

ESSAYS IN MICROECONOMICS

THOMAS M. EISENBACH

A DISSERTATION

PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF ECONOMICS

ADVISER: MARKUS K. BRUNNERMEIER

MAY 2011

UMI Number: 3458609

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3458609

Copyright 2011 by ProQuest LLC.

All rights reserved. This edition of the work is protected against unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

© Copyright by Thomas M. Eisenbach, 2011.

All rights reserved.

Abstract

This dissertation consists of three self-contained essays in microeconomics.

The first chapter studies the rollover risk of financial institutions. It presents a competitive-equilibrium model of financial institutions optimally choosing their debt maturity structure. Rollover risk arises endogenously from the interaction of creditors in a global-game framework. When only idiosyncratic risk is present short-term debt acts as an effective disciplining device but once aggregate risk is added a two-sided inefficiency arises. Good aggregate states lead to excessive risk-taking while bad aggregate states suffer from fire-sale liquidation. In the competitive equilibrium with endogenous liquidation values, the two-sided inefficiency reinforces itself through a feedback effect. It increases the volatility of liquidation values and thereby amplifies the impact of aggregate risk.

The second chapter is coauthored with Martin Schmalz and studies the effects of anxiety in decision-making. We model an anxious agent as one who is more risk averse for imminent than for distant risk. Such preferences can lead to dynamic inconsistencies with respect to risk trade-offs. We derive implications for financial markets such as over-trading and price anomalies around announcement dates which are found empirically. We show that strategies to cope with anxiety can explain costly delegation of investment decisions. Finally, we model how an anxiety-prone agent may endogenously become overconfident and take excessive risks.

The third chapter is coauthored with Dirk Bergemann, Joan Feigenbaum and Scott Shenker and studies markets for digital goods. We consider the optimal design of flexible use in a digital-rights-management policy. Consumers can acquire a digital good either as a licensed or an unlicensed copy. The availability of unlicensed copies is increasing in the flexibility accorded to licensed copies. We show

that this results in a key trade-off between the value of licensed copies and the threat of piracy. We augment the basic model by introducing a secure platform that is required to use the digital good. We show crucial differences in the equilibrium depending on whether platform and content are sold by two separate firms or by a single integrated one.

Acknowledgments

I am indebted to my advisers Markus Brunnermeier and Stephen Morris for their guidance. I would like to thank my coauthors Dirk Bergemann, Joan Feigenbaum, Martin Schmalz and Scott Shenker for their contributions to our joint work. For helpful comments and discussion I would like to thank Magdalena Berger, Dong Choi, Paolo Colla, Douglas Diamond, Itay Goldstein, Zhiguo He, David Hirshleifer, Jakub Jurek, Bryan Kelly, Jia Li, Konstantin Milbradt, Benjamin Moll, Martin Oehmke, Justinas Pelenis, Wolfgang Pesendorfer, José Scheinkman, Klaus Schmidt, Felipe Schwartzman, Hyun Shin, David Sraer, Jeremy Stein, Ilya Strebulaev, Amir Sufi, Wei Xiong, Adam Zawadowski and Sergey Zhuk as well as seminar audiences at Carnegie Mellon Tepper, Chicago Booth, Duke Fuqua, London Business School, London School of Economics, New York Fed, Princeton, Stanford GSB, UC Irvine Merage, UI Urbana-Champaign, University of Munich and Wharton.

Contents

Abstract	iii
Acknowledgments	v
List of Figures	viii
1 Rollover Risk: Optimal but Inefficient	1
1.1 Introduction	1
1.2 Model	8
1.3 Endogenous Rollover Risk	13
1.4 Equilibrium without Aggregate Risk	17
1.5 Equilibrium with Aggregate Risk	23
1.6 Numerical Simulation	30
1.7 Discussion	33
1.8 Conclusion	36
1.9 Appendix	37
2 Anxiety in the Face of Risk	44
2.1 Introduction	44
2.2 Model	52
2.3 Finance Applications	54
2.4 Overconfidence	64

