha(\overline{\phi}-c)\log[a]}{1-a}$$

After combining the two expressions:

$$\mathbb{E}[\phi_{out}|\phi_{out} < \phi_{in}] = \frac{1}{1+\alpha}\big(2\overline{\phi} + (1-\alpha)(c)\big)$$

This can be rewritten as

$$\mathbb{E}[\phi_{out}|\phi_{out} < \phi_{in}] = c + \frac{2\alpha}{1+\alpha}\left(\overline{\phi}-c\right)$$

which is reminiscent to a markup-pricing formula. Similarly,

$$\begin{aligned}\mathbb{E}[\phi_{in}|\phi_{in} \le \phi_{out}] &= c + \frac{2\alpha}{1-\alpha}\left(\overline{\phi}-c\right)\left(\frac{\alpha\ln[\alpha]}{1-\alpha}+1\right) \\ &= c + \frac{2\alpha}{1-\alpha}\left(1+\frac{\alpha\ln[\alpha]}{1-\alpha}\right)\left(\overline{\phi}-c\right)\end{aligned}$$

Because $\alpha < 1$, the mark-up is always smaller for the insider.

3.A.4 Proof of Lemma 7

Bank's profit

The outsider profit can be rewritten for $j \in \{1,2\}$ as

$$\pi_{out} = \gamma_{-j}(1-\alpha_{-j})\left(\alpha_{-j}\overline{\phi} + (1-\alpha_{-j})c - c\right) = \gamma_{-j}\alpha_{-j}(1-\alpha_{-j})\left(\overline{\phi} - c\right)$$

The overall profit:

$$\pi^j = \pi^j_{in} + \pi^j_{out} = \left(\gamma_j\alpha_j + \gamma_{-j}\alpha_{-j}(1-\alpha_{-j})\right)\left(\overline{\phi} - c\right) \tag{3.17}$$

Notice that

$$\gamma_j\alpha_j + (1-\gamma_j)\alpha_{-j} = \alpha \quad \Rightarrow \quad \alpha_{-j} = \frac{\alpha - \gamma_j\alpha_j}{1-\gamma_j}$$

therefore

$$\begin{aligned}
\pi_j &= \left(\gamma_j\alpha_j + \gamma_{-j}\alpha_{-j}(1-\alpha_{-j})\right)\left(\overline{\phi} - c\right) = -\frac{\alpha^2 + \alpha_j^2\gamma_j^2 - \alpha(1-\gamma_j + 2\alpha_j\gamma_j)}{1-\gamma_j}\left(\overline{\phi} - c\right) \\
&= \left[\alpha - \frac{\alpha^2 + \alpha_j^2\gamma_j^2 - 2\alpha\alpha_j\gamma_j}{1-\gamma_j}\right]\left(\overline{\phi} - c\right) = \left[\alpha - \frac{(\alpha - \alpha_j\gamma_j)^2}{1-\gamma_j}\right]\left(\overline{\phi} - c\right)
\end{aligned}$$

whenever $\alpha_j = \alpha$:

$$\begin{aligned}
\left[\alpha - \frac{(\alpha - \alpha_j\gamma_j)^2}{1-\gamma_j}\right] &= -\frac{\alpha^2(1+\gamma_j) - 2\alpha^2\gamma_j}{1-\gamma_j} + \alpha = -\frac{\alpha^2\left(1+\gamma_j^2 - 2\gamma^j\right)}{1-\gamma_j} + \alpha \\
&= -\frac{\alpha^2\left(1-\gamma_j\right)^2}{1-\gamma_j} + \alpha = \alpha - \alpha^2(1-\gamma_j) = \alpha(1-\alpha(1-\gamma_j)) \\
&= \alpha(1-\alpha) + \gamma_j\alpha^2
\end{aligned}$$

as we have

$$\frac{\partial\pi}{\partial\gamma_j} = \alpha^2\left(\overline{\phi} - c\right) > 0$$

,the profit always increases in own market share.
Whenever $\gamma_1 = \gamma_2 = 1/2$,

$$\begin{aligned}
\pi_j &= \alpha(1 - \frac{1}{2}\alpha)\left(\overline{\phi} - c\right) \\
&= \frac{1}{2}\alpha(2-\alpha)\left(\overline{\phi} - c\right)
\end{aligned}$$

Decomposition of profits

We decompose profits in a symmetric equilibrium into three parts: (i) profit from myopic customers, who always stay with the Bank, irrespectively of prices; (ii) profit on sophisticated customers because they decide to stay in equilibrium (and pay ϕ_{in}), or (iii) because they switch from the other bank (and pay ϕ_{out}). That is, for any given pair of equilibrium overdraft fees $\{\phi_{in}, \phi_{out}\}$

$$\mathbb{E}\pi(\phi_{in}, \phi_{out}) = \underbrace{\alpha\pi(\phi_{in})}_{\text{Profit from myopes}} + \underbrace{p^{ow}(1-\alpha)\pi(\phi_{out})}_{\text{Profit switching sophisticates}} + \underbrace{p^{iw}(1-\alpha)\pi(\phi_{in})}_{\text{Profit from staying sophisticates}}$$

where $p^{out.w}$ and $p^{in.w}$ are the probabilities that outsider and insider wins the competition.

Using results from previous section, it is possible to rewrite expected profit from serving a unit myopic / sophisticated switcher / sophisticated remainer customer as

$$\begin{aligned}
\mathbb{E}\pi^{myop} &= \alpha\left(\mathbb{E}[\phi_{in}] - c\right) \\
\mathbb{E}\pi^{\text{soph,switch}} &= Pr[switch]\left(\mathbb{E}[\phi_{out}|switch] - c\right) \\
\mathbb{E}\pi^{\text{soph,stay}} &= Pr[stay]\left(\mathbb{E}[\phi_{in}|stay] - c\right)
\end{aligned}$$

where

$$\mathbb{E}[\phi_{in}] = \int_{\phi'_{in}}^{\overline{\phi}} \phi_{in} f_{\phi_{in}} d\phi_{in} + \alpha\overline{\phi}$$

and the other variables are defined in the previous Lemma. The total expected income *from sophisticated customers* is

$$\mathbb{E}\pi^{soph} = p + R + \mathbb{E}[\phi|switch] \cdot Pr[switch] + \mathbb{E}[\phi|stay] \cdot Pr[stay]$$

After substitution and algebraic simplifications:

$$\begin{aligned}
\mathbb{E}\pi^{myop} &= \alpha^2\left(1 - \ln[\alpha]\right)\left(\overline{\phi} - c\right) \\
\mathbb{E}\pi^{\text{soph,switch}} &= (1-\alpha)\,\alpha\left(\overline{\phi} - c\right) \\
\mathbb{E}\pi^{\text{soph,stay}} &= \alpha\left(1 - \alpha + \alpha\ln[\alpha]\right)\left(\overline{\phi} - c\right)
\end{aligned}$$

3.B Appendix B - Proofs, Adverse selection

3.B.1 Proof of Lemma 9

The Lemma is established through a series of claims.[28] Let $\underline{F}_\theta$ and $\overline{F}_\theta$ denote the infimum and the supremum of the support of the probability distributions F_θ for $\theta \in \{\text{“}in\text{”}, \text{“}out\text{”}\}$.

Claim 1 (Participation constraints) *(i) Insider would never offer any $\phi^H_{in} < c^H$ for the high types, therefore $\underline{F}^H_{in} \geq c^H$. Similarly, it would never offer any $\phi^L_{in} < c^L$, so $\phi^L_{in} \geq c^L$. (ii) Outsider would never offer any $\phi_{out} < c^{HL}$, so $\underline{F}_{out} \geq c^{LH}$.*

Proof. (i)The bank would get negative profit if it wins. (ii) By (i), $\phi^L_{in} \geq c^L$, and because $c^L > c^{LH}$, by bidding c^{LH} or below, the outsider gets all low-types. Therefore, even if it wins all high-types with probability 1, it would make a loss at any $\phi < c^{LH}$. ■

Claim 2 (Common lower boundaries) *(i) The lower boundaries of insider-H's and outsider's distributions must coincide: $\underline{F}_{out} = \underline{F}^H_{in} := \underline{F}$ (ii) There must be no probability mass by outsider at the lower boundary.*

Proof. (i) Suppose $\underline{F}^H_{in} < \underline{F}_{out}$. Then insider would win at $\underline{F}_{out}$ with probability 1, and make strictly larger profit, contradicting equilibrium. Similarly, there cannot be any $\underline{F}_{out} < \underline{F}_{in}$, since outsider could raise it to $\frac{\underline{F}_{out}+\underline{F}^H_{in}}{2}$, still win with probability 1, and obtain larger profit.

(ii) Suppose $\underline{F} > c^{LH}$ and outsider places a positive probability mass on $\underline{F}$. Then insider would lose at $\underline{F} + \epsilon$ with $\epsilon \to 0$ with a probability strictly larger than 0. Instead, she could improve her payoff by playing $\underline{F} - \epsilon$ and win with probability 1. ■

Claim 3 *Insider makes strictly positive profit.*

Proof. Because of Claim 2, by playing $\underline{F}$ the insider wins with probability 1. Because $\underline{F} \geq c^{HL}$ due to Claim 1, insider obtains a profit at least $c^{HL} - c^H > 0$ on each (high-type) customer. ■

Claim 4 *There is a common upper boundary of the continuous parts of the distributions: $\overline{F}_i = \overline{F}_o := \overline{F} = c^L$*

[28] This is an existence-proof, and does not address the uniqueness of equilibrium. In particular, we ‘guess’ some properties of the equilibrium — namely that the insider fee to low-types is a degenerate distribution at c^L — and then verify that this is indeed an equilibrium. Also, continuity of the distributions is imposed. We believe that this simplified proof carries the main intuition while avoids less interesting technical details.

Proof. That the upper boundaries are common is trivial. Suppose $\overline{F} > c^L$. Then outsider obtains positive profit. By playing $\overline{F}$ outsider loses with certainty, contradicting positive profit, therefore there must not be probability mass by the outsider on $\overline{F}$. This implies insider loses competition for high-types at $\overline{F}$ with probability 1, and makes zero profit on high-types. That cannot be part of an equilibrium, as there is a profitable deviation to charge $c^{HL} - \epsilon$ and obtain strictly positive profit. Contradiction to $\overline{F} > c^L$. Now suppose $\overline{F} < c^L$. Given that $\phi_{in}^L = c^L$, outsider would get all low-types. If there is no mass-point by insider on $\overline{F}$, outsider would also lose all high-types and make negative profit, contradicting equilibrium. But insider cannot have mass point at $\overline{F}$, as in that case outsider must have zero mass on $\overline{F}$, and insider loses with certainty, and obtains lower profit than at c^{HL}. ■

Claim 5 *Outsider is making zero profit.*

Proof. By charging c^L outsider loses with certainty and obtains zero profit. ■

3.B.2 Proof of Lemma 10

Switching probabilities and expected fees

From the CDF's in Theorem 3 we can calculate the PDF's and the joint PDF:

$$
\begin{aligned}
f_{in}(\phi_{in}^H) &= \frac{\beta}{1-\beta} \frac{c^L - c^H}{\left(\phi_{in}^H - c^H\right)^2} \\
f_{out}(\phi_{out}) &= \beta \frac{c^L - c^H}{(\phi_{out} - c^H)^2} \\
f_{io}(\phi_{in}^H, \phi_{out}) &= \frac{\beta^2}{1-\beta} \frac{\left(c^L - c^H\right)^2}{\left(\phi_{in}^H - c^H\right)^2 (\phi_{out} - c^H)^2}
\end{aligned}
$$

The following computations are similar to the case with customer naiveté. The *truncated* probability (over the mixture) that insider vs. outsider wins is:

$$
\begin{aligned}
\tilde{Pr}[\phi_{in}^H < \phi_{out}] &= \int_{c^{LH}}^{c^L} \int_{c^{LH}}^{\phi_{out}} f_{io} d\phi_{in}^H d\phi_{out} = \frac{1-b}{2} \\
\tilde{Pr}[\phi_{out} < \phi_{in}^H] &= \int_{c^{LH}}^{c^L} \int_{c^{LH}}^{\phi_{in}^H} f_{io} d\phi_{out} d\phi_{in}^H = \frac{1-b}{2}
\end{aligned}
$$

The (unconditional) probabilities are

$$Pr[\phi_{in}^{H} < \phi_{out}] = \frac{1-\beta}{2} + \beta = \frac{1+\beta}{2}$$
$$Pr[\phi_{out} < \phi_{in}^{H}] = \frac{1-\beta}{2}$$

Note that the *conditional* probability that insider/outsider wins *given* mixture is:

$$Pr[\phi_{in}^{H} < \phi_{out}|mix] = Pr[\phi_{out} < \phi_{in}^{H}|mix] = \frac{1-\beta}{2}/(1-\beta) = \frac{1}{2}$$

The expected fee of insider conditional on winning is

$$\begin{aligned}
\mathbb{E}[\phi_{in}^{H}|\phi_{in}^{H} < \phi_{out}] &= \frac{\int_{c^{LH}}^{c^{L}} \int_{c^{LH}}^{\phi_{out}} \phi_{in}^{H} f_{io} d\phi_{in}^{H} d\phi_{out} + \beta \int_{c^{LH}}^{c^{L}} \phi_{in}^{H} f_{in} d\phi_{in}^{H}}{\tilde{Pr}[\phi_{in}^{H} < \phi_{out}]} \\
&= \left[\frac{1}{2}\big((1-\beta)(c^{H}) + 2\beta\Delta c\big) + \frac{\beta^{2}\ln[\beta]\Delta c}{1-\beta} + \beta\left(c - \frac{\beta\ln[\beta]}{1-\beta}\Delta c\right)\right] * \frac{2}{1+\beta} \\
&= c^{H} + \frac{2\beta}{1+\beta}\Delta c
\end{aligned}$$

The expected fee of outsider conditional on winning:

$$\begin{aligned}
\mathbb{E}[\phi_{out}|\phi_{out} < \phi_{in}^{H}] &= \frac{\int_{lb}^{ub} \int_{lb}^{\phi_{out}} \phi_{out} f_{io} d\phi_{out} d\phi_{in}^{H}}{\tilde{Pr}[\phi_{out} < \phi_{in}^{H}]} \\
&= \left[\frac{1}{2}\big((1-\beta)c^{H} + 2\beta\Delta c\big) + \frac{\beta^{2}\ln[\beta]\Delta c}{1-\beta}\right] * \frac{2}{1-\beta} \\
&= c^{H} + \frac{2\beta}{1-\beta}\Delta c + \frac{2\beta^{2}}{(1-\beta)^{2}}\ln[\beta]\Delta c
\end{aligned}$$

Finally, we need to know the (unconditional) expected fees:

$$\mathbb{E}[\phi_{out}] = \tilde{\mathbb{E}}[\phi_{out}] + \beta(c^{L}) = (1-\beta)c^{H} + \beta c^{L} - \beta\ln[\beta]\Delta c$$
$$\mathbb{E}[\phi_{in}^{H}] = c^{H} - \frac{\beta\ln[\beta]\Delta c}{1-\beta}$$

3.C Appendix C - Proofs, Full model

3.C.1 No pure-strategy for L-types

For didactic reasons we give a proof sketch for the claim that for $\alpha < \beta$ there is no equilibrium in which insider follows a pure strategy (degenerate distribution) for low-types. For $\alpha > \beta$, there is an equilibrium with $\phi_{in}^L = \overline{\phi}$. Although this also follows from the main theorem below, this short proof provides important insight.

Proof. *(Sketch)* Suppose that insider's strategy to bid for low-types is a degenerate distribution, and it places probability mass 1 on some value ϕ_{in}^L.

Suppose first that $\alpha < \beta$, and $\phi_{in}^L > \overline{F_{in}^H}$. There will be no probability mass by insider on $\overline{F_{in}^H}$, but to guarantee at least the minimax payoff to insider, outsider must have probability mass at $\overline{F_{in}^H}$, loses with certainty, and makes zero profit on high-types. Given that pricing, outsider would find it optimal to move all mass instead to $\phi_{in}^L - \epsilon$, and obtain larger profit. (At $\overline{F_{in}^H}$ and $\phi_{in}^L - \epsilon$ outsider would serve the same set of customers, but at a higher price). This is a possible deviation whenever $\phi_{in}^L > \overline{F_{in}^H}$. Now suppose (with slightly imprecise but intuitive notation) that $F_{in}^L = \overline{F_{in}^H}+$, so that outsider cannot profitably raise its prices as previously described. This however cannot be equilibrium either: in this case insider would have a profitable deviation, as by raising to $F_{in}^L = \overline{\phi}$ it would obtain higher profit.

Finally, suppose $\alpha < \beta$ and $\phi_{in}^L < \overline{F_{in}^H}$. Outsider cannot be indifferent between $\phi_{in}^L - \epsilon$ and ϕ_{in}^L, as in the first case it would get all, in the second case none of the low-type customers. It can be indifferent only if the profit from low-types is zero, that is, if $F_{in}^L = c^L$. However, in this case insider would have a profitable deviation to raise $\overline{\phi}$ and serve only naive customers. Contradiction. ■

Notice that if insider charges $\phi_{in}^L = \overline{\phi}$ and there is no mass by the outsider on $\overline{\phi}$. Then low-types always switch. For any $\alpha > \beta$ we can show that under those conditions there exist an equilibrium with some distributions F_{in}^H and F_{out}. These distributions satisfy that $\sup\{\phi | F_{in}^H(\phi) < 1\} = \overline{\phi}$ and $Pr[\phi_o = \overline{\phi}] = 0$ (that is, $\overline{\phi}$ is the upper boundary of insider-HIGH, but not played by outsider with positive probability). Under those conditions insider cannot decrease prices for low-types because given outsider's distribution, the profit would be lower. Therefore, outsider indeed obtains all low-type customers, and the proposed pricing is equilibrium.

3.C.2 Proof of Theorem 4

The Theorem is proven in 4 parts.First, we characterize the structure of equilibrium. Then, we show that only one 'minimax payoff' can be binding. Finally, we derive the distributions for $\alpha \leq \beta$ and for $\alpha > \beta$ separately.

3.C.2.1 PART A — Equilibrium structure

Recall that we solve for three generic, *independent* distributions: F_{out}, F_{in}^H and F_{in}^L. We start by establishing some claims regarding the equilibrium:

Lemma 11 *The three distributions, F_{out}, F_{in}^H and F_{in}^L must satisfy that (1) the supports of insider's distributions for high and low types do not overlap; (2) The supports of insider's distributions for high and low types cannot be disjoint. (3) The support of F_{out} coincides with the union of the supports of F_{in}^H and F_{in}^L.* [29] *(4) There is no probability mass by the insider at the minimum boundary for any of the distributions.*

The claims are proven separately below.

Claim 1 *Insiders' two distributions for low and high type cannot have overlapping parts, that is, $\overline{F}_i^H \leq \underline{F}_i^L$.*

Proof. Intuitively, the proof establishes that insider cannot be indifferent between two fees offered to high-types to low-types at the same time, while facing with the (same) probability distribution F_{out}. Suppose $\underline{F}^L < \overline{F}^H$. As we restrict to independent randomizations, at any $\phi \in \left(\underline{F}^L, \overline{F}^H\right)$, insider must be indifferent independently for the low-types and the high-types. That means, for an arbitrary ϕ and $\phi' \in \left(\underline{F}^L, \overline{F}^H\right)$

$$\begin{aligned}(1-F_{out}(\phi))(1-\alpha)(1-\beta)(\phi-c^\theta)+\alpha(1-\beta)(\phi-c^\theta)\\ =(1-F_{out}(\phi'))(1-\alpha)(1-\beta)(\phi'-c^\theta)+\alpha(1-\beta)(\phi'-c^\theta)\end{aligned}$$

and

$$\begin{aligned}(1-F_{out}(\phi))(1-\alpha)\beta(\phi-c^\theta)+\alpha\beta(\phi-c^\theta)\\ =(1-F_{out}(\phi'))(1-\alpha)\beta(\phi'-c^\theta)+\alpha\beta(\phi'-c^\theta)\end{aligned}$$

[29] One must be careful with the mathematical language here, because the randomization does not necessarily happens over a compact interval. Notice however that the usual definition of the support of a distribution, *supp*(F) is *the closure* of the set of possible values with nonzero mesure. That is, specifically, $\overline{F} := \sup\{x : F(x) < 1\} \in supp(F)$ even if $\overline{F}$ is never played.

After simplifications, and rearranging terms, for each $\theta \in \{L, H\}$ we obtain:

$$(1 - F_{out}(\phi))(1-\alpha)(\phi - c^\theta) + \alpha(\phi - c^\theta) = (1 - F_{out}(\phi'))(1-\alpha)(\phi' - c^\theta) + \alpha(\phi' - c^\theta)$$
$$(1-\alpha)(\phi - \phi') - (1-\alpha)\left[F_{out}(\phi)(\phi - c^\theta) - F_{out}(\phi')(\phi' - c^\theta)\right] = \alpha(\phi' - \phi)$$
$$\left[F_{out}(\phi)(\phi - c^\theta) - F_{out}(\phi')(\phi' - c^\theta)\right] = -\frac{\alpha(\phi' - \phi)}{(1-\alpha)} + (\phi - \phi')$$
$$c^\theta \left(F_{out}(\phi') - F_{out}(\phi)\right) = -\frac{\alpha(\phi' - \phi)}{(1-\alpha)} + (\phi - \phi') - F_{out}(\phi)(\phi) + F_{out}(\phi')\phi'$$

From the last equation the contradiction is obvious, as the RHS is constant, while the LHS is different for H and L. ■

Claim 2 *Insider's low and high distributions cannot be disjoint with a gap between the two intervals, i.e.* $\overline{F}_i^H \geq \underline{F}_i^L$.

Proof. Suppose they are disjoint, $\overline{F}_{in}^H < \underline{F}_{in}^L$. There cannot be any probability mass by outsider on any $\phi < \underline{F}_{in}^L$, as it would find optimal to put this mass on $\underline{F}_{in}^L - \epsilon$ instead. However, this cannot be optimal for insider. As insider would win with the same probability over high types for every $[\overline{F}_{in}^H, \underline{F}_{in}^L)$, it would find it optimal to move some probability mass to the left, and increase its payoff. Contradiction to equilibrium. ■

Notice that Claim 1 and 2 together implies that $\overline{F}_{in}^H = \underline{F}_{in}^L$.

Claim 3 $\underline{F_{in}^H} = \underline{F_{out}}$ *and* $\overline{F_{in}^L} = \overline{F_{out}}$

Proof. The proof is analogous to previous results. Whenever $\underline{F}_{in}^H < \underline{F}_{out}$, insider has incentives to put the probability mass on $[\underline{F}_{in}^H < \underline{F}_{out})$ to F_{out} instead. Similarly, if $\underline{F}_{in}^H > \underline{F}_{out}$, outsider would place the mass - for example - on $(\underline{F}_{out}, \frac{\underline{F}_{out} + \underline{F}_{in}^H}{2}$ to $\frac{\underline{F}_{out} + \underline{F}_{in}^H}{2}$ instead. The equality of upper boundaries can be seen analogously. ■

Claim 4 *There is no probability mass by the insider at the minimum boundary for any of the distributions.*

Proof. (Sketch) Suppose there is mass on $\underline{F}_{in}^H$ by insider-high. Then outsider loses at $\underline{F}_{in}^H$ with some positive probability. Instead of playing $\underline{F}_{in}^H$, it could put all probability mass to $\underline{F} - \epsilon$, and win with probability 1. Contradiction to equilibrium.

Suppose there is mass on $\hat{\phi}$ by insider-low. Then outsider loses low-types with some positive probability. Instead, it could place all mass at $\hat{\phi} - \epsilon$, win all low-types, and not lose on high-types. Contradiction to equilibrium. ■

3.C.2.2 PART B - Only one binding minimax payoff

Lemma 12 *Only one of insider's "minimax-profit" can be binding. To be specific, either $\underline{\pi}(L)$ or $\underline{\pi}(H)$ is binding, where*

$$\pi_L^0 = \alpha\beta\left(\overline{\phi} - c^L\right)$$
$$\pi_H^0 = \alpha(1-\beta)\left(\overline{\phi} - c^H\right)$$

Proof. Suppose that some ϕ (denoted $\hat{\phi}$) is played under both H and L distribution. This exists, and as a consequence of previous claim, $\hat{\phi} = \overline{F}_i^H = \underline{F}_i^L$. Then the required probability mass on F_o from the right of $\hat{\phi}$ to make the bank indifferent between playing $\hat{\phi}$ and their minimax payoff for low and high type respectively is different, as follows:

$$\rho_{out}^L \beta(\hat{\phi} - c^L) = \alpha\beta\left(\overline{\phi} - c^L\right)$$
$$\rho_{out}^H (1-\beta)(\hat{\phi} - c^H) = \alpha(1-\beta)\left(\overline{\phi} - c^H\right)$$

As $\rho_{out}^L \neq \rho_{out}^H$, (it can be shown that $\rho_L < \rho_H$) only one of them can be binding. Specifically, if insider plays $\overline{\phi}$ with positive mass for the low-type, then $\underline{\pi}(L)$ must be the binding one. ■

3.C.2.3 PART C - The main proof, case $\alpha \leq \beta$

As per our assumption, insider randomizes independently for low-type and for high-type, facing with the same probability distribution F_{out}. This generates two independent indifference conditions. Suppose[30] that $\underline{\pi}(L)$ is binding. Then the indifference condition for low-types is:

$$\left(1 - F_{out}^L(\phi)\right)(1-\alpha)\beta\left(\phi - c^L\right) + \alpha\beta\left(\phi - c^L\right) = \alpha\beta\left(\overline{\phi} - c^L\right)$$

This defines the upper part of the outsider's CDF (against 'low'-types):

$$F_{out}^L(\phi) = \frac{\phi - c^L - \alpha\left(\overline{\phi} - c^L\right)}{(1-\alpha)(\phi - c^L)} = \frac{1}{1-\alpha} - \frac{\alpha}{1-\alpha}\frac{\overline{\phi} - c^L}{\phi - c^L}$$

Suppose that the cutoff-point between F_{in}^H and F_{in}^L is some $c^L < \hat{\phi} < \overline{\phi}$ and that the outsider's CDF over $\left(\underline{F}, \hat{\phi}\right)$ takes the functional form $F_{out}^H(\cdot)$. In what follows, we write everything as a function of an *arbitrary cutoff-value* $\hat{\phi}$.

[30] This is the only *guess-and-verify in equilibrium* part left in the proof, which is enough for existence, but cannot guarantee the uniqueness.

The insider's indifference condition for the HIGH-types can be written as

$$\left(1 - F^H_{out}(\phi)\right)(1-\alpha)(1-\beta)\left(\phi - c^H\right) + \alpha(1-\beta)\left(\phi - c^H\right)$$
$$= (1-\alpha)(1-\beta)(1 - F^L_{out}(\hat{\phi}))\left(\hat{\phi} - c^H\right) + \alpha(1-\beta)\left(\hat{\phi} - c^H\right)$$

We can express the CDF $F^H_{out}(\phi)$ as a function of the extra argument $\hat{\phi}$.

$$F^H_{out}(\phi, \hat{\phi}) = \frac{(-\phi\hat{\phi} + c^L(\phi - \alpha\hat{\phi}) + \alpha\hat{\phi}\overline{\phi} + c^H((-1+\alpha)c^L + \hat{\phi} - \alpha\overline{\phi}))}{((-1+\alpha)(c^H - \phi)(c^L - \hat{\phi}))}$$

Now we calculate the lower bound of the distribution, again, as a function of $\hat{\phi}$, through the following equation:

$$F^H_{out}(\phi, \hat{\phi}) = 0 \tag{3.18}$$

This gives the (unique) solution for the lower boundary $\underline{F}_{out}$:

$$\underline{F}_{out}(\hat{\phi}) = \frac{(\alpha\hat{\phi}(c^L - \overline{\phi}) + c^H(c^L - ac^L - \hat{\phi} + \alpha\overline{\phi}))}{(c^L - \hat{\phi})}$$

As at the lower boundary the outsider wins with probability 1 , we can compute outsider's payment as a function of $\hat{\phi}$, denoted by π^0_{out}.

$$\pi^0_{out} = \frac{((-1+\alpha)(-\beta(c^H - c^L)(c^L - \hat{\phi}) + \alpha(c^H - \hat{\phi})(c^L - \overline{\phi})))}{(c^L - \hat{\phi})}$$

The critical threshold value $\hat{\phi}$ must also fulfil that at this value, the insider-low distribution has no mass-point:

$$F^L_{in}(\hat{\phi}) = 0$$

Recall outsider's indifference condition, adapted to these two specific points:

$$(1-\alpha)\beta(\hat{\phi} - c^L) = \overline{\rho_{in}}(1-\alpha)\beta\left(\overline{\phi} - c^L\right)$$

which gives the value of the mass-point of insider's LOW-distribution, as a function of $\hat{\phi}$.

$$\overline{\rho}_{in} = \frac{\hat{\phi} - c^L}{\overline{\phi} - c^L}$$

Whenever outsider charges $\overline{\phi} - \epsilon$, with $\epsilon \to 0$, it obtains profit

$$\overline{\rho}(1-\alpha)\beta(\overline{\phi} - c^L)$$

Let $\pi_{out}^{0,L}$ the limit of this profit with $\epsilon \to 0$. Finally, $\hat{\phi}$ is determined through the following indifference condition equation:

$$\pi_o^{0,H} = \frac{\hat{\phi} - c^L}{\overline{\phi} - c^H}(1-\alpha)\beta(\overline{\phi} - c^L)$$

The solution of this equation is

$$\hat{\phi}^\star = c^L + \frac{\alpha\left(\overline{\phi} - c^L\right)}{\beta}$$

Substituting back $\hat{\phi}^\star$ into F_{out}^H gives the probability mass from the left of the critical value, that is, the probability that insider loses the high-types by playing $\overline{\phi}$. This probability is

$$F_{out}^H(\hat{\phi}) = \frac{1-\beta}{1-\alpha}$$

which implies that insider wins the complementary probability $(\frac{\beta-\alpha}{1-\alpha})$, generating the following profit from high-types:

$$\pi_{in} = (1-\beta)(\alpha(\overline{\phi} - c^L) + \beta\Delta c)$$

Notice that this new equilibrium payoff exceeds the "minimax payoff" $\alpha(1-\beta)(\overline{\phi} - c^H)$ whenever $\alpha > \beta$. *This confirms the claim that the minimax payoff for the low-type is binding whenever* $\alpha > \beta$. The insider can achieve larger profit in equilibrium, than it could do with serving only naive high-types!

—*Outsider-HIGH distribution*—

Based on this, we can write insider's indifference condition as

$$(1-F_o^H)(1-\alpha)(1-\beta)(\phi - c^H) + \alpha(1-\beta)(\phi - c^H) = (1-\beta)(\alpha(\overline{\phi} - c^L) + \beta\Delta c)$$

This pins down the outsider's CDF over the interval $[\underline{F}, \hat{\phi})$.

$$\begin{aligned} F_{out}^H(\phi) &= \frac{\phi - c^H - \beta\Delta c - \alpha(\overline{\phi} - c^L)}{(1-\alpha)(\phi - c^H)} \\ &= \frac{1}{1-\alpha} - \frac{\alpha(\overline{\phi} - c^L) + \beta\Delta c}{(1-\alpha)(\phi - c^H)} \\ &= \frac{1}{1-\alpha} - \frac{\alpha(\overline{\phi} - c^H) + (\beta-\alpha)\Delta c}{(1-\alpha)(\phi - c^H)} \end{aligned}$$

—Insider's distributions—

Insider's distributions are derived using outsider's indifference condition: at every ϕ, outsider must be indifferent between playing ϕ or its alternative payoff, which is pinned down by the mass-point by insider on $\overline{\phi}$. That leads to the two independent indifference conditions:

$$\left(1 - F^H_{in}(\phi)\right)(1-\alpha)(1-\beta)(\phi - c^H) + (1-\alpha)\beta(\phi - c^L) = (1-\alpha)\alpha(\overline{\phi} - c^L)$$
$$\left(1 - F^L_{in}(\phi)\right)(1-\alpha)\beta(\phi - c^L) = (1-\alpha)\alpha(\overline{\phi} - c^L)$$

so the respective distributions are:

$$\begin{aligned} F^H_{in}(\phi) &= \frac{(\phi - c^H - \beta\Delta c - \alpha(\overline{\phi} - c^L))}{((1-\beta)(\phi - c^H))} \\ &= \frac{1}{1-\beta} - \frac{\alpha(\overline{\phi} - c^L) + \beta\Delta c}{(1-\beta)(\phi - c^H)} \end{aligned}$$

and

$$\begin{aligned} F^L_{in}(\phi) &= \frac{\beta(\phi - c^L) - \alpha(\overline{\phi} - c^L)}{\beta(\phi - c^L)} \\ &= 1 - \frac{\alpha}{\beta}\frac{\overline{\phi} - c^L}{\phi - c^L} \end{aligned}$$

—Verify distributions —

It is useful to check some properties of the derived distributions. With simple algebra it is possible to check that (i) $F^H_{in}(\phi) = 0$ and $F_{out}(\phi) = 0$ has the same solutions; (ii) $F^H_{in}(\hat{\phi}) = 1$ and $F^L_{in}(\hat{\phi}) = 0$; and (iii) $F_{out}(\overline{\phi}) = 1$.

3.C.2.4 PART D - The main proof, case $\alpha > \beta$

The anticipated equilibrium: insider is mixing over $(\phi'_{in}, \overline{\phi})$ for the High-type, and according to a degenerate distribution $\phi^L_{in} = \overline{\phi}$ for the Low-type. Outsider plays F_o over $\phi'_{in}, \overline{\phi}$. Insider will be placing a probability mass on $\overline{\phi}$.

Customers' behaviour: given this pricing equilibrium, low-type customers switch to outsider with probability 1, while low-type customers according to some probability.

Next, we establish overdraft fee distributions using the outsider's and the insider's indifference property.

Outsider's indifference property: Because there is no mass-point by the insider at $\underline{F}$, by bidding the lower boundary ϕ'_{in} the outsider wins and sophisticated customers switch with probability 1. The profit for outsider in this case is:

$$\pi^0_{out} = (1-\alpha)\left(\beta(\phi'_{in} - c^L) + (1-\beta)(\phi'_{in} - c^H)\right)$$

For any $\phi > \phi'_{in}$ it must be the case that

$$\mathbb{E}(\pi_{out}(\phi)) = \pi^0_{out} \tag{3.19}$$

where

$$\begin{aligned}\mathbb{E}\left(\pi_{out}(\phi)\right) &= \overbrace{Pr\left[\phi < \phi^H_{in}\right]}^{\text{O wins}} * \overbrace{(1-\alpha)\pi^{LH}(\phi)}^{\text{profit from both types}} + \overbrace{Pr\left[\phi_{out} > \phi^H_{in}\right]}^{\text{I wins}} * \overbrace{(1-\alpha)\beta\pi^L(\phi)}^{\text{profit from low types}} \\ &= (1 - F_{in}(\phi))(1-\alpha)\left(\beta\pi^L(\phi) + (1-\beta)\pi^H(\phi)\right) + F_{in}(\phi)(1-\alpha)\beta\pi^L(\phi) \\ &= (1-\alpha)\pi^{LH}(\phi) - F_{in}(\phi)(1-\alpha)(1-\beta)\pi^H(\phi)\end{aligned}$$

So the equilibrium equation (3.19) simplifies to

$$\pi^{LH}(\phi'_{in}) = \pi^{LH}(\phi) - F_{in}(\phi)(1-\beta)\pi^H(\phi)$$

Leading to the following outsider CDF:

$$F_{in}(\phi) = \frac{\pi^{LH}(\phi) - \pi^{LH}(\phi'_{in})}{(1-\beta)\pi^H(\phi)}$$

According to our specification $\pi^{LH}(\phi) = \phi - c^{LH}$, so after substitutions:

$$F_{in}(\phi) = \frac{\phi - \phi'_{in}}{(1-\beta)(\phi - c^H)} = \frac{1}{1-\beta} - \frac{\alpha}{1-\beta} \cdot \frac{\overline{\phi} - c^H}{\phi - c^H} \tag{3.20}$$

This satisfies the requirement that $F_{in}(\phi'_{in}) = 0$. Solving $F_{in}(\phi) = 1$ for ϕ gives:

$$\phi^{max} = \frac{\phi'_i}{\beta} - \frac{1-\beta}{\beta}(c^H) = \overline{\phi} + \frac{\alpha - \beta}{\beta}\left(\overline{\phi} - c^H\right)$$

, from which it is clear that

$$\phi^{max} < \overline{\phi} \quad \Leftrightarrow \quad \alpha < \beta \quad \text{and} \quad \phi^{max} > \overline{\phi} \quad \Leftrightarrow \quad \alpha > \beta$$

This implies that whenever $\alpha > \beta$, the upper boundary of the mixture is $\overline{\phi}$, and there is a mass-point by insider on $\overline{\phi}$. In this case the CDF at $\overline{\phi}$ is

$$F_{in}(\overline{\phi}) = \frac{1-\alpha}{1-\beta}$$

therefore, the probability mass on $\overline{\phi}$ must be

$$p_{in} = 1 - \frac{1-\alpha}{1-\beta} = \frac{\alpha-\beta}{1-\beta} \tag{3.21}$$

Insider's indifference property: Outsider will never bid above $\overline{\phi}$, so when insider's offer is $\left(\overline{\phi}, \overline{\phi}\right)$ then it serves only myopes for both types and its profit is

$$\pi_{in}^0 := \alpha\pi^{LH}(\overline{\phi}) = \alpha\left(\beta\pi^L(\overline{\phi}) + (1-\beta)\pi^H(\overline{\phi})\right)$$

To account for the possibility that $\overline{F}_{in} < \overline{\phi}$, suppose outsider is mixing over $[\phi'_{in}, u]$ according to F_{out}. For any offer by the insider to the high types $\phi \in \{\phi'_{in}, u\}$ the insider wins with probability $Pr(\phi < \phi_{out}) = 1 - F_{out}(\phi)$ and obtains profit

$$\begin{aligned}\mathbb{E}\pi &= (1-F_{out})\left((1-\beta)\pi^H(\phi) + \alpha\beta\pi^L(\overline{\phi})\right) + F_{out}\alpha\left((1-\beta)\pi^H(\phi) + \beta\pi^L(\overline{\phi})\right) \\ &= \alpha\beta\pi^L(\overline{\phi}) + (1-\beta)\pi^H(\phi) - F_{out}(1-\alpha)(1-\beta)\pi^H(\phi)\end{aligned}$$

The equilibrium condition:

$$(1-\beta)\pi^H(\phi) - F_{out}(1-\alpha)(1-\beta)\pi^H(\phi) = \alpha(1-\beta)\pi^H(\overline{\phi})$$

$$F_{out}(\phi) = \frac{\pi^H(\phi) - \alpha\pi^H(\overline{\phi})}{(1-\alpha)\pi^H(\phi)} = \frac{1}{1-\alpha} - \frac{\alpha}{1-\alpha} \cdot \frac{\overline{\phi} - c^H}{\phi - c^H}$$

Solving $F_{out}(\phi) = 0$ gives

$$\phi = \alpha\overline{\phi} + (1-\alpha)c^H = \phi'_{in}$$

which implies the density satisfies $F(\phi'_{in}) = 0$ and $F(\overline{\phi}) = 1$, so there is no mass-point in outsiders' CDF if they mix over $[\phi'_{in}, \overline{\phi}]$.

3.C.3 Probabilities and expected values

This section provides some additional analytical characterization of the equilibrium of the full model. The first result establishes switching probabilities. In the limit of $\alpha \to 0$ and $\beta \to 0$ we get back the results from the pure Adverse Selection and Customer naiveté models respectively.