<i>CONTENTS</i>	vii
2.5 Conclusion	70
2.6 Appendix	73
3 Flexibility as an Instrument in DRM	76
3.1 Introduction	76
3.2 Model	82
3.3 Optimal Flexibility and Price	85
3.4 Sales of Platforms for Digital Goods	87
3.5 Conclusion	92
3.6 Appendix	93
Bibliography	96

List of Figures

1.1	Project time-line for bank i	9
1.2	Implemented and efficient liquidation policy without aggregate risk	21
1.3	Inefficiency in state H	26
1.4	Inefficiency in state L	26
1.5	Percentage of ex-ante surplus destroyed by the inefficiency	32
1.6	Amplification effect for $p = 0.7$ and $q = 0.05$	33

Chapter 1

Rollover Risk:

Optimal but Inefficient¹

1.1 Introduction

Short-term debt and the rollover risk it creates have been prominent features of the financial crisis of 2007–2008. This paper presents a model of banks optimally choosing the maturity structure of their debt.² The banks interact in a competitive equilibrium framework with endogenous interest rates and liquidation values. The rollover risk for a given maturity structure arises endogenously from the coordination problem of a bank’s creditors which is captured by a global game. The model distinguishes between two exogenous sources of risk faced by an individual bank, idiosyncratic risk specific to the bank’s assets as well as aggregate risk if

¹An early version of this paper was presented under the title “The Good and Bad of Liquidity Risk” at the conference EconCon at Princeton University in August 2010.

²Throughout the paper I mostly use the term “bank.” However, the paper applies not only to traditional banks but any type of leveraged market-based financial institution.

assets are correlated across banks. After receiving additional information about the two sources of risk, the banks have a strong incentive to take excessive risks and therefore use short-term debt as a disciplining device.

The main contribution of this paper is to show an important inefficiency that arises from the use of short-term debt in the presence of aggregate risk. While short-term debt acts as an effective disciplining device when banks only face idiosyncratic risk, it is severely undermined when aggregate risk is added. The problem is that the disciplining effect is too weak in good aggregate states and too powerful in bad aggregate states. This leads to a two-sided inefficiency: In good aggregate states the banks take excessive risks in the form of projects with negative net present value. Bad aggregate states suffer from fire sales as projects with positive net present value are liquidated. As a result, economic surplus is destroyed in both situations.

In addition, the paper uses the competitive equilibrium framework with endogenous liquidation values to highlight that the inefficiency reinforces itself through a feedback effect. Given the presence of aggregate risk, even the first-best allocation has liquidation values that vary across aggregate states. However, the use of short-term debt – whose problem originates in this variation – further increases the volatility of liquidation values and thereby amplifies the impact of aggregate risk.

To be more specific, I model a group of banks, each with the opportunity to invest in a project of its own. After the investment decision is made, additional information about each project's expected payoff becomes available and a bank can decide whether to continue or liquidate its project. A key assumption of my model that differs from most of the literature on rollover risk is that liquidation is not inherently inefficient. Liquidated assets are employed in a secondary sector so the liquidation value reflects the true economic value of the assets in alternative uses.

This implies that liquidation is only inefficient if the assets' value in the secondary sector is less than their expected value in current use as the bank's project. Importantly, this also implies that *not* liquidating is inefficient if the project's expected payoff is less than the assets' value in the secondary sector.

The secondary sector exhibits decreasing marginal productivity which implies that the liquidation value a bank receives is decreasing in aggregate asset sales. If projects are correlated across banks, leading to uncertainty about the equilibrium level of aggregate asset sales, this gives rise to aggregate risk with states of the world that differ in total asset sales and therefore liquidation values. Each individual bank then faces two sources of risk, idiosyncratic risk about its own project payoff and aggregate risk about the liquidation value determined by aggregate asset sales.

Unless a bank is fully equity financed, it has the wrong incentives when it comes to continuing or liquidating its project. Similar to a risk-shifting problem, the bank has an incentive to continue excessively risky projects at the cost of debt holders, i.e. projects whose expected payoff has turned out to be less than the liquidation value. Therefore a bank's choice of maturity structure and the implied exposure to rollover risk play an important role for the realized economic surplus of the bank's project.