Lemma 13 *The probability that outsider (insider) wins is as follows:*

- *Whenever $\alpha > \beta$ (mass point by insider for low-types)*

$$Prob[\phi_{out} < \phi_{in}^H] = \frac{1+\alpha-2\beta}{2(1-\beta)} \qquad Prob[\phi_{in}^H \le \phi_{out}] = \frac{1-\alpha}{2(1-\beta)}$$

- *Whenever $\alpha < \beta$, outsider (insider) wins HIGH types with probability:*

$$Prob[\phi_{out} < \phi_{in}^H] = \frac{1-\beta}{2(1-\alpha)} \qquad Prob[\phi_{in}^H \le \phi_{out}] = \frac{1+\beta-2\alpha}{2(1-\alpha)}$$

and wins LOW types with probability:

$$Prob[\phi_{out} < \phi_{in}^L] = \frac{\alpha^2+(\beta-2)\beta}{2(\alpha-1)\beta} \qquad Prob[\phi_{in}^L \le \phi_{out}] = \frac{(\beta-\alpha)^2}{(2(1-\alpha)\beta)}$$

Proof. *We want to calculate $Prob[\phi_{out} < \phi_{in}^H]$. First, suppose $\alpha > \beta$. Insider is playing $\overline{\phi}$ with probability ρ_{in}, in that case outsider wins with probability 1. In case insider is mixing (with probability $1-\rho_{in}$) continuously, we have*

$$Prob[\phi_{out} < \phi_{in}^H] = \int_{\phi_{in}'}^{\overline{\phi}} \int_{\phi_{in}'}^{\phi_{in}^H} f_{o,i} d\phi_{out} d\phi_{in}^H \tag{3.22}$$

where

$$f_{in} = F_{in}' = \frac{\alpha}{1-\beta}\frac{\overline{\phi}-c^H}{(\phi-c^H)^2}$$
$$f_{out} = F_{out}' = \frac{\alpha}{1-\alpha}\cdot\frac{\overline{\phi}-c^H}{(\phi-c^H)^2}$$

The joint PDF is, due to the independence assumption:

$$f_{o,i}(\phi_{out},\phi_{in}^H) := \frac{\alpha^2}{(1-\alpha)(1-\beta)}\cdot\frac{\left(\overline{\phi}-c^H\right)^2}{(\phi_{out}-c^H)^2\left(\phi_{in}^H-c^H\right)^2}$$

The internal integral of (3.22), *(with respect to ϕ_{out}) is*[31]

$$\int_{\phi'_{in}}^{\phi_{in}^H} f_{oi} d\phi_{out} = -\frac{\alpha^2}{(1-\alpha)(1-\beta)} \cdot \frac{(\overline{\phi}-c^H)^2}{(\phi_{out}-c^H)(\phi_{in}^H-c^H)^2}$$
$$= -\frac{\alpha^2}{(1-\alpha)(1-\beta)} \cdot \frac{\left(\overline{\phi}-c^H\right)^2}{\left(\phi_{in}^H-c^H\right)^2} \cdot \left(\frac{1}{\phi_{in}^H-c^H} - \frac{1}{\phi'_{in}-c^H}\right)$$

For the full integral we use the following interim results:

$$\int_{\phi'_{in}}^{\overline{\phi}} \frac{1}{(\phi_{in}^H-c^H)^3} d\phi_{in}^H = \left[-\frac{1}{2(\phi_{in}^H-c^H)^2}\right]_{\phi'_{in}}^{\overline{\phi}} = -\frac{1}{2}\frac{\alpha^2-1}{\alpha^2}\frac{1}{(\overline{\phi}-c^H)^2}$$
$$\int_{\phi'_{in}}^{\overline{\phi}} \frac{1}{(\phi_{in}^H-c^H)^2} d\phi_{in}^H = \left[-\frac{1}{\phi_{in}^H-c^H}\right]_{\phi'_{in}}^{\overline{\phi}} = -\frac{\alpha-1}{\alpha}\frac{1}{\overline{\phi}-c^H}$$

This leads to the following formula for the likelihood that outsider wins:

$$LIK[\phi_{out} < \phi_{in}^H] = \frac{1-\alpha}{2(1-\beta)}$$

The overall probability, including the mass point, is:

$$Prob[\phi_{out} < \phi_{in}^H] = \frac{1+\alpha-2\beta}{2(1-\beta)}$$

Now suppose $\alpha < \beta$. In this case outsider is playing ϕ^{max} with probability mass defined in Theorem 4. If playing the mass point, outsider loses with certainty. With the complementary probability, both insider and outsider is mixing continuously over $(\phi'_{in}, \phi^{max}]$ and the expression for the likelihood, the joint PDF is exactly as before, but with an outer integral boundary changed to ϕ^{max}.

$$Prob[\phi_{out} < \phi_{in}^H] = \int_{\phi'_{in}}^{\phi^{max}} \int_{\phi'_{in}}^{\phi_{in}} f_{oi} d\phi_{out} d\phi_{in}$$

This has the following solution for the likelihood, using the same steps as before:

$$LIK[\phi_{out} < \phi_{in}^H] = \frac{1-\beta}{2(1-\alpha)}$$

■

Next, we calculate conditional expected payments.

[31] We use here the formula $\int \frac{f'}{f^2} = -\frac{1}{f}$, in the expression $f' = 1$.

Lemma 14 *Conditional on winning the competition, the outsider and insider obtains the following expected overdraft fee if* $\alpha \geq \beta$*:*

$$\mathbb{E}[\phi_{out} < \phi_{in}^H]^{(1)} = c^H + \frac{2\alpha}{1+\alpha-2\beta}\left(\overline{\phi} - c^H\right) + \frac{2\alpha\beta\ln[\alpha]}{(1-\alpha)(1+\alpha-2\beta)}\left(\overline{\phi} - c^H\right)$$
$$\mathbb{E}[\phi_{in}^H < \phi_{out}]^{(1)} = c^H + \frac{2\alpha}{1-\alpha}\left(\overline{\phi} - c^H\right) + \frac{2\alpha^2\ln[\alpha]}{(1-\alpha)^2}\left(\overline{\phi} - c^H\right)$$

Expected overdraft fees in principle can be calculated for the $\alpha \leq \beta$ case, but it is analytically very difficult, and left for future work.

Proof.

We calculate expected fee offered by insider resp. outsider *conditional on winning the price competition*, that is, the following conditional expected values:

$$\mathbb{E}[\phi_{out}|\phi_{out} < \phi_{in}^H]$$
$$\mathbb{E}[\phi_{in}^H|\phi_{in}^H \leq \phi_{out}]$$

Whenever mass points are present at the boundary, we need to consider the total expected value as a sum of two components. From Theorem 4, if $a > b$ insider places a positive mass $\rho_{in} = \frac{\alpha-\beta}{1-\beta}$ on $\overline{\phi}$ and in this case outsider wins at every value ϕ_{out} over the mixture. The expected values are:

$$\mathbb{E}[\phi_{out}|\phi_{out} < \phi_{in}^H] = \frac{1}{Pr[\phi_{out} < \phi_{in}^H]}\left(\int_{\phi'_{in}}^{\overline{\phi}}\int_{\phi'_{in}}^{\phi_{in}^H} \phi_{out} f_{oi} d\phi_{out} d\phi_{in}^H + \rho_i \int_{\phi'_{in}}^{\overline{\phi}} \phi_{out} f_{\phi_{out}} d\phi_{out}\right)$$
$$= \frac{2(1-\beta)}{1+\alpha-2\beta}\left(\Phi + \frac{\alpha-\beta}{1-\beta}\mathbb{E}[\phi_{out}]\right)$$

and

$$\mathbb{E}[\phi_{in}^H|\phi_{in}^H < \phi_{out}] = \frac{\int_{\phi'_{in}}^{\overline{\phi}}\int_{\phi'_{in}}^{\phi_{out}} \phi_{in} f_{oi} d\phi_{in} d\phi_{out}}{\int_{\phi'_{in}}^{\overline{\phi}}\int_{\phi'_{in}}^{\phi_{out}} f_{oi} d\phi_{in} d\phi_{out}}$$

Because the joint PDF is symmetric in the two variables, the double-integral part is also symmetric, and we have

$$\Phi := \int_{\phi'_{in}}^{\overline{\phi}} \int_{\phi'_{in}}^{\phi_{in}} \phi_{out} f_{oi} d\phi_{out} d\phi_{in} = \int_{\phi'_{in}}^{\overline{\phi}} \int_{\phi'_{in}}^{\phi_{out}} \phi_{in} f_{oi} d\phi_{in} d\phi_{out}$$
$$= \frac{1}{2}\left(\frac{(3\alpha-1)(-c^H)+2\alpha\overline{\phi}}{1-\beta} + \frac{2\alpha^2\ln(\alpha)(\overline{\phi}-c^H)}{(1-\alpha)(1-\beta)}\right)$$
$$= \frac{1}{2(1-\beta)}\left((1-\alpha)c + 2\alpha\left(\overline{\phi}-c^H\right) + \frac{2\alpha^2\ln(\alpha)(\overline{\phi}-c^H)}{1-\alpha}\right)$$

It is relatively easy to calculate the expected ϕ_{out} over the entire mixture

$$\mathbb{E}(\phi_{out}) := \int_{\phi'_{in}}^{\overline{\phi}} \phi_{out} f_{\phi_{out}} d\phi_{out} = c^H - \frac{\alpha\ln(\alpha)}{1-\alpha}\left(\overline{\phi}-c^H\right)$$

Therefore the overall expectations are:

$$\mathbb{E}[\phi_{out}|\phi_{out} < \phi_{in}^H] := \frac{2(1-\beta)}{1+\alpha-2\beta}\left(\Phi + \frac{\alpha-\beta}{1-\beta}\mathbb{E}(\phi_{out})\right) = \frac{2(1-\alpha)^2}{(1-\beta)^2}\Phi + \frac{\alpha-\beta}{1-\beta}\mathbb{E}(\phi_o)$$
$$\mathbb{E}[\phi_{in}^H|\phi_{in}^H < \phi_{out}] := \frac{2(1-\beta)}{1-\alpha}\Phi$$

$\blacksquare$

Chapter 4

Fire-sale in a liquidation game with leverage requirements

4.1 Introduction

In the aftermath of the financial crisis, the view that interconnectedness is an important determinant of financial stability became conventional wisdom among academics and policy-makers. A large theoretical and empirical literature on systemic risk started to emphasize how various forms of business relations in the financial sector can turn to a transmission channel through which shocks propagate in the financial system, eventually leading to systemic bank failures and causing real economic losses. For example, since the seminal contribution of Allen and Gale (2000), it is well known that interbank markets facilitate liquidity risk sharing, but can also be the source of 'direct' contagious failures which may eventually destabilize the financial system as a whole.

Another potential layer of interconnectedness, which is the subject of this paper, is indirect linkages through common investments, or *'asset commonalities'*. If an investor is forced to liquidate their asset due to some funding pressure, prices may depart from fundamental values. Mark-to-market evaluation of portfolios forces other investors of the same asset to re-evaluate their portfolio, which decreases equity value. In turn, the drop in equity induces additional funding pressure, and those - otherwise healthy - institutions may be forced to engage in further asset liquidation. With multiple owners of the same asset class, the situation is exacerbated by a coordination problem: if many investors find it optimal to liquidate their asset at the same time, the price drop may be severe enough forcing even more institutions to sell, and pushing prices into a downward price spiral. Furthermore, institutions trying to avoid losses may find it optimal to sell other (otherwise unaffected) assets, transmitting the shock to even more sectors and institutions, expanding isolated problems to a potentially system-wide contagion. There is a strong feedback-effect towards the direct contagion mechanism as well: depressed asset prices and increased volatility raises haircuts on these contaminated assets in the overnight repo market, which accelerates the 'dry-up of liquidity' on the interbank market, and reinforces the need for forced asset sale at the first place.

In this paper we explicitly model the asset liquidation *decision* of financial institutions under funding pressure in a duopoly settings, when multiple asset classes are available to adjust the portfolio. In the model, the 'funding pressure', which is the key market friction behind this phenomenon, is captured by a leverage constraint: following an asset-price shock, the banking system may be forced to engage in systemic deleveraging to restore leverage targets by selling assets and repaying debt.[1] The relevance of this mechanism in propagating crises is convincingly demonstrated - both empirically and theoretically - in an influential paper by Adrian and Shin [2010]. The investment portfolio on banks' balance sheets differ in ex-ante liquidity, measured as the market price impact following an asset sale during 'normal times'. Equity-maximizer financial institutions adjust their portfolio by choosing to sell assets such that the impact on equity is minimized. In the presence of asset commonalities, if all banks end up selling the same asset class ('commonality'), liquid assets suddenly may appear illiquid and can be sold only at a significant fire-sale discount, a phenomenon which was widely observed during the financial crisis. This endogenous determination of the fire-sale price has to be taken into consideration by rational financial institutions.

The joint deleveraging decision of interlinked financial institutions induces a non-cooperative game which we dub '*the liquidation game*'. The main result of this paper is that as long as the equilibrium liquidation decision of the banks is non-trivial in the sense that liquidating only one single asset does not strictly dominate, the emerging Nash-equilibrium is not Pareto-optimal. Individual banks could achieve higher ex-post equity value by choosing another feasible liquidation strategy, which, however, cannot be maintained as an equilibrium. The market outcome in equilibrium is reminiscent to a Prisoner's dilemma: cooperation, which in this context would mean self-restraint in selling the more liquid asset commonalities and relying more on idiosyncratic but less liquid assets to restore leverage, could increase the payoff for each players, but cannot be maintained as an equilibrium. In

[1]The model can be generalized to other sources of similar funding frictions: for example, there is strong empirical evidence that decreasing value of asset-under-management induces fund outflows for investment funds, which forces them to liquidate part of the portfolio, even at diminished prices.

the unique Nash-equilibrium banks 'defect', and over-liquidate the commonality.

The comparison of the equilibrium and the social planner's optimal solution reveals an even more striking feature: the potential loss from the inefficient equilibrium may even be larger, if markets appear to be ex-ante more liquid. Intuitively, more liquid commonality raises the incentives to tilt the liquidation strategy towards that asset class, which leads to an even larger equilibrium price effect, and further diminishes equity. This finding has slightly uncomfortable consequences for financial stability: higher liquidity, although almost unanimously called for by policy-makers after the crisis, can even be detrimental in highly integrated markets, if fire-sale decisions following a potential shock are jointly determined in an equilibrium.

There is a growing literature on deleveraging and asset liquidation strategies. Many of these papers extend the now-standard framework of Eisenberg and Noe [2001], originally designed to characterize a payment equilibrium in a network of borrowing-lending relationships between economic agents. The general models provide important characterization of existence and uniqueness of equilibrium, but do not emphasize the strategic inefficiencies arising as a consequence of decentralized decision making, which is the main contribution of this paper. Specifically, in most of the existing models banks respond *deterministically* to shocks, and strategic interactions are not taken into consideration. In contrast to the existing literature and complementing some recent results, the purpose of this paper is to characterize the welfare aspects of *equilibrium* deleveraging, and demonstrate possible inefficiencies. We focus on contagion due to asset fire sale in a multiple asset commonality settings, and ignore direct contractual obligations or interbank markets and cross-holdings on the liability side. Although we acknowledge the importance of the mutually reinforcing effects between contagion channels due to multiple network layers, our focus is the welfare loss caused by the *behaviour* in crisis, namely the deleveraging *decisions*, instead of mechanical network externalities and other contagious effects which are more extensively studied elsewhere.

4.2 Literature review

Asset-price contagion: Theoretical models of deleveraging financial institutions describe how the presence of asset commonalities and mark-to-market evaluation can lead to negative fire-sale spillovers as a source of contagion in the banking sector. The seminal work is due to Cifuentes et al. [2005], who extend the Eisenberg-Noe framework (Eisenberg and Noe [2001]) of financial contagion with an illiquid asset and leverage-targeting banks. They prove the existence of clearing vector and fire sale prices as a solution of a fixed point problem.[2] The fixed-point technique of these papers become standard in the systemic risk and network literature. A highly influential paper by Acemoglu et al. [2015] shows that small shock and large shock regimes may have different effect on regular network structures in a framework which incorporates the possibility of early liquidation of banks' investment, and exogenous liquidation prices. Glasserman and Young [2015] considers various network structures and shock distributions to characterize the network effects of contagion. In a recent paper, Awiszus and Weber [2015] incorporates the contagious effects of bankruptcy costs, fire sale losses, interbank networks and cross-holdings into a comprehensive structural model of systemic risk. All of the above mentioned papers however feature only one illiquid asset, and banks are not strategic, decision-making actors. The recurring scheme in the literature is to assume proportional repayment on interbank loans, limited liability, and in case of cash shortage, a liquidation of the single illiquid investment, which may or may not be partial. Multiple asset extensions are rare in the literature since they complicate the analysis considerably.

Greenwood et al. [2015] considers a fairly general framework with multiple assets classes, but they assume that the bank maintains a fixed portfolio structure during deleveraging, and instead of solving for equilibrium, they consider domino-like contagion effects. Furthermore, they do not analyse the impact of heterogeneity in the ex-ante illiquidity of the assets. The model is designed for empirical applications, and their illustrative calculations, as well as the application and extension by Duarte and Eisenbach [2014] demonstrate the sizeable impacts of fire-sale losses on banks' equity. By making the proportional deleveraging assumption, however, in this model the bank is still a passive participant in the sense that it only suffers the exogenous shock and respond according to a deterministic rule. Therefore the methodology of these papers is insufficient to study equilibrium actions.

A similar model by Caccioli et al. [2014] focuses on the stability of various network structures. Using simulation techniques they demonstrate that the system-wide stability is 'hump-shaped' as a function of diversification (a result similar to Elliott et al. [2014]), and crucially depends on the leverage and the 'crowding' parameter (number of assets versus institutions) of the network. Their theoretical analysis of stability, which borrows its methodology from the epidemics literature, leads to similar results. In contrast to our paper, banks have no active role and an insolvent bank always liquidates all of its assets. Caccioli et al. [2015] applies a stress-testing framework to emphasize the importance of interactions of the two main contagion channels (and the two layers of networks), and show that the pres-

[2] The uniqueness of this equilibrium under mild conditions is proven by Amini et al. [2016].

ence of counterparty risk (direct connections), although not considered very risky on its own, strongly amplifies the risk inherent in common asset holdings.

A significant step toward genuinely multi-asset extensions is Chen et al. [2014], who study asset-price contagion with the possibility of rebalancing the investment portfolio. Their model characterizes the contagion chain caused by an exogenous change of the state of the economy. Fire-sale prices are determined endogenously through asset market equilibrium conditions. The authors' focus is optimal asset holding network structures, and they show that in a low-leverage regime more diversification, while in the high-leverage regime more idiosyncratic asset holdings are beneficial for the systemic contagion perspective. The most closely related papers to ours are Feinstein [2015] and Feinstein and El-Masri [2015] who consider equilibrium liquidation strategies as an explicit generalization of the Eisenberg-Noe framework, prove the existence of equilibrium, and characterize the equilibrium via numerical examples. Our paper considers a simpler setup, but goes further in characterizing the equilibrium outcomes and its inefficiency under certain circumstances, by comparing the equilibrium with a social optimum benchmark.

The network externalities also question the standard risk-mitigating effect of diversification, by introducing excess covariance due to common assets. For example, Raffestin [2014] derives analytically the covariances in a framework where banks are subject to stochastic shocks and connected by a network with home biases. He finds that intermediate number of bankruptcies are less likely under high diversification, but the probability of extreme failures is large, therefore little diversification may be socially optimal. Tasca et al. [2014] brings in leverage into the picture by using the Merton-model to calculate joint default probability of banks, and focus on the trade-off between the effects of leverage and diversification (asset commonalities). In the 'safe' regime diversification can compensate for increased leverage, but not in the 'risky' regime. The key regulatory insight is that the individually optimal diversification might be systemically under-diversified. Wagner [2011] shows that the risk of joint liquidation (essentially the coordination problem with asset commonalities) and the resulting fire-sale prices (due to limited cash in the market) creates incentives for investors to hold heterogeneous portfolios, so the classical optimality of full diversification breaks down.

Financial Networks and Systemic Risk: The paper is also related more generally to the research on network effects and systemic risk. Studies on the impact of interconnections between financial institutions and their role in propagating crises have been booming in the last decade, and it might be surprising why recognizing its importance and formal modelling started relatively late. The sluggish start of this stream of research may be attributable to the fact that mainstream economists had to overcome the classical view on the 'macroeconomic effects of microeconomic shocks', namely, that small idiosyncratic shocks cancel each other out on average, and macroeconomic consequences are negligible. Solid theoretical background to disprove this view is only given in the influential papers by Gabaix [2009] and Acemoglu et al. [2012]. Similarly, Stiglitz [2010] describes a fairly general framework to emphasise that full integration - as a tool for risk sharing - is generally not optimal in the presence of nonconvexities.

In the banking research, a large body of empirical and theoretical literature focuses on the direct (interbank) linkages as a mechanism for facilitating contagion. Early papers include Allen and Gale [2000] and Freixas et al. [2000]. For example, Allen and Gale [2000], building on their previous liquidity-based crisis models [Allen and Gale, 1998] show that the effect of liquidity shocks and the extent of a crisis depends crucially on the connections within banks: a more completely connected market is more robust than an incomplete network structure. On the contrary, others argue -especially the branch of literature which considers contagion in financial market as an *'epidemic'* and draws the analogy that a systemic crisis is similar to the spreading out of a disease - that dense interconnections increase the likelihood of a system-wide contagion. This approach naturally led to an extensive search for the *'key player'* (see Zenou [2014] for a recent comprehensive review), the banks who are *'too-interconnected-to-fail'*, and a vast simulation-based literature (Upper [2011]). Conflicting views are somewhat reconciled in the recent papers by Acemoglu et al. [2015] and Elliott et al. [2014] who give novel characterization of the effect of interbank network connections to systemic risk. Acemoglu et al. [2015] distinguishes two shock-size regimes, and shows that completely connected networks go through a phase transition, becoming from the most resilient to the least resilient networks as the shock switches from small to large. The basic intuition is that for small shocks the system-wide excess liquidity is sufficient to absorb it, and more connected networks can facilitate the utilization of system-wide cash reserves. However, for a large shock, 'weakly connected' networks turns out to be more resilient, because a second shock-absorber - senior claimants - can be forced to bear losses to protect the rest of the system. Elliott et al. [2014] explores the integration and diversification properties of network structures, and shows that intermediate levels of both integration and diversification makes the network more suspicious to system-wide contagion. Similarly to earlier intuition, larger diversification helps to utilize more counterparties to bear the losses of external shocks.