A bank can choose any combination of long-term and short-term debt to finance its investment. While long-term debt has the same maturity as the project's final payoff, short-term debt has to be rolled over after the additional information about the project's expected payoff and the liquidation value becomes available. Rollover risk arises since it may not be possible to satisfy all withdrawals of short-term creditors, even by liquidating all of the bank's assets.

I model the resulting coordination problem among short-term creditors as a global game and derive a unique equilibrium with very intuitive properties. After bad news the short-term creditors withdraw their loans and the bank has to be liquidated while after good news the creditors roll over and the project is continued. Due to the two sources of risk there are two ways in which news can be bad. A creditor run can be triggered by bad idiosyncratic news about the bank itself or by bad aggregate news about the liquidation value. In addition, the two sources of risk interact in determining a bank's rollover risk. A bank is more vulnerable to idiosyncratic news for bad aggregate news and more vulnerable to aggregate news for bad idiosyncratic news.

Since the global game equilibrium is unique and has continuous comparative statics, a bank's initial choice of debt maturity structure directly translates into its exposure to rollover risk when the additional information becomes available. By choosing a greater fraction of short-term debt, the bank increases the risk that it suffers a run and has to liquidate its project.

To distinguish between the different effects of the two sources of risk on a bank's maturity structure choice, I first analyze the benchmark case without aggregate risk, i.e. with projects that are uncorrelated across banks. Without uncertainty about the liquidation value, a bank has full control over the amount of rollover risk it exposes itself to. Therefore the bank chooses its financing exactly so as to implement the efficient liquidation policy where the project is liquidated if and only if the expected payoff turns out to be less than the liquidation value. Optimally exposing itself to rollover risk allows the bank to fully eliminate its incentive problem, maximizing its project's ex-ante and interim net present value.

Adding aggregate risk in form of correlated projects and a random liquidation value has two important effects. First, the optimal liquidation policy now depends on the realization of the liquidation value. If the liquidation value turns out to be high, efficiency requires liquidating projects that should be continued if the liquidation value were low. At the same time, the bank's rollover risk given the maturity structure chosen ex ante now varies with the realization of the liquidation value. If the liquidation value turns out to be high, creditors are less worried about the bank's liquidity, making a run less likely.

The key problem is that these two effects go in opposite directions. For a high liquidation value, *more* projects should be liquidated but the bank's increased stability leads to *less* liquidation. For a low liquidation value, *less* projects should be liquidated but the bank's reduced stability leads to *more* liquidation. Aggregate risk effectively drives a wedge between the efficient liquidation policy and the achievable liquidation policy. A bank optimally chooses its maturity structure but can no longer achieve the efficiency of the benchmark case without aggregate risk.

Given the optimal maturity structure, the disciplining effect of short-term debt is weaker than required in good aggregate states, allowing the bank to continue projects with negative net present value. This means that the bank is taking excessive risks with assets that have more valuable use elsewhere. As a mirror image, the disciplining effect is stronger than required in bad aggregate states, forcing the bank to liquidate projects with positive net present value. Here assets are sold at fire-sale prices which correspond to their actual value in alternative uses but are below their value in current use.

If projects are correlated then even the first-best allocation implies more asset sales and lower liquidation values in bad aggregate states than in good aggregate

states. However, the competitive equilibrium always has *more* liquidation in bad states and *less* liquidation in good states than is efficient. This means that compared to the first-best allocation, liquidation values are higher in good states and lower in bad states, increasing volatility. Not only is the volatility of liquidation values causing the inefficiency in the first place, it is also further amplified in the competitive equilibrium. As a result, banks face greater aggregate risk than they would in the first-best allocation. Nevertheless, the competitive equilibrium is constrained efficient so there is no scope for policy intervention to improve welfare by changing banks' use of short-term debt.