Fire-sale spillovers: Early theoretical models of fire sale go back to Shleifer and Vishny [1992], who characterize fire sale as a situation in which assets are sold on a price lower than their 'value in best use' (i.e. fundamental value). In this sense the concept is a strong relative to *asset illiquidity.* Their classic interpretation is the following: when firms under distress are forced to sell assets, it is likely that similar firms - potentially the highest valuation users of the asset - are also under pressure, and likely not be able to raise sufficient funding to purchase it. Therefore, it must be sold to outsiders, with lower valuation (for example, because of information asymmetries, or the lack of expertise to operate the given asset). In their original model the inability to raise funding by the insider firm comes from debt overhang (Hart [1993], Hart and Moore [1995]). Exacerbating the situation, it is very likely that under a sectoral shock all owners of a specific asset type has to liquidate simultaneously, causing an even larger downward price pressure.

This paper was followed by a large empirical and theoretical literature on fire sale and its consequences on asset prices, emergence of financial crises and the real macroeconomic feedback effects. This literature is recently reviewed by Shleifer and Vishny [2011]. Notable follow-up research in the context of recent crisis emphasizes the connection to the limits of arbitrage (e.g. Shleifer and Vishny [1997], Gromb and Vayanos [2002]): when fire sale occurs, financiers may not be able to distinguish illiquidity from fundamental price drop and withdraw funds from arbitrageurs exactly when they need it to exploit the mispricing. If this withdrawal is simultaneous, investors unwind positions simultaneously exacerbating the mispricing and causing severe fire sales.

Diamond and Rajan [2011] extends the literature with many interesting aspects. Fire sale offers highly profitable investment opportunities to arbitrageurs with liquid cash. In anticipation of future fire sale, those investors find it optimal to withhold from buying the asset now. As a result, prices should decrease immediately - even before the actual insolvency -, increasing the expected rate of return for the whole market, which explains the 'adverse effect of future illiquidity on current lending'. They show that management has strong incentives to risk-shifting: 'hold on' to the illiquid asset and risking the future fire sale and insolvency ('illiquidity seeking').

The fire sale phenomenon also links to the literature emphasizing the connections between market liquidity and funding liquidity. Acharya and Viswanathan [2011] emphasizes how the possibility of risk shifting implied by the combination of short-term debt and leverage leads to credit rationing and forced deleveraging, which - in the framework of limited market participation a'la Shleifer and Vishny [1992] and cash-in-the-market pricing a'la Allen and Gale [1994] leads to the increased severity of crises. Brunnermeier and Pedersen [2009] connects the market liquidity and funding liquidity more directly: tightening funding liquidity makes investors reluctant to open capital-intensive positions, which lowers the market liquidity and increases volatility. In response to the growing illiquidity, financiers - following their risk models - increase margins, reinforcing the negative spillover of prices.

Empirical evidence on asset-price contagion: There exists considerable empirical evidence that asset-price contagion exists. For example, De Marco [2013] estimates how the European sovereign debt crisis spilled over to a supply shock on commercial bank's asset side causing a contagion from one asset class to the other. Manconi et al. [2012] provides evidence of the contagion from securitized bonds to corporate bond market: when securitized bonds became toxic, institutional investors sold corporate bonds, lowering prices on an otherwise healthy asset class. Duarte and Eisenbach [2014] estimates that an exogenous price shock of repo-financed assets leads to a significant drop on the equity, using a panel-extension of the cross-sectional framework in Greenwood et al. [2015]. Jotikasthira et al. [2012] considers international fund flows and shows that global funds reallocate investments on fire sale prices as a reaction of changing fund flows and this reallocation takes place on fire sale prices especially on emerging markets, inducing a sizeable excess correlation between those markets, and also with the domestic country of the global funds.

Focusing a bit more closely to our research question, Merrill et al. [2014] analyses the residential mortgage-backed securities market during the crisis and shows that some institutions were in many cases indeed incentivized to sell the illiquid (i.e. further from fundamental price) asset due to the risk-sensitive capital requirement regulations: if the illiquid asset induces high capital requirements, the bank might be better off selling it rather than the liquid, but non-risky asset.

As an interesting insight, Cella et al. [2013] traces back the issue to the different objectives between agents with different investment horizons: during crisis periods short-term investors -in the fear of short-term price declines - are expected to coordinate on selling, which is not offset by the amount of liquidity provided by long-term investors (who, as usually argued, are probably also facing equity problems). Empirical tests show significant differences between price impact functions of assets primarily held by investors with various investment horizon. An implication for the theoretical research stream is to keep in mind that not only network *structure*, but the *type* of nodes (here: short- or long term investors) could be an important determinant of systemic risk.

Hau and Lai [2012] documents compelling evidence for significant price contagion via common asset *ownership* in the US stock market. They show that funds with large exposure to distressed financials were forced to sell *nonfinancial* assets on fire sale prices, mainly because of funding constraints, and find that discounts are the largest for those stocks that performed well during (i.e. were fundamentally *not* exposed to) the crises. This implies that banks indeed optimize and choose assets to avoid capital losses (i.e.: sell liquid assets, as illiquidity is understood as deviation from fundamentals), with obvious real-economic consequences.

The direct network connections also matter in selling decision: for example Favara and Giannetti [2015] shows that lenders differ in internalizing fire sale externalities: those with a larger share of collateralized debt internalize more the feedback effects on collateral values, and have more incentives to renegotiate debt instead of turning to asset sales.

The research question of optimal deleveraging during a crisis is only relevant, if crises are at least partially related to the solvency of institutions. I discuss to empirical contributions which confirm this view. Boyson et al. [2014] directly tests two competing hypotheses on the nature of crises namely whether the crisis of origin on individual bank level is *liquidity shortage*, or *insolvency*-driven, and the evidence provided supports rather the latter one. Consistently with this view, they find that banks asset selling choice reflects '*cherry picking*', namely they select asset for sale where the selling do not deplete the capital, and in effect, stabilize solvency position in the first place.

Finally, I close the empirical background review with a great insight by Adrian and Shin [2010]. This paper provides direct empirical evidence for the connection between deleveraging and fire sales. The authors show that leverage is procyclical in financial institutions, that is, bank's increase leverage during good times and -more importantly- deleverage during bad times, when prices fall. On aggregate the consequences could be massive fire-sale.

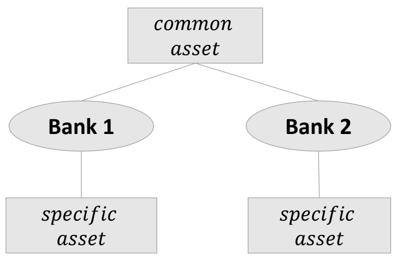
The rest of the paper is organized as follows: section 2 introduces the model and assumptions. Section 3 solves for equilibrium and for social optimum. Section 4 provides a numerical characterization of the results, while Section 5 concludes.

4.3 Model

4.3.1 Motivating example

To fix ideas, we start with a stylized, illustrative example. Consider an economy with two banks and two asset classes: a ‘pooled’, or ‘common’ asset (e.g. treasury bonds), which represents the asset commonality, and a ‘bank-specific’ asset class (e.g. loan portfolio), the idiosyncratic, uncorrelated assets of the two banks. The structure is illustrated in Figure 4.1. We assume that ex ante the common asset is perceived as more liquid, that is, in the ’business-as-usual’ regime it can be liquidated with a smaller price impact.

Figure 4.1: A stylized banking system.



We start by emphasizing the game-theoretic nature of the problem. The two banks simultaneously decide whether to sell their illiquid specific assets, or the more liquid common asset. By this we model a hypothetical ‘crisis’ situation when the whole banking system is engaged in a systematic de-leveraging, after the economy is hit by an (unmodelled) shock. Let the payoff matrix of the induced non-cooperative game be as follows:

	S	**C**
S	-3,-3	-5,-2
C	-2,-5	-4,-4

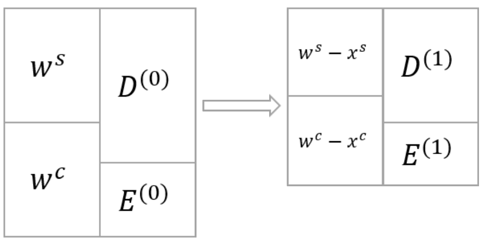
where ‘S’ and ‘C’ denote the action of liquidating the specific or common asset respectively. This illustrative payoff matrix is intuitively plausible in a crisis situation: selling the specific asset is costly (-3), but does not impose any negative externalities to the other player. Selling the (more liquid) common asset is less costly, as long as only one player chooses to do so, but in this case the other player suffers both from the high price impact of the specific asset (-3) and is hit by the negative price shock as a result of the other player’s liquidation of the common asset (-2). If, however, the players coordinate on selling the liquid asset, they both suffer from the externality and the liquid asset becomes ex post illiquid, as a result of the coordination problem. The payoff matrix describes a standard prisoners' dilemma situation: the only Nash-equilibrium of the (one-shot, not repeated) game is (C,C), and the equilibrium outcome is not Pareto-optimal.

From this illustrative discussion one would expect that to the extent that optimal decision of banks exhibit similar characteristics, there will be *too much liquidation* in equilibrium from the liquid asset, which is therefore not Pareto optimal. In the following we build foundations for the payoff-matrix arbitrarily imposed in this section to demonstrate that this is indeed the case.

4.3.2 The banking model

Consider the following environment: Banks are financed with a combination of debt and equity, and invest into two asset classes: a 'specific' (type s) and a 'common' (type c) asset.[3] Banks are subject to a leverage constraint: the ratio of total asset over equity must not exceed a pre-defined, bank-specific level $\overline{\ell}_i$. We do not interpret this necessarily as a regulatory requirement: it could be an internal target chosen by the bank's investment strategy. We consider a two-period model ($t \in \{0, 1\}$). In period 0 the leverage target is not satisfied (following for example an exogenous, unmodelled drop of asset prices). Banks respond by liquidating some of their assets, from which the proceeds are fully used to repay debt and thereby shrink their balance sheet, and importantly, decrease leverage, as documented in Adrian and Shin [2010] (see Figure 4.2). Asset prices in period 1 (denoted by $p^{(1)}$) are determined for each asset independently, by an inverse demand function which depends on the total liquidated asset. We start the analysis with a two-bank two-asset economy, which we refer to as the '2x2 model'.

Figure 4.2: Shrinking balance sheet
as a result of deleveraging



[3] An alternative interpretation is that common assets are investments into highly correlated portfolios, or for example in a market index.

In the model the banks are free to choose the liquidated quantities from each assets and rebalance the portfolio during the deleveraging process, which is consistent with empirical evidence discussed earlier (for example, Hau and Lai [2012], Boyson et al. [2014]). Banks' liquidation decisions are restricted only by being non-negative[4], and by a short-selling constraint.

We introduce the notation $p_i^{j(t)}$ for the price in period $t \in \{0, 1\}$, for asset $j \in \{s, c\}$ of bank $i \in \{1, 2\}$ where i may be omitted for the common asset. The total assets of bank i in period 0, with initial quantities w_i^s, w_i^c from specific and common asset respectively is therefore

$$A_i^0 := w_i^s p_i^{s(0)} + w_i^c p^{c(0)}$$

while in period 1, after liquidating quantities x_i^s, x_i^c it is

$$A_i^1 := (w_i^s - x_i^s) p_i^{s(1)} + (w_i^c - x_i^c) p^{c(1)}$$

All proceeds from liquidation are used to repay debt, so if D_i^0 denotes initial debt for bank i, the new debt level in period 1 is:

$$D_i^1 = D_i^0 - x_i^s p_i^{s(1)} - x_i^c p^{c(1)}$$

Equity (E) is expressed as total assets minus debt. Each equity-maximizer bank $i \in \{1, 2\}$ maximizes after-liquidation equity, subject to the leverage constraint, resource constraints and the no short-selling constraint. Formally, the problem of Bank i is:

$$\begin{aligned} \underset{x_i^c, x_i^s}{\text{maximize}} \quad & (w_i^s - x_i^s) p_i^{s(1)} + (w_i^c - x_i^c) p^{c(1)} - (D_i^0 - x_i^s p_i^{s(1)} - x_i^c p^{c(1)}) \\ \text{subject to} \quad & \frac{(w_i^s - x_i^s) p_i^{s(1)} + (w_i^c - x_i^c) p^{c(1)}}{(w_i^s - x_i^s) p_i^{s(1)} + (w_i^c - x_i^c) p^{c(1)} - (D_i^0 - x_i^s p_i^{s(1)} - x_i^c p^{c(1)})} \leq \bar{\ell}_i \\ & 0 \leq x_i^s \leq w_i^s \\ & 0 \leq x_i^c \leq w_i^c \end{aligned} \tag{4.1}$$

Intuitively, banks maximize after-sale equity which satisfies the leverage constraint under the new equilibrium asset prices. The objective function simplifies to $w_i^s p_i^{s(1)} + w_i^c p^{c(1)}$ (see Appendix), which also implies that banks' objective is equivalent to

[4] Relaxing this constraint by allowing negative liquidation (i.e. asset buy) would simplify the mathematical problem, but make the interpretation of results less straightforward, due to the possibility of positive price impacts and the emergence of 'bubbles' due to the extra demand.

minimizing (subject to constraints) the weighted price impact where the weights are the initial asset holdings. This is an economically important observation: the banks' loss, which must be written off against equity, is only the value loss due to decreasing prices; asset liquidation and debt repayment per se do not change equity value.

Our goal is to identify *equilibrium selling strategies* as a solution of each bank's optimization problem, which are linked by the joint determination of asset prices. This is formally the Nash-equilibrium of induced non-cooperative game.

Definition 1 (Liquidation equilibrium.) *A 'liquidation equilibrium' in the 2x2 model is a pair of selling vectors $x_1 \in \mathbb{R}^2, x_2 \in \mathbb{R}^2$ such that each vector is a solution of the maximum problem (4.1) $\forall i \in \{1, 2\}$, taking the other player's action x_{-i} as given. It is thereby the Nash-equilibrium of the induced game.*

In a liquidation equilibrium, none of the banks have incentive to deviate from the current liquidation strategy, taking the other player's decision as given.

Our objective is to analyse the effect of strategic behaviour and equilibrium decision of banks. Therefore, we restrict attention to environments where the equilibrium action is not constrained by the bank's original endowments, or any other parameters of the model in a trivial way. Before stating the assumptions on the parameter space, we introduce a convenient notation for the leverage constraint. Let $\mathbf{w}_i := (w_i^s, w_i^c)$, $\mathbf{x}_i := (x_i^s, x_i^c)$ and $\mathbf{p}_i^{(t)} := (p_i^{s(t)}, p^{c(t)})$, and a subscript $(-i)$ denote the player other than i. Then rearranging the leverage constraint leads to

$$\Lambda_i(\mathbf{x}_i, \mathbf{x}_{-i}) := (1 - \ell)\mathbf{w_i}'\mathbf{p}_i^{(1)} - \mathbf{x}_i'\mathbf{p}_i^{(1)} + \ell_i D_i \leq 0$$

Assumption 1 (Sufficient resources.) *Each bank i can restore leverage even if the other bank $(-i)$ liquidates all of its assets. Furthermore, for any liquidation choice of the other bank $(-i)$, each bank i can restore leverage either by selling only specific, or by selling only common assets. Formally[5]:*

$$\forall i : \Lambda_i((w_i^s, 0), \mathbf{x}_{-i}) < 0 \quad \forall \mathbf{0} \leq \mathbf{x}_{-i} \leq \mathbf{w}_{-i}$$

and

$$\forall i : \Lambda_i((0, w_i^c), \mathbf{x}_{-i}) < 0 \quad \forall \mathbf{0} \leq \mathbf{x}_{-i} \leq \mathbf{w}_{-i}$$

Assumption 1 ensures that the leverage constraint can be satisfied by selling either the common or the specific asset. It guarantees that the equilibrium action pro-

[5]We apply the convention that $\leq$ denotes element-wise relation when applied to vectors.

file is not restricted by the available resources in the economy. As a result, our analysis focuses on social inefficiencies which characterize banks' optimal equilibrium *behaviour*. The assumption as stated ensures that our results does not simply reflect unbalanced investments on the asset side of the bank's balance sheets.

Now we formalize the 'interesting problem' assumption. We will require that each bank is forced to liquidate a non-negative amount from at least one asset. This should be seen as a rather technical assumption ensuring that banks start from a situation in which the constraints are not satisfied, i.e. from a 'shocked' state - thereby it replaces the explicit modelling of an economic shock. Typically, this could be the result of a system-wide drop in asset prices.

Assumption 2 (Interesting problem.) *Without selling any assets, the leverage constraint is not satisfied.*

$$\forall i : \Lambda_i(\mathbf{0}, \mathbf{x}_{-i}) > 0 \quad \forall \mathbf{x}_{-i} \leq \mathbf{w}_{-i}$$

By assumption 2 we formally restrict the parameter space such that banks are forced to engage in strictly positive deleveraging, thereby make the formal problem interesting and consistent with our narrative. The next assumptions characterize the price impact function.

Assumption 3 (Linear price impact.) *The price impact is a linear function of the total quantities offered for sale. That is, for each assets j*

$$p_i^{j(1)} = p_i^{j(0)} + \xi_i^j \left(\sum_{k=1}^{2} x_k^j \right) \tag{4.2}$$

where $p_i^{j(0)}$ is the original price, ξ_i^j is a price impact coefficient, x_1^j and x_2^j are liquidated quantities, $p^{j(1)}$ is the updated (equilibrium) price of asset j for $j \in \{c, s\}$.

Assumption 4 (Small price impact) *We assume that the price impact is relatively small. Precisely, we assume the following expression to be positive:*

$$\forall i\, \forall j : p_i^{j(0)} + \xi^j \left(\sum_i w_i^j \right) + l_i w_i^j \xi^j > 0$$

Intuitively this assumption requires that even a hypothetical 'leveraged' short sale (which means more liquidation that is allowed by our constraints) would keep the prices in the positive range.

4.3.3 Model solution

We start solving the model with a detailed characterization of the mathematical problem:

Lemma 1 *The maximum problem 4.1*

(i) under assumptions 1, 2, and 3, maximizes a linear functional over a convex set. Thereby, the optimum must be at the boundary of the constraint set.

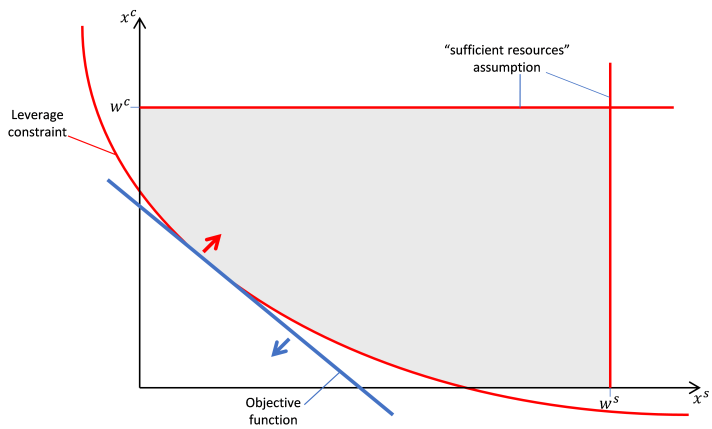
(ii) The constraint $\Lambda_i(\cdot)$ is monotonously decreasing in x_i under the (sufficient) assumption (4).

(iii) Under assumptions 1 and 2, in optimum, the leverage constraint is binding, and fulfils with equality. The 'no-short-selling' constraint $w^j > x^j$ is always non-binding.

Proof. *See Appendix 4.A.1* ■

The Lemma is illustrated in Figure 4.3. The formal proofs are in Appendix, here we discuss only the intuition. Item (i) follows from simple algebraic manipulations. Item (ii) is a key technical condition to ensure that asset liquidation and debt repayment indeed improves leverage. Item (iii) intuitively says that since the objective function by construction is decreasing in the quantity offered for sale, it is never optimal to liquidate more assets than the minimum quantity which just restores leverage. As a consequence, it will be binding in optimum.

Figure 4.3: Illustration of the geometry of the optimum problem



The fact that in optimum the leverage constraint is binding makes it possible to consider an equivalent problem, in which the value of required liquidation of specific assets is given as a function of the liquidated quantity of common assets. This formalization allows us to express the induced game in an equivalent form with one-dimensional strategy spaces. In this equivalent model bank's choose the quantity of common assets (the 'liquidation game'), and the required quantity of specific assets for liquidation is determined as the minimum quantity which solves the leverage constraint with equality for each banks. This, in turn, determines the payoff of a given strategy profile.

Lemma 2 *For each action profile* $\mathbf{x}^c \in [0, w_1^c] \times [0, w_2^c] \subset \mathbb{R}^2$ *there is a unique value of* x_1^s *and* x_2^s *which solves the model. It will be denoted by the function* $\tilde{x}^s(\mathbf{x}^c)$: $[0, w_1^c] \times [0, w_2^c] \to [0, w_1^s] \times [0, w_2^s]$, *and can be expressed in analytical form as:*

$$\tilde{x}^s = \frac{p^s - (1-\ell)w^s\xi^s - \sqrt{[p^s - (1-\ell)w^s\xi^s]^2 + 4\xi^s\left((1-\ell)\left(w^s p^{s(0)} + w^c p^{c(1)}\right) - x^c p^{c(1)} + \ell d\right)}}{-2\xi^s}$$

Proof. *See Appendix 4.A.3* ■

Lemma 2 states that x_i^s can be written as a function of the strategy profile $\mathbf{x}^c := (x_i^c, x_{-i}^c)$. Substituting back to the optimization problem, for each player i, *taking the action of* $(-i)$ *as given*, it becomes a one-dimensional, parametric, constrained maximization, equivalent with our original problem:

$$\begin{aligned} \underset{s_i^c}{\text{maximize}} \quad & w_i^s p_i^{s(1)}\left(\tilde{x}_i^s(\mathbf{x}^c)\right) + w_i^c p^{c(1)}(\mathbf{x}^c) & \\ \text{subject to} \quad & w_i^c \geq x_i^c \geq 0 & <\mu> \\ & w_i^s \geq \tilde{x}_i^s(\mathbf{x}^c) \geq 0 & <\lambda> \end{aligned} \tag{4.3}$$

The optimization problem of the reduced-form 2x2 model can be solved analytically using the Kuhn-Tucker conditions for optimality. The optimum is characterized with the following first-order conditions (i suppressed for simplicity):

$$\frac{\partial \mathcal{L}}{\partial x^c} = w^s\xi^s\frac{\partial x^s}{\partial x^c} + w^c\xi^c + \lambda\frac{\partial x^s}{\partial x^c} \leq 0$$

$$\lambda \geq 0; x^s \geq 0 \text{ but } \lambda(-x^s) = 0 \qquad <CS>$$

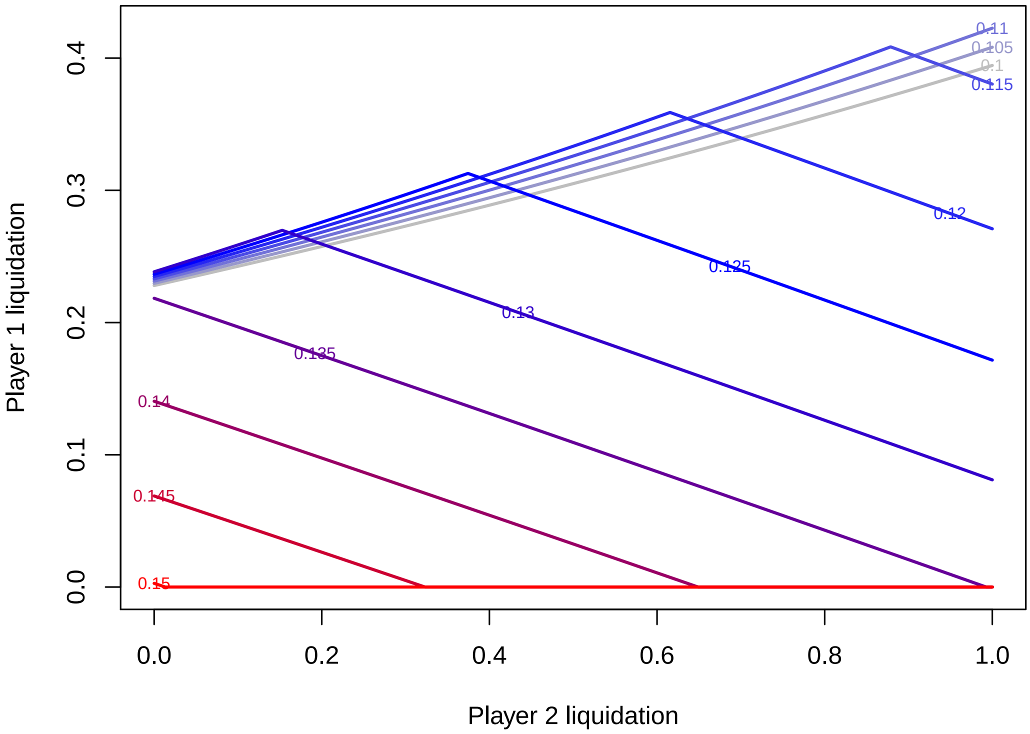
$$x^c \geq 0 \text{ but } x^c\left(\frac{\partial \mathcal{L}}{\partial x^c}\right) = 0 \qquad <CS>$$

We solve these conditions in the next Lemma. The optimum solution for each player i, viewed as a function of other player's $(-i)$ selling decision, is the best-response function of the liquidation game.

Lemma 3 (Best response function) *The best response function of the induced liquidation game, — a solution of problem 4.3 as a function of* s^c_{-i} *—, is a continuous, piecewise defined quadratic expression. The analytical form can be found in Appendix 4.A.4*

Proof. *See Appendix 4.A.4* ∎

Figure 4.4: Some best response functions



Best response functions for various price impact coefficients for the common asset (ξ^C*).*

Figure 3 provides an illustration of a range of best response functions for some values of the ex-ante perceived illiquidity of the common asset. For small price impact, only the common asset is liquidated, and therefore the best response function is increasing as the price shock by other player's action hits the player. For intermediate price impact ($0.11-0.13$ in the figure), the common asset is liquidated exclusively only as long as the counterparty liquidates relatively little: for large enough actions of player 2, player 1 gradually substitutes it by liquidating the specific asset. In the range $0.135-0.15$ both assets are liquidated and the best response function turns to decreasing. Clearly, in this range the player partially substitutes liquidation of common assets with idiosyncratic assets. Finally, for the largest price impacts on the diagram, Player 1 liquidates only the specific asset, and the best response function changes to flat at zero.

The illustrative best response functions illuminate a key aspect of the liquidation game. When player 2 increases the amount of liquidated common assets, there are two effects in force, which are reminiscent to standard microeconomic theory. First, there is an 'income effect', which is the straight network externality caused by common asset ownership: increased liquidation by some investors hurt other investors due to market price impact and mark-to-market evaluation, so they need to increase asset sale, which induces strategic complementarity. For sufficiently high ex-ante liquidity the optimum is to liquidate exclusively the common asset, and this is the only effect in place, therefore, the best response functions are increasing. However, in a multiple asset settings we can identify a second effect, which we can dub as 'substitution effect': as a result of increased liquidation by other market participants, investors may find it more attractive to increase the weight of specific assets in the liquidated portfolio and decrease liquidation from the common asset. The substitution effect may dominate the income effect, which leads to decreasing best response function and the appearance of strategic substitutes.

Our first theorem proves the existence of Nash-equilibrium of this game.

Theorem 1 *In the 2x2 model with linear price impact (assumption 3), for a given set of parameters $\Theta := \{W, p_0, D_0, \bar{\ell}, \xi\}$ fulfilling assumptions 1 and 2, there exist a liquidation equilibrium as a fixed point of the best-response correspondence.*

Proof. Consider the best-response correspondence

$$\Phi(s_1^c, s_2^c) : [0, w_1^c] \times [0, w_2^c] \to \mathbb{R}^2$$

Our assumptions 1 and 2 guarantee that

$$\operatorname{Im} \Phi(\cdot) \subseteq [0, w_1^c] \times [0, w_2^c]$$

That is, the image of the fixed point correspondence is a subset of the domain of $\Phi(\cdot)$. Lemma 3 establishes continuity of best response functions, so the mapping $\Phi(\cdot)$ is a continuous function. We have a continuous function which maps a compact convex set into itself, so the Brouwer fixed point theorem guarantees the existence of a fixed point. ■

4.3.4 The social planner's problem

The key insight from this paper is the inefficiency of the Nash-equilibrium of the liquidation game. To illustrate this point, we compare the equilibrium outcomes with a solution to a social planner's problem who maximizes joint equity of the two banks. The main result of this section demonstrates that a social planner, for a significant subset of the parameter space, could achieve a strictly better outcome and improve on the coordination failure in a market equilibrium. This is improvement in a a Pareto-sense, not just on aggregate equity, as in the social planner's optium solution all individual players will be better off.

We start with formalizing the social planner's problem which maximizes the joint equity subject to the leverage constraints, no-short-selling and nonnegativity assumptions. To write the problem compactly we extend the notation with

$$\mathbf{W} = \left(\begin{smallmatrix} w_1^s, 0, w_1^c \\ 0, w_2^s, w_2^c \end{smallmatrix}\right),\ \mathbf{X} = \left(\begin{smallmatrix} x_1^s, 0, x_1^c \\ 0, x_2^s, x_2^c \end{smallmatrix}\right),\ \mathbf{d} = \left(\begin{smallmatrix} D_1 \\ D_2 \end{smallmatrix}\right) \text{ and } \mathbf{p} = \left(\begin{smallmatrix} p_1^s, p_2^s, p^c \end{smallmatrix}\right)'$$

The social planner's problem can be written as

$$\begin{array}{lll} \underset{\mathbf{X}}{\text{maximize}} & \mathbf{1}' \left(\mathbf{W}\mathbf{p}^{(1)} - \mathbf{D} \right) & \\ \text{subject to} & (1-\ell_i)\mathbf{w_i}'\mathbf{p}^{(1)} - \mathbf{s_i}'\mathbf{p}^{(1)} + \ell_i D_i \leq 0 & \forall i \in \{1,2\} < \lambda_i > \\ & w_i^j \geq s_i^j \geq 0 & \forall i \in \{1,2\} \forall j \in \{s,c\} < \eta_i^j > \end{array}$$

where $\mathbf{1} = (1,1,1)'$. The objective function is the total after-liquidation equity of the financial system. The first set of constraints are the rearranged leverage constraints, which must hold for all institutions individually. The last set of constraints are the usual 'no-short-selling' and 'no-buy' constraints.

The solutions of the social planner's problem are characterized with the usual first order conditions (see Appendix 4.A.5). The main result of this section compares the socially optimal solution with the liquidation equilibrium in the previous subsection:

Theorem 2 *The liquidation equilibrium is not Pareto-optimal in all such cases in which in equilibrium a positive quantity is chosen from both assets for liquidation.*
Proof. *See Appendix 4.A.5* ∎

Intuitively, we are in a prisoners'-dilemma like situation as demonstrated in the introduction: the Pareto-optimal liquidation strategy is not individually optimal, therefore cannot be maintained as equilibrium, since banks have incentives to deviate and sell more from the relatively liquid common asset. However, if both are doing this, the asset becomes relatively illiquid, and the total payoff is lower.

Theorem 2 only characterizes the situation in which a positive quantity is chosen from both assets. The following theorem establishes the optimality of corner-solutions as well:

Theorem 3 *The 'liquidation equilibrium' coincides with the social planner's solution in all such cases when*

i zero quantity is chosen from the common assets for liquidation in equilibrium.

ii zero quantity is chosen from the specific asset by the social planner

Proof. *See Appendix 4.A.6* ■

4.4 Model analysis

In this chapter we analyse and illustrate the equilibrium solution and compare with social planners' outcome. The derivation of our main result has no restriction whatsoever on the parameter values of the economy apart from Assumptions 1...4. To make the comparison more transparent, without loss of generality, we can introduce a few normalizations on the parameter space.

First, note that absolute price levels do not play a role in this context, only relative price *changes* determine the result. Therefore, we can normalize all initial prices to one:

$$\mathbf{p}^{(0)} := \mathbf{1}$$

We cannot, however, normalize all initial quantities, since that would pin down the investment proportions of idiosyncratic versus common asset. Without loss of generality, however, we can normalize idiosyncratic asset quantities to one, and leave the total quantities of common assets as a variable. In this section we focus on symmetric equilibrium with ex-ante identical banks. [6] We set $\mathbf{w}_c = w_c$ as a parameter, and normalize

$$\mathbf{w}_s = \mathbf{1}$$

In the main model, the debt amount D_0 was also treated as a free parameter. It is however easier to interpret and compare the outcomes for different parameter settings if instead of the amount of initial debt, the initial leverage (following the initial shock) is parameterized. In particular, we define (without loss of generality)

[6] Otherwise another parameter would be required which determines the allocation of common asset between the two banks.

a new variable, 'shock size', defined as $\kappa := \frac{\ell^0}{\bar{\ell}}$. With this definition, the initial debt level is:

$$D_0 = A - \frac{A}{\kappa \bar{\ell}}$$

For example, if we normalize all $p_i^{j(0)} = 1$, $w_i^j = 1$ and $\bar{l} = 2$, a '*10% shock*' corresponds to a shock-size of $\kappa = 1.1$ which implies an initial debt of 1.1, and an initial leverage of $\ell^0 = 2.22$. This should be restored to $\ell = 2$, a *10% decrease* in leverage.

Finally, we can slightly modify the definition of the price impact parameter (ξ). So far ξ^s and ξ^p could be arbitrary numbers. For the purpose of numerical analysis, it is better to tie their values to total asset quantities. Intuitively one would expect that a given (numerical) quantity for sale must have smaller price effect, if it is a relatively small portion of the *total* asset. In order to incorporate this to the model without any affect to the validity of proofs, we introduce a new variable

$$\phi^j = \frac{\xi^j}{W^j}$$

with W^j being the total quantity of asset j, and in all numerical illustrations we apply the price impact formula with ϕ in place of ξ.

$$\mathbf{p}^{(1)} = \mathbf{p}^{(0)} + \phi \mathbf{X}^j$$

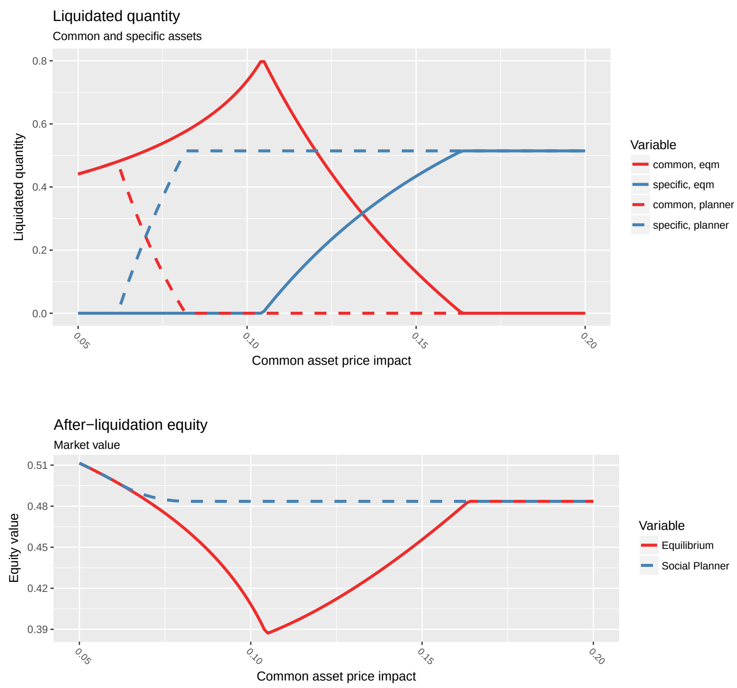
To sum up, after introducing the normalizations the economy can be described by the following parameters:

$$\Theta^N := \{\boldsymbol{\xi}, W^c, \kappa, \bar{\boldsymbol{l}}\}$$

The most critical assumption of this model is the 'sufficient resources' assumption, which allows us to ignore cases where a selling decision is made due to inefficient asset holdings, not by optimal choice. Intuitively, this requirement can be interpreted as a constraint on the price impact coefficients which we can meaningfully consider in the numerical analysis. The larger the leverage (left) or the larger the total asset commonality in the economy (right), the smaller price impact coefficients is sufficient to satisfy the assumptions.

In the following sub-sections we illustrate the effect of various parameters of the economy on the equilibrium liquidation quantity (thick blue and red) and the social optima (dashed blue and red) curves (top figures). Further, we calculate the value of equity in equilibrium (blue) versus social optimum (red) equity levels, expressed as a percentage of the original equity (bottom figures).

Figure 4.5: Equilibrium as a function of common asset price impact



4.4.1 Common asset price impact

We start by varying the common asset price impact coefficient (ξ^c), while keeping other parameters, notably the price impact for the specific asset, fixed. As discussed before, the portfolio is symmetric and normalized ($w_1^s = w_2^s = 1, w_1^c = w_2^c = 1$) and initial prices are normalized ($p^c = p^s = 1$). By the symmetry of the problem, the figures can be understood as the outcome for both $i = 1, 2$. The leverage constraint is set to $\bar{\ell} = 3$ and the banking sector starts from a 'shocked' state where the leverage constraint is not fulfilled by a factor of $\kappa = 1.1$ which we can interpret as the size of shock. Therefore, $\ell^0 = \kappa\bar{\ell}$.

Figure [4.5] presents liquidated quantities (top figure) for equilibrium (solid) and socially optimal (dashed) solution, and after-liquidation equity for optimum (red) and equilibrium (blue) as the specific asset's price impact is fixed at $\xi^s = 0.15$ and the common price impact varies in the range $\xi^c \in \{0.05, 0.2\}$.

For very high ex-ante common asset price impact (the right of axis x), the equilibrium action is to liquidate only specific asset. In this region, obviously, any further increase of the common price impact coefficient ξ^c has no further effect on the optimal choice or the equity value. Note however, that this region does not comply with the 'narrative' of the paper, that is, that usually the common asset is perceived as ex ante more liquid. As the price impact decreases to the interesting region (towards the left of x axis), the equilibrium action is to sell a combination of the two assets (there is an internal solution to the optimum problem). In this case the negative externality pushes down the new equity value below previous levels.

As the price impact coefficient of common asset further decreases, it becomes more-and-more attractive to substitute for it, which exacerbates the coordination problem. As a result, counter-intuitively, banks' fire-sale losses increase as the (absolute) price impact decreases, that is, more liquid markets eventually hurt the banks. Intuitively, assets which are perceived as more liquid attracts more sellers, therefore they become more illiquid ex post. That means formally, the derivative of the objective function at optimum with respect to the price impact function of common asset is negative.[7] This is a key result, so we state this as a theorem[8]:

Theorem 4 *Let*

$$\Phi(\mathbf{x}) = (w_i^s - x_i^s)p_i^{s(1)} + (w_i^c - x_i^c)p^{c(1)} - (D_i - x_i^s p_i^{s(1)} - x_i^c p^{c(1)})$$

denote the objective function of maximum problem (4.1)*. Then, under assumption 4 we have*

$$\frac{\partial \Phi^\star}{\partial \xi^c} < 0$$

for the interior optimum region, where $\Phi^\star = \Phi^\star(x^\star(\xi^\star))$ *is the optimum value of the objective function viewed as a function of equilibrium selling quantities.*
Proof. *See Appendix 4.A.7.* ∎

In contrast to the equilibrium outcome, a social planner could avoid the increased fire-sale losses by selling the specific asset for both banks for the critical range of price impact coefficients.

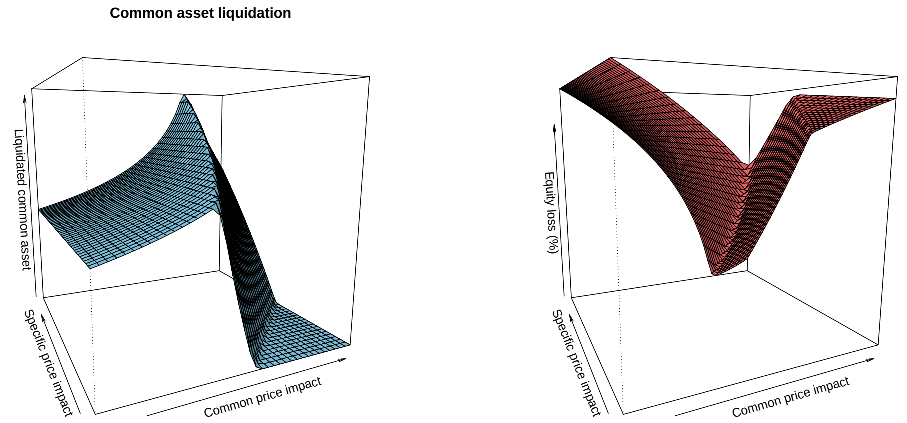
Finally, the left-side of the diagram illustrates that for sufficiently low levels of common asset price impacts the optimum is the same as the equilibrium, as predicted by Theorem 3. If the common asset is sufficiently liquid, the benefits from liquidity outweigh the costs associated with network externalities.

So far the the price impact for the specific asset was fixed at an arbitrary level. Figure 4.6 illustrates the role of overall liquidity profiles (combinations of price impact coefficients). The left figure is the equilibrium of the liquidation game

[7] This is not obvious, since $\frac{\partial \Phi}{\partial \xi^c} > 0$.