Related Literature The role of short-term debt as a disciplining device has been discussed in a literature going back to [Calomiris and Kahn \(1991\)](#).³ A common feature of this literature is that the benefit of a disciplining effect comes at the cost of inefficient liquidation and the choice of maturity structure has to trade off the two. My paper differs, first, in the fact that liquidation is not per se inefficient and, second, in the distinction between two sources of risk. In particular, my model has an efficient outcome if only idiosyncratic risk is present. The new inefficiency in my model arises because of the inability of the disciplining mechanism to deal with two sources of risk. This leads to an inefficient outcome in good as well as bad aggregate states which the optimal maturity structure has to trade off. Another recent paper on optimal maturity structure choice is [Brunnermeier and Oehmke \(2010\)](#). Their model doesn't have a disciplining problem so the optimal maturity structure is a corner solution of either all short-term or all long-term debt.

³See, e.g. [Rajan \(1992\)](#), [Leland and Toft \(1996\)](#) and [Diamond and Rajan \(2001\)](#). For a recent approach with interesting dynamic effects see [Cheng and Milbradt \(2010\)](#).

The paper is also related to [Shleifer and Vishny \(1992\)](#) who study the interaction of debt as a disciplining device with endogenous liquidation values. In their model, disciplining is only necessary in the good state and liquidation always happens at a (potentially inefficient) discount in the bad state. The focus of their paper is how equilibrium liquidation values limit debt capacity. In my model assets are always sold to outsiders but not necessarily at a discount. More importantly, the incentive problem in my model is present in all aggregate states so the optimal maturity structure has to trade-off the two inefficiencies of too much liquidation in one state and too little liquidation in the other.

Related from a technical point of view are several papers also using a global game setup to model the coordination problem among creditors, notably [Morris and Shin \(2004\)](#), [Rochet and Vives \(2004\)](#) and [Goldstein and Pavan \(2005\)](#).⁴ In my paper, the global game is not as much front and center but rather used as a convenient modeling device. The convenience stems from the fact that under weak assumptions the global game has a unique equilibrium and that this equilibrium has continuous comparative statics. This allows me to study an ex-ante stage where the maturity structure is chosen optimally, taking into account the effect on the global-game equilibrium at a later stage. Finally, since the global game itself is restricted to a single time period, I avoid the complications in dynamic global games pointed out by [Angeletos, Hellwig, and Pavan \(2007\)](#).

The rest of the paper is structured as follows. In Section 1.2 I lay out the model and highlight the important features. I then proceed according to backwards induction and first derive the endogenous rollover risk for a given maturity structure

⁴In a related model not using a global game setup, [He and Xiong \(2010\)](#) study the inter-temporal coordination problem among creditors with different maturity dates and derive very similar comparative statics.

in Section 1.3. Then I analyze the optimal maturity structure choice and competitive equilibrium in Section 1.4 for the efficient benchmark case without aggregate risk. In Section 1.5 I add aggregate risk and illustrate the two-sided inefficiency this causes. To quantify the effects of the model, I run a numerical simulation in Section 1.6. Finally, in Section 1.7 I discuss the robustness of the main results as well as some extensions. Section 1.8 concludes.

1.2 Model

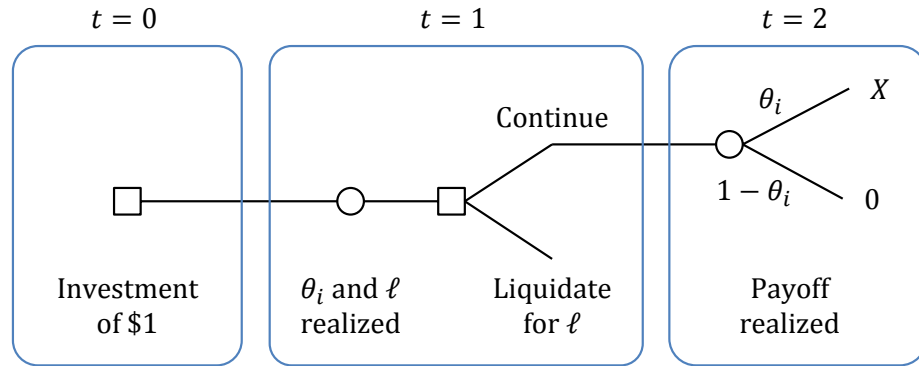
There are three time periods $t = 0, 1, 2$ and all agents are risk neutral with a discount rate equal to the risk-free rate of zero. There is a continuum of identical banks $i \in [0, 1]$, each with the opportunity to invest in a project.