[8] In the appendix, I formulate the statement of the Theorem and give an analytical condition which guarantees the Theorem to hold. Unfortunately, this condition cannot be expressed in a closed analytical form, so instead of a full analytical proof, I provide an illustration that the condition indeed holds in the interesting parameter region.

Figure 4.6: Equilibrium as a function of price impacts



(a) Equilibrium of the (restricted) liquidation game

(b) Ex-post equity values in liquidation equilibrium

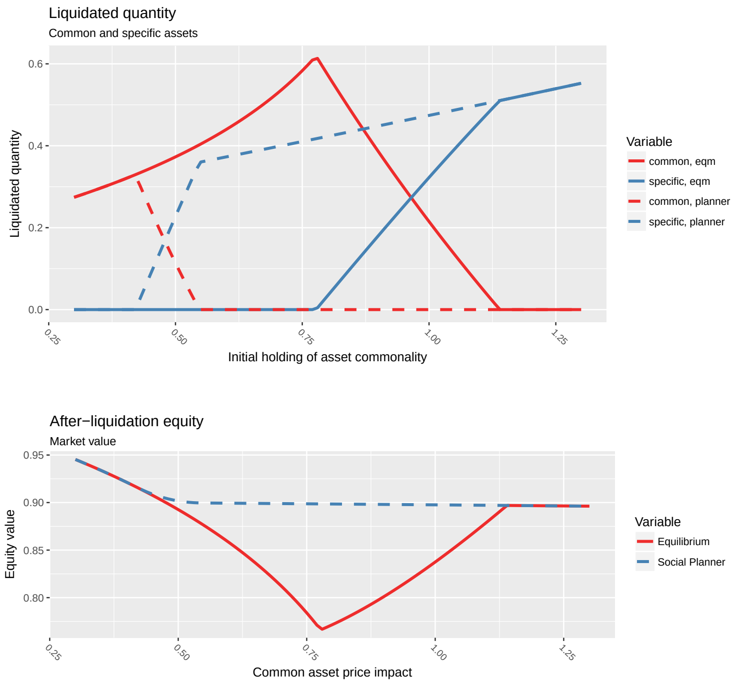
(common assets offered for sale), while the right figure demonstrates the associated equity loss. It is not surprising that higher overall illiquidity leads to larger equity losses. The notable result is that increasing illiquidity of specific assets largely extends the problematic region in both directions, but more significantly, for larger (absolute) common price impacts. The conclusion is that an overall deterioration of market liquidity involving all assets in the market, makes it more likely (for a larger subset of parameters) that banks end up in an inefficient liquidation spiral, an effect which is beyond the simple consequences of deteriorating liquidity.

4.4.2 Weight of the common asset

In this chapter we vary the total quantity of common asset in the market keeping the weight of specific assets normalized at $w_i^s = 1$, and all other parameters constant. The total quantity of common asset W^c can be interpreted as a simple measure of diversification on the market: since w_i^s is normalized, the larger W^c is, the higher percentage of wealth is invested in the (perfectly correlated) common asset, so the larger is the correlation between the two banks' total asset portfolio. To interpret the results we note that by the normalization we adopted for numerical analyses, varying quantity of common asset on the balance sheet while keeping prices and all other parameters fixed also changes the absolute size of the balance sheet.

The diagram is topologically similar to varying the price impact parameter.

Figure 4.7: Equilibrium as a function of asset commonality



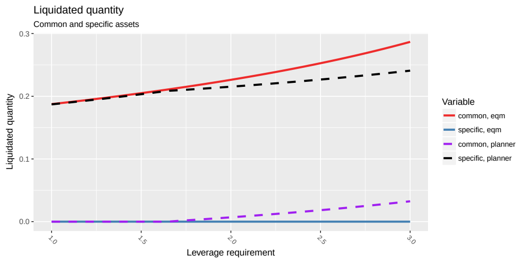
The reason is fundamental: as it is obvious from the form of the objective function, bank's are concerned about *weighted* price impacts, so increasing weight of common asset has a similar effect as increasing price impact of common assets. The interpretation is, however different, and is analogous to earlier findings in the literature (e.g. Elliott et al. [2014]), which points out that financial networks are most susceptible for contagion for intermediate levels of diversification.

Although this model is based on completely different principles than earlier literature, the intuition is the same: for high levels of diversification, banks internalize a relatively high proportion of network externalities, and optimally choose to sell the specific asset, and reach the optimal outcome. For relatively low levels of diversification, the network externality is small, and does not dominates the benefits from better liquidity of the common asset. The banking system is subject to a possibly inefficient equilibrium outcome for *intermediate levels* of diversification.

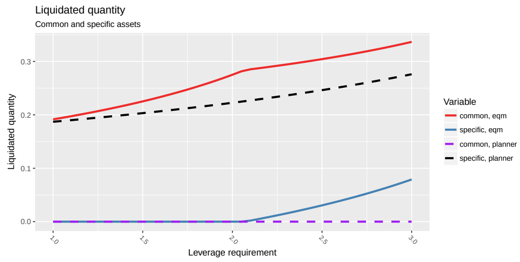
4.4.3 Varying leverage

Next we analyse the effect of varying leverage levels of banks (see figure 4.8). We start the graph from the limiting case $\ell = 1$. By definition in this case $d_0 = 0$

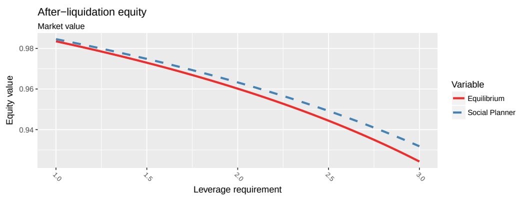
Figure 4.8: Equilibrium as a function of leverage



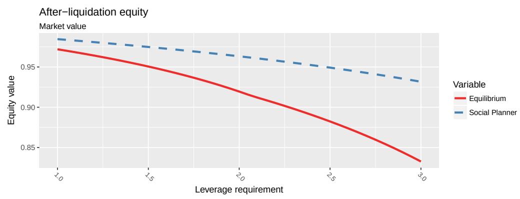
(a) Low-ξ^c regime



(b) High-ξ^c regime



(c) Low-ξ^c regime



(d) High-ξ^c regime

so there is no liquidation, and no equity loss.[9] Note how Theorem 2 holds for the depicted case: optimal liquidation of common asset are always below the equilibrium levels. The relative equity levels (bottom figure) show that the inefficiency sharply increases for larger levels of leverage. Note that we kept the relative shock size constant in the numerical analysis, and not the absolute shock size. I repeated the analysis with fixed absolute shock size, and the results are similar (although the absolute magnitude is obviously smaller).

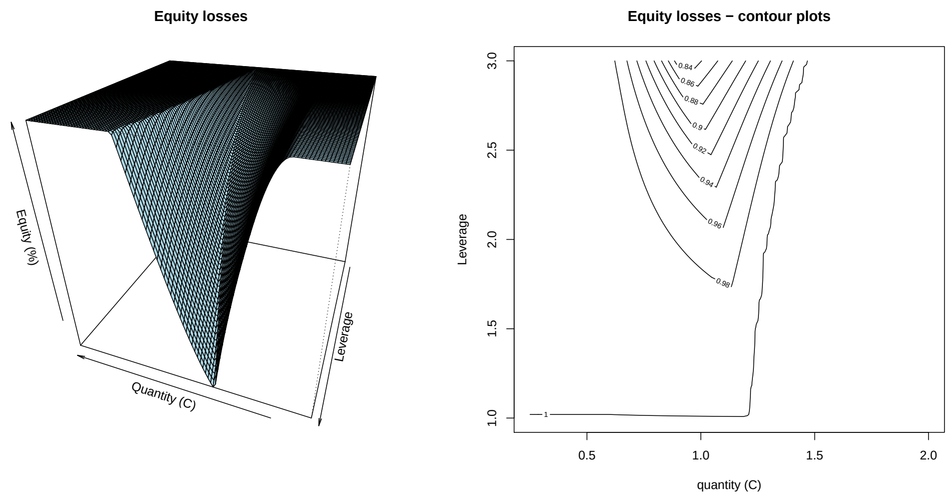
In figure 4.9 we illustrate the effect of leverage and market diversification (total asset commonality) on the relative equity loss of prevailing equilibrium compared to the social optima. It is not surprising based on the previous figure that larger leverage may have significantly larger negative effect, and this diagram also makes clear that the potential parameter region susceptible to inefficient equilibrium is much larger as well. In addition, if leverage is higher, the 'worst outcome' tend to occur for lower levels of market diversification.

4.4.4 Varying shock size

The shock in this context should be understood as how far the initial position is from fulfilling the leverage requirement. The effect on equity is straightforward: higher shocks induces more liquidation. The composition of optimal selling vector in figure 4.10 demonstrates that it is not scale-free with respect to the shock size.

[9]Formally our assumptions are not fulfilled at this point, so we consider it as a limiting case.

Figure 4.9: Equilibrium as a function of leverage and commonality



(a) Relative equity losses

(b) Contour plots

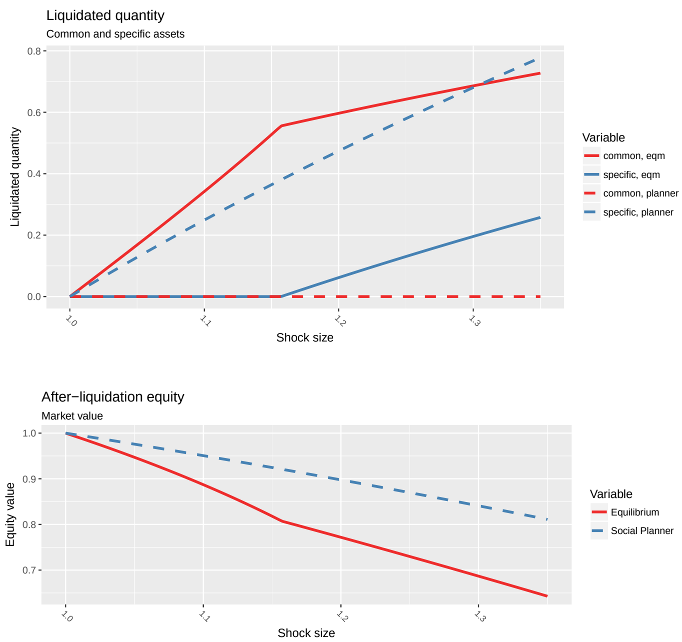
For small shock, it is optimal to sell only the specific asset, while for larger shocks, a substitution towards common asset starts.

4.5 Conclusions

In this paper we analysed a simple banking system with fire-sale spillovers due to common illiquid asset holdings when banks have multiple illiquid assets and choose the liquidation strategy optimally. Even the very simple, two-bank model can produce interesting and counter-intuitive results. Our main theorem states that for a large range of parameter values the only prevailing equilibrium of the 'liquidation game' is in which banks sell too much of the common asset, compared to what a social planner would find optimal. The most striking result is that in these situations, at least under symmetric equilibrium, the socially optimal solution is a Pareto-improvement as well, that is, banks individually would get better off by choosing the liquidation strategy designed by the social planner, but this cannot be maintained as a Nash-equilibrium. Even more counter-intuitively, the efficiency losses due to coordinating on the inefficient equilibrium are larger when the market conditions improve: if the common assets are perceived more liquid, the temptation to coordinate on selling it is higher, and the ex-post illiquidity and induced equity losses are higher.

Our illustrative numerical analysis indicate that even this simple model can

Figure 4.10: Equilibrium as a function of shock size



suggest two findings which are analogous to earlier theoretical results in the financial contagion literature. First, the equity losses are most severe for intermediate levels of market diversification; in contrast, for very large diversification the fire sale externalities are sufficiently internalized, while for very low levels of diversification the externalities are sufficiently small and do not influence outcomes. Note that our setup intentionally excludes the 'extremes' (like a perfectly diversified banking sector), in which pure resource constraints would alter our conclusions. By excluding those parameter regions by assumption, our results are purely due to bank's strategic considerations. The second result is a weak confirmation of the findings of Chen et al. [2014]: as the system becomes more leveraged, it can benefit more from less diversification and more idiosyncratic portfolios.

4.A Appendix - Proofs

4.A.1 Notation

Consider the optimum problem 4.1. After straightforward algebraic simplifications, and switching to matrix notation the problem of bank i can be written as:

$$\begin{array}{lll} \underset{\mathbf{x_i}}{\text{maximize}} & \mathbf{w}_i'\mathbf{p}_i^{(1)}(\mathbf{x}_i) - D_i & \\ \text{subject to} & (1-\ell_i)\mathbf{w}_i'\mathbf{p}_i^{(1)} - \mathbf{x}_i'\mathbf{p}_i^{(1)} + \ell_i d_i \leq 0 & <\lambda> \\ & \mathbf{x}_i \geq 0 & <\eta> \\ & \mathbf{w}_i - \mathbf{x}_i \geq 0 & <\mu> \end{array} \tag{4.4}$$

where the t=1 price $\mathbf{p}_i^{(1)}$ is determined by the linear price impact (Assumption 3)

$$\mathbf{p}_i^{(1)}(\mathbf{x}_i + \mathbf{x}_{-i}) = \mathbf{p}_i^{(0)} + \langle \boldsymbol{\xi_i} \rangle (\mathbf{x}_i + \mathbf{x}_{-i})$$

The vectors $\{\mathbf{w}_i, \mathbf{x}_i\}$ denote Bank i's asset endowment and liquidation decision, $\mathbf{x}_{-i}$ is the other bank's (total) liquidation from the same assets, $\mathbf{p}_i$ denotes prices of these assets, while $\langle \boldsymbol{\xi} \rangle$ is a diagonal matrix constructed from the price impact vector $\boldsymbol{\xi}$. To simplify discussion we introduce some further notation. Let $\Lambda_i(\mathbf{x}_i) : \mathbb{R}^2 \to \mathbb{R}$ denote the ex-post (after liquidation) leverage constraint (RHS of the constraint associated with λ) in the optimization problem of Bank i given that a selling vector $\mathbf{x}_i$ is chosen, that is,

$$\Lambda_i(\mathbf{x}_i) := (1-\ell_i)\mathbf{w_i}'\mathbf{p}_i^{(1)}(\mathbf{x}_i) - \mathbf{x_i}'\mathbf{p}_i^{(1)}(\mathbf{x}_i) + \ell_i D_i$$

Furthermore, let the objective be

$$\Phi_i(\mathbf{x}_i) := \mathbf{w}_i'\mathbf{p}_i^1(\mathbf{x}_i) - d_i$$

Note that $p_i^{(1)} = p_i^{(1)}(x_i, x_{-i})$ is a function of the liquidated asset quantities in both expressions. With this notation we can describe the decision problem of each banks i in a compact form, as a standard constrained maximum problem:

$$\begin{array}{lll} \underset{\mathbf{x_i}}{\text{maximize}} & \Phi_i(x_i) & \\ \text{subject to} & -\Lambda_i(x_i) \geq 0 & <\lambda> \\ & \mathbf{x}_i \geq 0 & <\eta> \\ & \mathbf{w}_i - \mathbf{x}_i \geq 0 & <\mu> \end{array} \tag{4.5}$$

These optimum problems are linked by the joint determination of asset prices

$$\mathbf{p}_i^{(1)} = \mathbf{p}_i^{(0)} + \langle \boldsymbol{\xi} \rangle (\mathbf{x_i} + \mathbf{x_{-i}})$$

For notational simplicity we suppress the dependency of value function and constraint on all other parameters describing the economy until the chapter dealing with comparative statics. For completeness, these parameters are

$$\Theta := \{\boldsymbol{\xi}, \mathbf{p}^{(0)}, \mathbf{W}, \mathbf{d}, \overline{\boldsymbol{\ell}}\}$$

Both the objective and constraint, therefore the optimum depends on the liquidation choice of the players, x_{-i}. If we want to emphasize this relationship, we write $\Phi(x_i, x_{-i})$ and $\Lambda(x_i, x_{-i})$ respectively. The object of interest is the set of maximizers, interpreted as a function of other bank's decision, denoted by $x_i^\star(x_{-i})$. This is the best response function of the game.

4.A.2 Proof of Lemma 1

Proof. (item i)

The objective function $\Phi_i(\mathbf{x}_i)$ is obviously a linear functional. Substituting the price equation into the constraint we get

$$\Lambda_i(x_i) = (1-\ell_i)\mathbf{w}_i' \left(\mathbf{p}_i^{(0)} + \langle \boldsymbol{\xi} \rangle (\mathbf{x}_i + \mathbf{x}_{-i})\right) - \mathbf{x}_i' \left(\mathbf{p}_i^{(0)} + \boldsymbol{\xi} (\mathbf{x}_i + \mathbf{x}_{-i})\right) + \ell_i d_i \quad (4.6)$$

The right-hand-side is a 2-variable continuous, twice differentiable function. We compute the Hessian with respect to the two variables $x_i^{(j)} : j \in \{s, c\}$:

$$\begin{aligned}
&\frac{\partial \Lambda_i(\mathbf{x}_i)}{\partial x_i^j} = (1-l) w_i^j \xi_i^j - \left(p_i^{(0),j} + \xi_i^j \left(x_i^j + x_{-i}^j\right)\right) - s_i^j \xi_i^j \\
&\frac{\partial^2 C}{\partial x^{(j)2}} = -2\xi_i^j > 0 \\
&\frac{\partial^2 C}{\partial x^{(j1)} \partial s^{(j2)}} = 0
\end{aligned}$$

By definition the price impact is negative ($\xi < 0$), so the Hessian is positive semi-definite[10]. This proves that $\Lambda_i(\mathbf{x}_i)$ is convex. The epigraph of a convex function is convex by definition of convexity, so the (leverage) constraint set is convex in $\mathbb{R}^2$. ■

Proof. (item iia)

We state conditions under which $\Lambda(\mathbf{x})$ is decreasing in x, that is, selling assets

[10]The geometric figure is an infinite paraboloid, the epigraph is ellipsoid

and repaying debt is leverage-decreasing. A sufficient condition is $\forall j \in \{s, c\}$:

$$\frac{\partial \Lambda_i(\mathbf{x}_i)}{\partial x^j} = (1-l)w_i^j \xi^j - (p_i^{(0),j} + \xi^j(x_i^j + x_{-i}^j)) - x_i^j \xi^j < 0$$

This holds whenever the price impact is sufficiently small, that is:

$$p_i^{(0),j} + 2\xi^j x_i^j + \xi^j x_{-i}^j - (1-l)w_i^j \xi^j > 0$$

We replace x with $w > x$ to get a sufficient restriction on exogenous the parameter space instead of restricting the endogenous action space:

$$p^j + 2\xi^j w_i^j + \xi^j w_{-i}^j - (1-l)w_i^j \xi^j > 0$$

which is stated in assumption 4. ■

Proof. (item iib)

The sufficient resources assumptions make the constraint associated with μ in the general set-up obsolete. The choices $(w_i^s, 0)$ and $(0, w_i^c)$ is clearly in the attainable set of the optimization problem, by Assumption 1. Take the first case (the second is proven analogously). Assume the optimum is some $(x_i^{s\star} > w_i^s, x_i^{c\star} > 0)$. This, however, cannot be optimum, because $\Phi(\mathbf{x}_i)$ is strictly decreasing in $\mathbf{x}_i$. Contradiction.

Furthermore, Assumption 1 and the continuity and monotonicity of Λ implies the existence of unique values $\overline{x}_i^c$ and $\overline{x}_i^s$ which exactly solves the leverage constraint.

$$\Lambda(\overline{x}_i^c, 0) = 0 \text{ and } \Lambda(0, \overline{x}_i^s) = 0$$

Clearly any choice $x_i^j > \overline{x}^j$ cannot be optimal. ■

Proof. (item iii) Assumption 1 and 2 together guarantees that

$$\mathbb{R}_-^2 \cap \{\mathbf{x}_i : \Lambda(\mathbf{x}_i) \leq 0\} = \varnothing$$

■

To sum up, we formalized the assumptions in such a way which ensures that (i) the problem can be solved if and only if banks liquidate a positive amount of assets, (ii) optimum is always on the boundary of the constraint set (iii) this boundary is always given by the leverage constraint, which is fulfilled with equality.