Project Bank i 's project requires an investment of 1 in the initial period $t = 0$ and has a random payoff in the final period $t = 2$ given by

$$\begin{cases} X & \text{with probability } \theta_i, \\ 0 & \text{with probability } 1 - \theta_i. \end{cases}$$

In the interim period $t = 1$, the project can still be abandoned and any fraction of its assets can be sold off to alternative uses at a liquidation value of $\ell < 1$. At the time of investment in $t = 0$, there is uncertainty about both the project's expected payoff $\theta_i X$ and the liquidation value ℓ , which is not resolved until additional information becomes available in the interim period $t = 1$. The structure of bank i 's project and its time-line is illustrated in Figure 1.1.

Importantly, in $t = 1$ the liquidation value ℓ is not directly linked to the expected payoff $\theta_i X$ of bank i 's project. In this model, liquidating a project entails taking the

Figure 1.1: Project time-line for bank i

assets out of their current use and selling them to be used for a different purpose – an actual reallocation of capital. Therefore, liquidation is not inherently inefficient: Efficiency requires that a project be abandoned and that its assets be liquidated whenever the expected payoff turns out to be less than the liquidation value and vice versa:

$$\theta_i X \leq \ell \quad \Rightarrow \quad \text{abandon}$$

$$\theta_i X > \ell \quad \Rightarrow \quad \text{continue}$$

Incentive Problem A debt-financed bank faces a basic incentive problem when it comes to continuing or liquidating its project which is similar to the risk-shifting problem of [Jensen and Meckling \(1976\)](#). Suppose that in the initial period $t = 0$ a bank has $\eta \in [0, 1]$ of equity and raises $1 - \eta$ in some form of debt. Denote by D_t the face value of this debt at $t = 1, 2$. After learning about θ_i and ℓ in $t = 1$, the bank wants to continue its project whenever the expected equity payoff from

continuing is greater than the equity payoff from liquidating:

$$\begin{aligned} & \theta_i (X - (1 - \eta) D_2) > \max \{0, \ell - (1 - \eta) D_1\} \\ \Leftrightarrow \theta_i > & \begin{cases} 0 & \text{for } 1 - \eta \leq \frac{\ell}{D_1} \\ \frac{\ell - (1 - \eta) D_1}{X - (1 - \eta) D_2} & \text{for } 1 - \eta > \frac{\ell}{D_1} \end{cases} \end{aligned}$$

Unless the bank is fully equity financed ($\eta = 1$), its decision doesn't correspond to the efficient one of continuing whenever $\theta_i X > \ell \Leftrightarrow \theta_i > \ell/X$. In particular, as long as $D_1 X > D_2 \ell$, i.e. X sufficiently larger than ℓ , the bank wants to take excessive risks in the interim period by continuing projects with negative net present value. Since this incentive problem is present for any $\eta < 1$ I consider the worst case and assume that banks have no initial equity.⁵

Uncertainty There are two possible aggregate states $s \in \{H, L\}$ in the interim period $t = 1$, with probabilities p and $1 - p$ for the high state and the low state, respectively. Conditional on the aggregate state s , the banks' success probabilities $\{\theta_i\}$ are i.i.d. with cumulative distribution function F_s on $[0, 1]$. The difference between the high state and the low state is that the distribution F_H strictly dominates the distribution F_L in terms of first-order stochastic dominance:

$$F_H(\theta) < F_L(\theta) \quad \text{for all } \theta \in (0, 1)$$

This means that higher success probabilities are more likely in state H than in state L and therefore that banks' projects are positively correlated through their success probabilities $\{\theta_i\}$.

⁵This assumption abstracts from the choice of leverage to focus on the choice of maturity structure. Section 1.7 discusses implications of allowing for equity financing.

Both the aggregate state s and each bank's success probability θ_i are realized at the beginning of the interim period $t = 1$, before the continuation decision about the project, but after the investment decision in $t = 0$.