4.A.3 Proof of Lemma 2

Focus on bank $i = 1$ decision problem, taking bank $i = 2$ decision as given. Since the leverage constraint fulfils with equality (lemma 1), any choice of x_1^p clearly defines the required sale from x^s by the equation

$$\Lambda(x_1^c, x_1^s, \cdot) = 0$$

Denote this quantity for each bank i by $\tilde{x}_i^s(x_i^c) : [0, w^{(c)}] \to [0, w^{(c)}]$

Assumption (1) and consequently 1 ensures that $\tilde{s}_i^s(\mathbf{s}^c)$ can be derived analytically by solving the constraint as equality for s^s as a function of s^p and s_{-i}^p:

$$
\begin{aligned}
0 &= (1-\ell_i)\mathbf{w_i}'\mathbf{p}^1 - \mathbf{x}_i'\mathbf{p}^1 + \ell d_i \\
0 &= (1-\ell)w^s\left(p^s + \xi^s x^s\right) + (1-\ell_i)w^c\left(p^c + \xi^c x^c + \xi^c x_{-i}^c\right) \\
&\qquad - x^s(p^s + \xi^s x^s) - x^c(p^c + \xi^c x^c + \xi^c x_{-i}^c) + \ell d \\
0 &= [-\xi^s](x^s)^2 + [(1-\ell)w^s\xi^s - p^s]\,x^s + (1-\ell)w^s p^s \\
&\qquad + (1-\ell)w^c\left(p^c + \xi^c x_i^c + \xi^c x_{-i}^c\right) - x^c\left(p^c + \xi^c x^c + \xi^c x_{-i}^c\right) + \ell d
\end{aligned}
$$

This is

$$\tilde{x}^s = \frac{p^s - (1-\ell)w^s\xi^s - \sqrt{[p^s - (1-\ell)w^s\xi^s]^2 + 4\xi^s\left((1-\ell)\left(\boldsymbol{w'p} + w^c\xi^c(x_i^c + x_{-i}^c)\right) - x^c\left(p^c + \xi^c\left(x^c + x_{-i}^c\right)\right) + \ell d\right)}}{-2\xi^s}$$

The geometry of the problem implies that we only need to consider the (-) sign.

Note that the expression for $\tilde{s}_i^s(\mathbf{s}^c)$ is not guaranteed to be positive for every possible action profile s^c. It is just the quantity which solves the constraint with equality. Negative values will be ruled out as possible solutions as part of the solution of the 'reduced' optimum problem discussed as part of the next lemma.

4.A.4 Proof of Lemma 3

From the previous Lemma x_i^s can be written $\forall i \in \{1, 2\}$ as a function of the strategy profile $\mathbf{x}^c := \{x_i^c, x_{-i}^c\}$. Let this function be $\tilde{x}_i^s(\mathbf{x}^c) : \mathbb{R}^2 \to \mathbb{R}$. Substituting back to the optimization problem, for each player i taking the action by $-i$ as given, it becomes a one-dimensional, parametric constrained maximization program, which is equivalent with our original problem. For all i,

$$\begin{aligned}
\underset{x_i^c}{\text{maximize}} \quad & w_i^s p^{1s}(\tilde{x}_i^s(\mathbf{x}^c)) + w_i^c p^{1c}(\mathbf{x}^c) \\
\text{subject to} \quad & 0 \leq x_i^c \leq w_i^c \qquad\qquad\qquad\qquad < \mu > \\
& 0 \leq \tilde{x}_i^s(\mathbf{x}^c) \leq w_i^s \qquad\qquad\quad\;\; < \lambda >
\end{aligned} \tag{4.7}$$

This leads to the following KKT first-order conditions (i and function arguments suppressed for simplicity):

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial s^c} = w^s \xi^s \frac{\partial \tilde{x}^s}{\partial x^c} + w^p \xi^p + \lambda \frac{\partial x^s}{\partial x^c} &\leq 0 \\
\lambda \geq 0; x^s \geq 0 \text{ but } \lambda(-x^s) &= 0 \qquad\qquad < CS > \\
x^c \geq 0 \text{ but } x^c \left(\frac{\partial \mathcal{L}}{\partial x^c} \right) &= 0 \qquad\qquad < CS >
\end{aligned}$$

The usual 'generic' approach to analytically solve the Karush-Kuhn-Tucker conditions is to evaluate all possible combinations with respect to the positivity of Lagrange-multipliers, then check whether the results are consistent with the imposed conditions. Finally, eliminate those which are not consistent, and evaluate the objective function over the set of candidate solutions. Following this method, we have to consider four cases:

	λ	η	Intuition
Case 1	$\lambda = 0$	$\eta = 0$	Non-binding constraints: $x^s > 0$ and $x^c > 0$
Case 2	$\lambda > 0$	$\eta = 0$	Binding constraint on $x^s \Rightarrow x^s = 0$
Case 3	$\lambda = 0$	$\eta > 0$	Binding constraint on $x^p \Rightarrow x^p = 0$
Case 4	$\lambda > 0$	$\eta > 0$	Both constraints binding: *ruled out by assumption*

Note that Case 4 is ruled out by the 'interesting problem' assumption. Below we solve case 1 to 3 in turn.

Case 1 If $\lambda = 0$ the first-order condition gives the interior optimum:

$$\frac{\partial \mathcal{L}}{\partial x^c} = 0$$

from which we obtain the solution

$$\frac{\partial x^s}{\partial x^c} = -\frac{w^c \xi^c}{w^s \xi^s} \tag{4.8}$$

This can be solved analytically. First, we calculate the partial derivative

$$\frac{\partial x^s}{\partial x^c} = -\frac{(1-l)w^c\xi^c - \left(p^c + \xi x^c_{-i}\right) - 2\xi^c x^c}{\sqrt{[(1-l)w^s\xi^s - p^s]^2 + 4\xi^s\left[(1-l)w^s p^s + (1-l)w^c\left(p^c + \xi^c\left(x^c + x^c_{-i}\right)\right) - x^c\left(p^c + \xi^c\left(x^c + x^c_{-i}\right)\right) + ld\right]}}$$

After rearranging equation (4.8), we obtain

$$\left[\frac{w^s\xi^s}{w^c\xi^c}\right]^2\left([(1-l)w^c\xi^c - (p^c + \xi x^c_{-i})]^2 - 4\,[(1-l)w^c\xi^c - (p^c + \xi x^c_{-i})]\,\xi^c x^c + 4\,(\xi^c x^c)^2\right) =$$
$$= [(1-l)w^s\xi^s - p^s]^2 + 4\xi^s\,[(1-l)w^s p^s + (1-l)w^c\,(p^c + \xi^c\,(x^c + x^c_{-i})) - x^c\,(p^c + \xi^c\,(x^c + x^c_{-i})) + ld]$$

We express the solution for x^c implicitly as

$$A[x^c]^2 + B[x^c] + C = 0 \tag{4.9}$$

where after solving the previous equation,

$$\begin{aligned}
A &= 4(\xi^c)^2\left[\frac{w^s\xi^s}{w^c\xi^c}\right]^2 + 4\xi^s\xi^c \\
B &= -4\xi^c\left[(1-l)w^c\xi^c - \left(p^c + \xi x^c_{-i}\right)\right]\left[\frac{w^s\xi^s}{w^c\xi^c}\right]^2 - 4\xi^s\left((1-l)w^c\xi^c - p^c - \xi^c x^c_{-i}\right) \\
C &= \left[(1-l)w^c\xi^c - \left(p^c + \xi x^c_{-i}\right)\right]^2\left[\frac{w^s\xi^s}{w^c\xi^c}\right]^2 - [(1-l)w^s\xi^s - p^s]^2 \\
&\qquad - 4\xi^s\left[(1-l)w^s p^s + (1-l)w^c(p^c + \xi^c x^c_{-i}) + ld\right]
\end{aligned}$$

The value of x^c in 4.9 is the optimum solution as long as $\tilde{x}^s(x^{c\star}) > 0$

Case 2 The first complementarity slackness condition implies with $\lambda \neq 0$ (that is, if the constraint is binding and $s^s = 0$) that s^p is a similar quadratic formula with the coefficients.[11]

$$\begin{aligned}
A &= -\xi^c \\
B &= [(1-l)w^c\xi^c - (p^c + \xi^c x^c_{-i}) \\
C &= [(1-l)w^s p^s + (1-l)w^c(p^c + \xi^c x^c_{-i}) + ld
\end{aligned}$$

Corollary: The best response function is continuous.

Proof. Both x^c_{case1} and x^c_{case2} are piecewise continuous. Furthermore, $\tilde{x}^s(x^c)$ is also continuous and monotone by assumption 4. We need to prove that there

[11] The easiest way to calculate these coefficients is to solve the leverage constraint Λ with equality and with $x^s_i = 0$.

are no 'jumps' when the optimal solution switches from case1 to case2. We have $x^s(x^c_{case2}) = 0$ by construction. The continuous, monotonic function $x^s(x^c)$ crosses zero at most once, and at this point, by construction, $x^c_{case2} = x^c_{case1}$ ■

Case 3 Finally, with $\eta \neq 0$ therefore $x^c = 0$. We note that the corresponding value for x^s is coming from equation 4.A.3.

To sum up, the optimal solution of the problem is:

$$x_i^{c\star}(\Theta, x^c_{-i}) = \begin{cases} x^{c\star}_{case1}\left(\Theta, x^c_{-i}\right) & \text{if } x^c > 0 \text{ and } x^s > 0 \\ x^{c\star}_{case2}\left(\Theta, x^c_{-i}\right) & \text{if } x^c > 0 \text{ and } x^s = 0 \\ 0 & \text{otherwise} \end{cases}$$

In this highly parametric form it is impossible to decide more precisely analytically which of the three cases are consistent with the condition and which of them provide the global maximum. Numerical methods are implemented for illustration in section 4.4.

4.A.5 Proof of Theorem 2

The Pareto-optimal outcome maximizes the joint equity subject to the leverage constraints and nonnegativity. To formalize this problem we extend the notation with $W = \binom{w_1'}{w_2'}$ and $X = \binom{x_1'}{x_2'}$. The problem can be written as

$$\begin{aligned} \underset{\mathbf{X}}{\text{maximize}} \quad & \mathbf{1}'(\mathbf{W}'\mathbf{p}^1 - \mathbf{d}) \\ \text{subject to} \quad & (1-\ell_i)\mathbf{w_{i0}}'\mathbf{p}^\star - \mathbf{x_i}'\mathbf{p}^\star + \ell d_{i0} \leq 0 \qquad && \forall i \in \{1,2\} < \lambda_i > \\ & x_i^j \geq 0 && \forall i \in \{1,2\} \forall j \in \{s,c\} < \eta_i^j > \end{aligned}$$

Similarly to the equilibrium case, this also can be written in a reduced form, maximizing with respect to $\mathbf{x}^c$

$$\begin{aligned} \underset{\mathbf{x}^c}{\text{maximize}} \quad & ({w_1}^s + {w_2}^s)p^s + w_1^s\xi_1^s s_1^s + w_2^s\xi_2^s s_2^s + (w_1^c + w_2^c)p^c + w_1^c\xi^c(x_1^c + x_2^c) + w_2^c\xi^c(x_2^c + x_1^c) \\ \text{subject to} \quad & x^c \geq 0 \qquad < \eta > \\ & x^s \geq 0 \qquad < \lambda > \end{aligned}$$

The Kuhn-Tucker method leads to the following first-order and complementarity slackness conditions:

$$w_1^s \xi^s \frac{\partial x_1^s}{\partial x_1^c} + w_2^s \xi^s \frac{\partial x_2^s}{\partial x_1^c} + (w_1^c + w_2^c)\xi^c + \lambda_1 \frac{\partial x_1^s}{\partial x_1^c} + \lambda_2 \frac{\partial x_2^s}{\partial x_1^c} = 0$$

$$w_1^s \xi^s \frac{\partial x_1^s}{\partial x_2^c} + w_2^s \xi^s \frac{\partial x_2^s}{\partial x_2^c} + (w_1^c + w_2^c)\xi^c + \lambda_1 \frac{\partial x_1^s}{\partial x_2^c} + \lambda_2 \frac{\partial x_2^s}{\partial x_2^c} = 0$$

$$\lambda_i \geq 0; x^s \geq 0 \text{ but } \lambda_i(-x_i^s) = 0 \qquad < CS >$$

$$x^c \geq 0 \text{ but } x^c \left(\frac{\partial \mathcal{L}}{\partial x^c} \right) = 0 \qquad < CS >$$

(Interior optimum): Consider the case where $\lambda_1 = \lambda_2 = 0$ (interior optimum). Using the analytical expression for x_i and earlier results we start by deriving the partial derivatives:

$$\frac{\partial x_i^s}{\partial x_i^c} = \pm \frac{(1-l_i)w_i^c \xi^c - (p^c + \xi x_{-i}^c) - 2\xi^c x_i^c}{\sqrt{[(1-l_i)w_i^s \xi^s - p^s]^2 + 4\xi^s \left[(1-l_i)w_i^s p^s + (1-l_i)w_i^c p^{c\star} - x_i^c p^{c\star} + l_i d_i\right]}}$$

$$\frac{\partial x_i^s}{\partial x_{-i}^c} = \pm \frac{(1-l_i)w_i^c \xi^c - \xi^c x_i^c}{\sqrt{[(1-l_i)w_i^s \xi^s - p^s]^2 + 4\xi^s \left[(1-l_i)w_i^s p^s + (1-l_i)w_i^c p^{c\star} - x_i^c p^{c\star} + l_i d_i\right]}}$$

This is a sufficient implicit characterization of the social planner's solution.

(Compare with equilibrium): Recall that the conditions for an interior optimum (i.e. non-binding constraints, $\lambda = 0$) of the liquidation equilibrium (see 4.A.4): $\forall i \in \{1, 2\}$

$$w_i^s \xi^s \frac{\partial x_i^s}{\partial x_i^c} + w_i^c \xi^c + \lambda \frac{\partial x_i^s}{\partial x_i^c} = 0$$

Consider $i = 1$. The difference between the equilibrium and optimum FOC's is

$$\mathcal{D}_1 := w_2^s \xi^s \frac{\partial x_2^s}{\partial x_1^c} + w_2^c \xi^c$$

clearly $\mathcal{D}_1 < 0$ as long as

$$\frac{\partial x_2^s}{\partial x_1^c} > 0$$

Consider the first FOC. Assumption 4 guarantees that $\frac{\partial x_1^s}{\partial x_1^c} < 0$. The terms not containing the partial derivatives are constant - only depend on parameters. As long as $\mathcal{D}_1 < 0$ the first order conditions imply that

$$\left[\frac{\partial x_1^s}{\partial x_1^c}\right]^{OPT} > \left[\frac{\partial x_1^s}{\partial x_1^c}\right]^{EQ}$$

This implies $[x_1^c]^{OPT} > [x_1^c]^{EQ}$ as long as the derivative is increasing in x^c, that is,

$$\frac{\partial x_i^s}{\partial x_i^c \partial x_i^c} > 0$$

Consider the partial derivative $\frac{\partial x_i^s}{\partial x_i^c}$. Denote by (arg) the expression under $\sqrt{\ }$ in the expression for $\frac{\partial x_i^s}{\partial x_i^c}$ for simplicity. Then

$$\frac{\partial x_i^s}{\partial x_i^c \partial x_i^c} = \left(-2\xi^c \sqrt{(arg)} - NOM * \frac{1}{2\sqrt{(arg)}} \frac{\partial (arg)}{\partial x_i^c} \right) * \frac{1}{\sqrt{arg}^2}$$

This is positive as long as

$$\frac{\partial (arg)}{\partial x_i^c} = 4\xi^s \left((1 - l_i) w_i^c \xi^c - p^{c\star} - x_i^c \xi^c \right) > 0$$

which is positive by Assumption 4. Therefore, the second derivative is positive.

4.A.6 Proof of Theorem 3

Proof. There is only one value of x_i^s which solves $\Lambda_i(x_i^s, 0) = 0$, so the theorem is trivial if both equilibrium and optimum induces zero liquidation of common asset. We just need to prove that

(i) it is never optimal to sell common asset by the social planner if it is not optimal by the individual banks.

(ii) it is never equilibrium action to sell specific asset if the social planner does not sell it

Proof of part (i):
The second complementarity slackness condition of bank's problem implies that

$$x^c = 0 \Rightarrow \frac{\partial \mathcal{L}}{\partial x^c} < 0$$

that is

$$w^s \xi^s \frac{\partial x^s}{\partial x^c} + w^c \xi^c + \lambda \frac{\partial x^s}{\partial x^c} < 0$$

but by Assumption 2 here $\lambda = 0$, so the condition for interior optimality of the social planner's problem implies:

$$\mathcal{D}_1 > 0$$

but $\mathcal{D}_1 < 0$ was established under the proof of earlier Theorem. Contradiction.

Proof of part (ii):

We need to prove that $\lambda^{soc} > 0 \Rightarrow \lambda^{eq} > 0$. That is, a binding constraint ($x^s = 0$) in the social planner's problem implies binding constraint in the equilibrium problem. The social planner's problem first order conditions rearranged:

$$\left(w_1^s \xi^s \frac{\partial x_1^s}{\partial x_1^c} + w_1^c \xi^c\right) + \left(w_2^s \xi^s \frac{\partial x_2^s}{\partial x_1^c} + w_2^c \xi^c\right) + \left(\lambda_1 \frac{\partial x_1^s}{\partial x_1^c} + \lambda_2 \frac{\partial x_2^s}{\partial x_1^c}\right) = 0$$

We have established earlier that $\frac{\partial x_1^s}{\partial x_1^c} < 0$ and $\frac{\partial x_2^s}{\partial x_1^c} > 0$ by assumption 4. We also know that $|\frac{\partial x_1^s}{\partial x_1^c}| > |\frac{\partial x_2^s}{\partial x_1^c}|$. In a symmetric equilibrium $\lambda_1 = \lambda_2$ so the third bracketed term is negative. The second term is also negative. The first term must be therefore positive. As a consequence, the optimum-problem can only be fulfilled if λ^{opt} is positive. ■

4.A.7 Proof of Theorem 4

Consider the objective function

$$\Phi = \mathbf{w}_i' \mathbf{p}^{(1)} - d_i$$

We write out explicitly (without indices for simplicity):

$$\Phi = w^s(p^s + \xi^s x^s) + w^c(p^c + \xi^c(x^c + x_{-i}^c)) - d$$

In optimum, we can write the derivative w.r.t. ξ^c as

$$\frac{d\Phi}{d\xi^c} = w^s \xi^s \frac{d\tilde{x}^{s\star}}{d\xi^c} + w^c \xi^c \left(\frac{dx_1^{c\star}}{d\xi^c} + \frac{dx_2^{c\star}}{d\xi^c}\right) + w^c(x_1^{c\star} + x_2^{c\star}) \tag{4.10}$$

As we focus on symmetric equilibrium, we can impose the following identities:

$$\frac{\partial x_1^{c\star}}{\partial \xi^c} = \frac{\partial x_2^{c\star}}{\partial \xi^c} := \frac{\partial x^{c\star}}{\partial \xi^c}$$
$$x_1^{c\star} = x_2^{c\star} := x^{c\star}$$

Recall that $\tilde{x}^s$ is a function of x^c (and x_{-i}^c), which also hold in any equilibrium. We write the total derivative w.r.t. ξ^c as

$$\frac{ds^{s\star}}{d\xi^c} = \frac{\partial \tilde{x}^{s\star}}{\partial \xi^c} + \frac{\partial \tilde{x}^{s\star}}{\partial x^c}\frac{\partial x^c}{\partial \xi^c} + \frac{\partial \tilde{x}^{s\star}}{\partial x_{-i}^c}\frac{\partial x_{-i}^c}{\partial \xi^c}$$

The term $\frac{\partial \tilde{x}^{s\star}}{\partial \xi^c}$ and $\frac{\partial \tilde{x}^{s\star}}{\partial x^c_{-i}}$ is straightforward to calculate. The term $\frac{\partial \tilde{x}^{s\star}}{\partial x^c}$ is given by the equilibrium condition which must be hold for equilibrium quantities:

$$\frac{\partial \tilde{x}^{s\star}}{\partial x^c} = -\frac{w^c \xi^c}{w^s \xi^s}$$

Using this, we can rewrite the *first term in* (4.10) as:

$$w^s \xi^s \left(\frac{\partial \tilde{x}^{s\star}}{\partial \xi^c} - \frac{w^c \xi^c}{w^s \xi^s} \frac{\partial x^c}{\partial \xi^c} + \frac{\partial \tilde{x}^{s\star}}{\partial x^c_{-i}} \frac{\partial x^c_{-i}}{\partial \xi^c} \right) = w^s \xi^s \frac{\partial \tilde{x}^{s\star}}{\partial \xi^c} + \left(w^s \xi^s \frac{\partial \tilde{x}^{s\star}}{\partial x^c_{-i}} - w^c \xi^c \right) \frac{\partial x^c}{\partial \xi^c}$$

The condition for the theorem becomes therefore

$$\Psi := w^s \xi^s \frac{\partial \tilde{x}^{s\star}}{\partial \xi^c} + \left(w^s \xi^s \frac{\partial \tilde{x}^{s\star}}{\partial x^c_{-i}} + w^c \xi^c \right) \frac{\partial x^c}{\partial \xi^c} + 2 w^c x^c < 0 \tag{4.11}$$

This can be fully calculated analytically in a straightforward way. We proceed with calculating the term $\frac{\partial x^{c\star}}{\partial \xi^c}$. We depart from the implicit analytical expression for $x^{c\star}$. Since we study symmetric equilibrium here, we can replace x_{-i} by x^c and so the implicit equation takes the form of (note that the arguments of A, B, C stand for the other player's action here):

$$A(x^c)[x^c]^2 + B(x^c)[x^c] + C(x^c) = 0$$

For notation let the LHS be function $g(\xi^c, \cdot)$. Using the implicit function theorem:

$$\frac{\partial x^c}{\partial \xi^c} = -\frac{\partial g / \partial \xi^c}{\partial g / \partial x^c}$$

The two partial derivatives are:

$$\begin{aligned}
\frac{\partial g}{\partial \xi^c} &= \frac{\partial A}{\partial \xi^c}[x^c]^2 + \frac{\partial B}{\partial \xi^c}[x^c] + \frac{\partial C}{\partial \xi^c} \\
\frac{\partial g}{\partial x^c} &= \frac{\partial A}{\partial x^c}[x^c]^2 + 2A(x^c)[x^c] + \frac{\partial B}{\partial x^c}[x^c] + B(x^c) + \frac{\partial C}{\partial x^c}
\end{aligned}$$

All elements of the expression are trivial to calculate. The derivatives w.r.t. ξ:

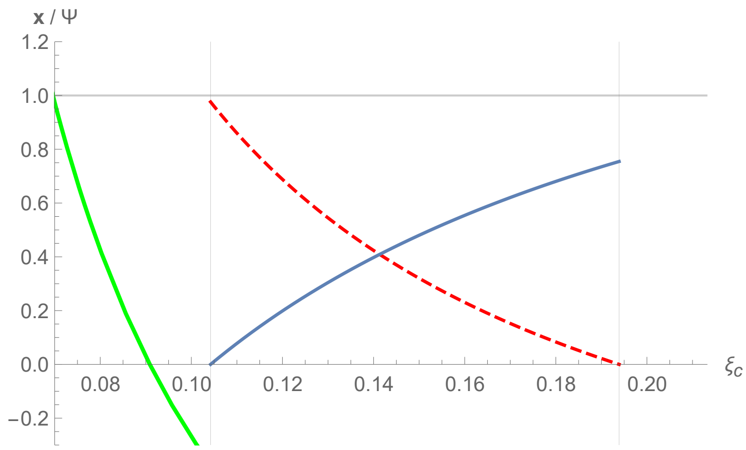
$$\begin{aligned}
\frac{\partial A}{\partial \xi^c} &= 4\xi^s \\
\frac{\partial B}{\partial \xi^c} &= 4 \left(x^c - (1-l) w^c - \left[\frac{\xi^s w^s}{\xi^c w^c} \right]^2 p^c \right) \\
\frac{\partial C}{\partial \xi^c} &= -\frac{2\xi^s \left((l-1)\xi^c \xi^s p^c w^c [w^s]^2 - 2(l-1)\xi^{c3}[w^c]^3 x^c + \xi^s p^c [w^s]^2 (p^c + \xi^c x^c) \right)}{\xi^{c3}[w^c]^2}
\end{aligned}$$

The derivatives w.r.t. x

$$\frac{\partial A}{\partial x^c} = 0$$
$$\frac{\partial B}{\partial x^c} = 4\xi^s\xi^c + 4\left(\frac{\xi^s w^s}{w^c}\right)^2$$
$$\frac{\partial C}{\partial x^c} = \frac{2\xi^s\left(2(l-1)\xi^{c2}w^{c3} + (l-1)\xi^c\xi^s w^c w^{s2} + \xi^s w^{s2}(p^c + \xi^c x^c)\right)}{\xi^c w^{c2}}$$

Using these numerical results it is possible to establish analytically the first two terms in Ψ are negative. Unfortunately, it is not possible to analytically show that $\Psi < 0$, so I use numerical illustrations. The following figure shows that whenever the conditions for an interior equilibrium holds, the value of Ψ (green, thick curve) remains negative. This is true for every parameter combinations tested during numerical analysis.

Figure 4.11: Value of Ψ in an interior equilibrium



Chapter 5

Concluding remarks

This thesis included three essays on banking theory.

The first essay described a novel mechanism which demonstrates how banking regulation might affect banks' incentives to exert risk management through an informational channel. Our research contributes to the broad literature on bank liquidity and capital management, and their role in financial stability. The mechanism we uncover in the paper is potentially of great interest for central banks and other regulatory bodies, as it investigates nontrivial, and possibly overlooked general equilibrium consequences of micro-prudential regulatory interventions. A potential follow-up research could extend significantly our preliminary empirical work. The new liquidity regulation introduced in Basel III is gradually rolled out to more and more countries, which creates the potential for extensive empirical research. It is important to mention that liquidity regulation is also being introduced in the asset management industry, which also creates a laboratory for empirical testing. Methodologically, the paper combines signalling with global games in a novel and tractable way. We consider the paper primarily a contribution to the banking literature, rather than to economic methodology, and keep the complexity of the interaction between signalling and global games as simple as possible. It is, however, of great interest to study the universe of possible models this machinery could produce in follow-up research.

The second essay departs from the empirical observation that pricing of retail banking products is significantly different across countries with otherwise comparable level of financial development. It is especially puzzling that these differences persist within the European Union, inside the single market for services. We believe there is no plausible theory yet to explain those country-wise differences. The model we propose in this paper takes the same basic view on retail banking markets as the previous theoretical literature, namely that personal current accounts are loss-leader primary markets, and banks recoup the costs on less competitive aftermarkets. However, we model this with an important distinction by incorporating aftermarket competition, and explicitly modelling two sources of market power which are specific to banking: customer naiveté, and adverse selection. We demonstrate that the competitive aftermarket assumption guarantees the existence of 'free banking' equilibrium even with relatively low number of naive customers, an effect which is not present in the previous literature. Furthermore, we show that the two forces in the model interact in equilibrium and reinforce each other. Specifically, the presence of adverse selection makes even more likely that base prices hit the lower bound. We believe the paper could be a starting point to a more in-depth analysis of the possible impacts of recently introduced regulations, both empirically and theoretically. For example, the Payment Services Directive introduced by the European Commission, and the 'Open banking' program in the UK specifically aims to reduce adverse selection in banking markets by requiring banks to develop API-s to access customer data by third parties as per the customers' request. The model predicts that this effort might lead to the end of free banking - which, perhaps surprisingly for many, is actually a welfare-improving intervention. The current version of our work leaves open further theoretical analysis of the welfare effects and the impact of regulatory intervention, which is left for future work. The gradual introduction of new programs in the EU offers an opportunity for extensive empirical research in understanding the nature of competition in retail banking.

The third essay identifies a 'liquidity trap' situation in the presence of asset commonalities. When financial institutions engage in selling part of their correlated portfolio as a result of a systemic shock, strategic interactions generate a game which is reminiscent to a Prisoners' dilemma, where players over-liquidate the more liquid asset commonality. In this thesis I derive the results for a simple case in a duopoly model of banking with an idiosyncratic asset and a common asset. Preliminary calculations, not included in the dissertation suggest that the results can be generalized to an arbitrary bipartite network structure of asset holdings. Perhaps even more interesting follow-up work could be to empirically test the economic relevance of the uncovered effects. The mutual fund portfolio database provides an excellent laboratory for these tests, which could be a basis of a follow-up empirical project.

Bibliography

Daron Acemoglu, Vasco M Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. The network origins of aggregate fluctuations. *Econometrica*, 80(5):1977–2016, 2012.

Daron Acemoglu, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. Systemic risk and stability in financial networks. *American Economic Review*, 105(2):564–608, 2015.

Viral V Acharya and S Viswanathan. Leverage, moral hazard, and liquidity. *The Journal of Finance*, 66(1):99–138, 2011.

Robert M Adams. Bank fees, aftermarkets, and consumer behavior. *Finance and Economics Discussion Series*, 2017-054, 2017.

Tobias Adrian and Hyun Song Shin. Liquidity and leverage. *Journal of financial intermediation*, 19(3):418–437, 2010.

Sule Alan, Mehmet Cemalcilar, Dean Karlan, and Jonathan Zinman. Unshrouding: Evidence from bank overdrafts in turkey. *The Journal of Finance*, 73(2):481–522, 2018.

Franklin Allen and Douglas Gale. Limited market participation and volatility of asset prices. *The American Economic Review*, pages 933–955, 1994.

Franklin Allen and Douglas Gale. Optimal financial crises. *Journal of finance*, pages 1245–1284, 1998.

Franklin Allen and Douglas Gale. Financial contagion. *Journal of political economy*, 108(1):1–33, 2000.

Hamed Amini, Damir Filipović, and Andreea Minca. Uniqueness of equilibrium in a payment system with liquidation costs. *Operations Research Letters*, 44(1):1–5, 2016.

George-Marios Angeletos and Alessandro Pavan. Selection-free predictions in global games with endogenous information and multiple equilibria. *Theoretical Economics*, 8(3):883–938, 2013.

George-Marios Angeletos, Christian Hellwig, and Alessandro Pavan. Signaling in a global game: Coordination and policy traps. *Journal of Political economy*, 114 (3):452–484, 2006.

Mark Armstrong and John Vickers. Consumer protection and contingent charges. *Journal of Economic Literature*, 50(2):477–93, 2012.

Kerstin Awiszus and Stefan Weber. The joint impact of bankruptcy costs, crossholdings and fire sales on systemic risk in financial networks. *Preprint*, 2015.

William H Beaver and Ellen E Engel. Discretionary behavior with respect to allowances for loan losses and the behavior of security prices. *Journal of Accounting and Economics*, 22(1):177–206, 1996.

Nicole Boyson, Jean Helwege, and Jan Jindra. Crises, liquidity shocks, and fire sales at commercial banks. *Financial Management*, 43(4):857–884, 2014.

Richard Brealey, Hayne E Leland, and David H Pyle. Informational asymmetries, financial structure, and financial intermediation. *The journal of Finance*, 32(2): 371–387, 1977.

Markus K Brunnermeier and Lasse Heje Pedersen. Market liquidity and funding liquidity. *Review of Financial studies*, 22(6):2201–2238, 2009.

Kenneth Burdett and Kenneth L Judd. Equilibrium price dispersion. *Econometrica: Journal of the Econometric Society*, pages 955–969, 1983.

Fabio Caccioli, Munik Shrestha, Cristopher Moore, and J Doyne Farmer. Stability analysis of financial contagion due to overlapping portfolios. *Journal of Banking & Finance*, 46:233–245, 2014.

Fabio Caccioli, J Doyne Farmer, Nick Foti, and Daniel Rockmore. Overlapping portfolios, contagion, and financial stability. *Journal of Economic Dynamics and Control*, 51:50–63, 2015.

Bruce I Carlin. Strategic price complexity in retail financial markets. *Journal of financial Economics*, 91(3):278–287, 2009.

Hans Carlsson and Eric Van Damme. Global games and equilibrium selection. *Econometrica: Journal of the Econometric Society*, pages 989–1018, 1993.

Cristina Cella, Andrew Ellul, and Mariassunta Giannetti. Investors' horizons and the amplification of market shocks. *Review of Financial Studies*, page hht023, 2013.

Chen Chen, Garud Iyengar, and Ciamac C Moallemi. Asset-based contagion models for systemic risk. *Operations Research*, 2014.

Rodrigo Cifuentes, Gianluigi Ferrucci, and Hyun Song Shin. Liquidity risk and concation. *Journal of the European Economic Association*, 3(2-3):556–566, 2005.

Filippo De Marco. Bank lending and the sovereign debt crisis. Technical report, Boston College working paper, 2013.

Douglas W Diamond and Philip H Dybvig. Bank runs, deposit insurance, and liquidity. *The journal of political economy*, pages 401–419, 1983.

Douglas W Diamond and Anil K Kashyap. Liquidity requirements, liquidity choice, and financial stability. *Handbook of Macroeconomics*, 2:2263–2303, 2016.

Douglas W Diamond and Raghuram G Rajan. A theory of bank capital. *The Journal of Finance*, 55(6):2431–2465, 2000.

Douglas W Diamond and Raghuram G Rajan. Fear of fire sales, illiquidity seeking, and credit freezes. *The Quarterly Journal of Economics*, 126(2):557–591, 2011.

Fernando Duarte and Thomas M Eisenbach. Fire-sale spillovers and systemic risk. *FRB of New York Staff Report*, (645), 2014.

Larry Eisenberg and Thomas H Noe. Systemic risk in financial systems. *Management Science*, 47(2):236–249, 2001.

Matthew Elliott, Benjamin Golub, and Matthew O Jackson. Financial networks and contagion. *American Economic Review*, 104(10):3115–3153, 2014.

Giovanni Favara and Mariassunta Giannetti. Forced asset sales and the concentration of outstanding debt: Evidence from the mortgage market. 2015.

Z. Feinstein. Financial contagion and asset liquidation strategies. *Preprint*, 2015.

Z. Feinstein and F. El-Masri. The effects of leverage requirements and fire sales on financial contagion via asset liquidation strategies in financial networks. *Preprint*, 2015.

Xavier Freixas, Bruno M Parigi, and Jean-Charles Rochet. Systemic risk, interbank relations, and liquidity provision by the central bank. *Journal of money, credit and banking*, pages 611–638, 2000.

Xavier Gabaix. The granular origins of aggregate fluctuations. Technical report, National Bureau of Economic Research, 2009.

Xavier Gabaix and David Laibson. Shrouded attributes, consumer myopia, and information suppression in competitive markets. *The Quarterly Journal of Economics*, 121(2):505–540, 2006.

Paul Glasserman and H Peyton Young. How likely is contagion in financial networks? *Journal of Banking & Finance*, 50:383–399, 2015.

Itay Goldstein and Chong Huang. Bayesian persuasion in coordination games. *The American Economic Review*, 106(5):592–596, 2016.

Itay Goldstein and Ady Pauzner. Demand–deposit contracts and the probability of bank runs. *the Journal of Finance*, 60(3):1293–1327, 2005.

Theoharry Grammatikos and Anthony Saunders. Additions to bank loan-loss reserves: Good news or bad news? *Journal of Monetary Economics*, 25(2):289–304, 1990.

Robin Greenwood, Augustin Landier, and David Thesmar. Vulnerable banks. *Journal of Financial Economics*, 115(3):471–485, 2015.

Paul A Griffin and Samoa JR Wallach. Latin american lending by major us banks: The effects of disclosures about nonaccrual loans and loan loss provisions. *Accounting Review*, pages 830–846, 1991.

Denis Gromb and Dimitri Vayanos. Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of financial Economics*, 66(2):361–407, 2002.

Michael D Grubb. Consumer inattention and bill-shock regulation. *The Review of Economic Studies*, 82(1):219–257, 2014.

Milton Harris and Artur Raviv. The theory of capital structure. *the Journal of Finance*, 46(1):297–355, 1991.

Oliver Hart. Theories of optimal capital structure: a managerial discretion perspective. *NBER Working Paper*, (R1806), 1993.

Oliver Hart and J Moore. Debt and seniority: An analysis of the role of hard claims in constraining management. *American Economic Review*, 85(3), 1995.

Harald Hau and Sandy Lai. The role of equity funds in the financial crisis propagation. 2012.

Zhiguo He. The sale of multiple assets with private information. *The Review of Financial Studies*, 22(11):4787–4820, 2009.

Paul Heidhues and Botond Kőszegi. Naivete-based discrimination. *The Quarterly Journal of Economics*, 132(2):1019–1054, 2017.

Paul Heidhues, Botond Kőszegi, and Takeshi Murooka. Exploitative innovation. *American Economic Journal: Microeconomics*, 8(1):1–23, 2016a.

Paul Heidhues, Botond Kőszegi, and Takeshi Murooka. Inferior products and profitable deception. *The Review of Economic Studies*, 84(1):323–356, 2016b.

Joseph P Hughes and Loretta J Mester. Bank capitalization and cost: Evidence of scale economies in risk management and signaling. *The review of Economics and Statistics*, 80(2):314–325, 1998.

Nicolas Inostroza and Alessandro Pavan. Persuasion in global games with application to stress testing. *Economist*, 2017.

Chotibhak Jotikasthira, Christian Lundblad, and Tarun Ramadorai. Asset fire sales and purchases and the international transmission of funding shocks. *The Journal of Finance*, 67(6):2015–2050, 2012.

Michael Kosfeld and Ulrich Schüwer. Add-on pricing in retail financial markets and the fallacies of consumer education. *Review of Finance*, 21(3):1189–1216, 2017.

Frédéric Malherbe. Self-fulfilling liquidity dry-ups. *The Journal of Finance*, 69(2): 947–970, 2014.

Alberto Manconi, Massimo Massa, and Ayako Yasuda. The role of institutional investors in propagating the crisis of 2007–2008. *Journal of Financial Economics*, 104(3):491–518, 2012.

Brian T Melzer and Donald P Morgan. Competition in a consumer loan market: Payday loans and overdraft credit. *Journal of Financial Intermediation*, 24(1): 25–44, 2015.

Craig B Merrill, Taylor Nadauld, René M Stulz, and Shane M Sherlund. Were there fire sales in the rmbs market? *Charles A. Dice Center Working Paper*, (2014-09), 2014.

Chun-Hui Miao. Consumer myopia, standardization and aftermarket monopolization. *European Economic Review*, 54(7):931–946, 2010.

Donald P Morgan, Michael R Strain, and Ihab Seblani. How payday credit access affects overdrafts and other outcomes. *Journal of money, Credit and Banking*, 44 (2-3):519–531, 2012.

Stephen Morris and Hyun Song Shin. Unique equilibrium in a model of self-fulfilling currency attacks. *American Economic Review*, pages 587–597, 1998.

Stephen Morris and Hyun Song Shin. Rethinking multiple equilibria in macroeconomic modeling. In *NBER Macroeconomics Annual 2000, Volume 15*, pages 139–182. MIT Press, 2001.

Chakravarthi Narasimhan. Competitive promotional strategies. *Journal of business*, pages 427–449, 1988.

Antonio Nicolo and Loriana Pelizzon. Credit derivatives, capital requirements and opaque otc markets. *Journal of Financial Intermediation*, 17(4):444–463, 2008.

Louis Raffestin. Diversification and systemic risk. *Journal of Banking & Finance*, 46:85–106, 2014.

Raghuram G Rajan. Insiders and outsiders: The choice between informed and arm's-length debt. *The Journal of Finance*, 47(4):1367–1400, 1992.

Stephen A Ross. The determination of financial structure: the incentive-signalling approach. *The bell journal of economics*, pages 23–40, 1977.

Myron S Scholes, G Peter Wilson, and Mark A Wolfson. Tax planning, regulatory capital planning, and financial reporting strategy for commercial banks. *The Review of Financial Studies*, 3(4):625–650, 1990.

Carl Shapiro. Aftermarkets and consumer welfare: Making sense of kodak. *Antitrust LJ*, 63:483, 1994.

Steven A Sharpe. Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *The journal of finance*, 45(4):1069–1087, 1990.

Andrei Shleifer and Robert Vishny. Fire sales in finance and macroeconomics. *Journal of Economic Perspectives*, 25(1):29–48, 2011.

Andrei Shleifer and Robert W Vishny. Liquidation values and debt capacity: A market equilibrium approach. *The Journal of Finance*, 47(4):1343–1366, 1992.

Andrei Shleifer and Robert W Vishny. The limits of arbitrage. *The Journal of Finance*, 52(1):35–55, 1997.

Victor Stango and Jonathan Zinman. What do consumers really pay on their checking and credit card accounts? explicit, implicit, and avoidable costs. *American Economic Review*, 99(2):424–29, 2009.

Victor Stango and Jonathan Zinman. Limited and varying consumer attention: Evidence from shocks to the salience of bank overdraft fees. *The Review of Financial Studies*, 27(4):990–1030, 2014.

Joseph E Stiglitz. Risk and global economic architecture: Why full financial integration may be undesirable. *American Economic Review*, 100(2):388–92, 2010.

Paolo Tasca, Pavlin Mavrodiev, and Frank Schweitzer. Quantifying the impact of leveraging and diversification on systemic risk. *Journal of Financial Stability*, 15: 43–52, 2014.

Christian Upper. Simulation methods to assess the danger of contagion in interbank markets. *Journal of Financial Stability*, 7(3):111–125, 2011.

Hal R Varian. A model of sales. *The American Economic Review*, 70(4):651–659, 1980.

Ernst-Ludwig Von Thadden. Asymmetric information, bank lending and implicit contracts: the winner's curse. *Finance Research Letters*, 1(1):11–23, 2004.

Wolf Wagner. Systemic liquidation risk and the diversity–diversification trade-off. *The Journal of Finance*, 66(4):1141–1175, 2011.

Marlon L Williams. Bank overdraft pricing and myopic consumers. *Economics Letters*, 139:84–87, 2016.

Yves Zenou. Key players. Technical report, CEPR Discussion Papers, 2014.