Liquidation Value The liquidation value for the banks' assets is determined endogenously from a downward-sloping aggregate demand for liquidated assets. The assets are reallocated to a secondary sector of the economy where they are employed with decreasing marginal productivity. For a total mass $\phi \in [0, 1]$ of assets sold off by all banks, the liquidation value $\ell(\phi)$ is given by a continuous and strictly decreasing function $\ell: [0, 1] \rightarrow [0, 1]$ which corresponds to the assets' marginal product in the secondary sector. Due to the exogenous correlation in banks' θ_i s there are fluctuations in equilibrium asset sales ϕ across the aggregate states H and L which implies volatility in the endogenous liquidation value with two different liquidation values $\ell_H = \ell(\phi_H)$ and $\ell_L = \ell(\phi_L)$ in the two states.

Given the model setup so far, each bank i is exposed to two sources of risk. It faces aggregate risk in terms of the state s which determines the distribution F_s as well as the liquidation value ℓ_s and it faces idiosyncratic risk in terms of its success probability θ_i drawn from F_s . The first source of risk is "aggregate" in the sense that its outcome affects all banks in the same way, the second source is "idiosyncratic" in the sense that its outcome affects only the particular bank itself.

Financing Each bank has to raise the entire investment amount of 1 through loans from competitive investors in $t = 0$. A bank can choose any combination of long-term debt and short-term debt to finance its project.⁶ Long-term debt matures

⁶Allowing only for short-term and long-term debt is essentially equivalent to assuming that θ_i and ℓ are not contractible. Section 1.7 discusses the implications of allowing more state-contingent contracting.

in the final period $t = 2$ at a face value of D_i^{LT} . Short term debt, on the other hand, has to be rolled over in the interim period $t = 1$ and, if rolled over, has a face value of D_i^{ST} in the final period $t = 2$. Instead of rolling over in $t = 1$ a short-term creditor has the right to withdraw the principal of his loan.⁷ This creates the possibility of the bank becoming illiquid in $t = 1$ since it may face more withdrawals from short-term creditors than it can satisfy by liquidating even the entire project.

Denoting by $\alpha_i \in [0, 1]$ the fraction of bank i 's project financed by short-term debt, the bank's choice of debt maturity structure in the initial period $t = 0$ amounts to a combination of short-term debt and long-term debt $(\alpha_i, 1 - \alpha_i)$. The face values D_i^{ST} and D_i^{LT} are then determined endogenously, taking into account both the idiosyncratic and aggregate risk, as well as the the rollover risk arising from the bank's maturity structure.

Competitive Equilibrium A competitive equilibrium consists of a maturity structure α_i for every bank i and liquidation values ℓ_H and ℓ_L in the two aggregate states such that (i) each bank chooses its maturity structure optimally given the equilibrium liquidation values, and (ii) the liquidation values result from the asset sales induced by the equilibrium maturity structure choices.

To reduce notational clutter I will drop the bank index i in the following sections that deal only with an individual bank.

⁷The assumption that the interim face value equals the principal is just a normalization and without loss of generality.

1.3 Endogenous Rollover Risk

Denoting the fraction of a bank's short-term creditors who withdraw their loans in $t = 1$ by λ , the bank has to liquidate enough of the project to raise $\alpha\lambda$ for repayment. Since the bank can raise at most ℓ by liquidating the entire project, it *can* become illiquid if $\alpha > \ell$ and it *will* be illiquid whenever⁸

$$\lambda > \frac{\ell}{\alpha}.$$

If the bank becomes illiquid in $t = 1$, there will be nothing left in $t = 2$ to repay the long-term creditors and, more importantly, any short-term creditors who decided to roll over their loan. The short-term creditors therefore face a coordination problem which I model as a global game by assuming a small amount of noise in each creditor's information. The bank's rollover risk is then derived from the equilibrium of the creditors' coordination game.

The short-term debt in this model is meant to represent market-based financing such as commercial paper. In these markets, most of the funds are allocated through intermediaries, e.g. money market funds in the commercial paper market. It turns out to be much more tractable to assume that the roll-over decision is taken by a fund manager on behalf of the actual investor.⁹ There is a continuum of fund managers with the following payoffs. If a fund manager withdraws his loan in $t = 1$ he receives a constant payoff of $w > 0$, a base salary. If the fund manager rolls over his loan in $t = 1$ the payoff depends on whether the bank repays the loan in $t = 2$: if

⁸The paper focuses on the case of $\alpha > \ell$ where the bank can become illiquid. Appendix 1.9.3 discusses the case of $\alpha \leq \ell$ and presents sufficient conditions for the optimal maturity structure to satisfy $\alpha > \ell$.

⁹The assumption of intermediation in the supply of short-term funding is similar to [Rochet and Vives \(2004\)](#) but doesn't affect the validity of the main results. Appendix 1.9.2 presents the model without fund managers.

the bank repays, the fund manager receives a payoff of bw , his base salary multiplied by a bonus factor $b > 1$; if the bank doesn't repay, the fund manager receives a payoff of zero.¹⁰ A higher b corresponds to higher-powered incentive structures and implies that the fund managers are willing to take greater risks. It can also be interpreted as a proxy for the risk tolerance of the short-term funding sector in general.¹¹

Each fund manager has to make his roll-over decision in the interim period $t = 1$ based on the following information. While the resolution of aggregate risk in the form of the liquidation value ℓ is perfectly observed by everyone and becomes common knowledge, the resolution of idiosyncratic risk in the form of the bank's success probability θ is not perfectly observed. Instead, each fund manager j receives a noisy signal $\tilde{\theta}_j = \theta + \varepsilon_j$, where the signal noise ε_j is i.i.d. uniformly on $[-\varepsilon, \varepsilon]$ for some arbitrarily small $\varepsilon > 0$.

Since the bank can become illiquid should too many loans be withdrawn, each fund manager's expected payoff of rolling over depends critically on the fraction λ of other fund managers who withdraw. Given that he only receives the bonus bw if the bank remains liquid in $t = 1$ and the project succeeds in $t = 2$, the expected payoff of rolling over is

$$\begin{aligned} & \Pr[\text{liquid} \mid \ell, \tilde{\theta}_j] \cdot \Pr[\text{success} \mid \ell, \tilde{\theta}_j] \cdot bw \\ &= \Pr\left[\lambda \leq \frac{\ell}{\alpha} \mid \ell, \tilde{\theta}_j\right] \cdot E[\theta \mid \ell, \tilde{\theta}_j] \cdot bw, \end{aligned}$$

while the payoff of withdrawing is w for sure.

¹⁰These simple payoffs are chosen in order to economize on exogenous parameters appearing in the model's expressions. More complicated payoffs, e.g. a withdrawal payoff also depending on the rollover outcome can easily be accommodated.

¹¹See [Krishnamurthy \(2010\)](#) for a discussion on the importance of lenders' risk tolerance in the short-term funding markets.

Using global game techniques I can derive the *unique* equilibrium of the fund managers' coordination game.¹² The equilibrium is symmetric in switching strategies around a signal threshold $\hat{\theta}$ such that each fund manager withdraws for all signals below the threshold and rolls over for all signals above. The equilibrium switching point $\hat{\theta}$ is determined by the fact that a fund manager exactly at the switching point has to be indifferent between rolling over and withdrawing, given his belief about the fraction λ of others withdrawing. Taking the limit as the signal noise ε goes to zero, the distribution of λ conditional on being at the switching point $\hat{\theta}$ becomes uniform on $[0, 1]$. The indifference condition for a fund manager at the switching point therefore simplifies to

$$\frac{\ell}{\alpha} \cdot \hat{\theta} \cdot bw = w,$$

which pins down the equilibrium switching point as

$$\hat{\theta} = \frac{\alpha}{\ell b}. \tag{1.1}$$

Proposition 1.1 *For $\varepsilon \rightarrow 0$, the unique equilibrium among short-term creditors is in switching strategies around the success probability threshold $\hat{\theta} = \frac{\alpha}{\ell b}$:*

- *For realizations of θ below $\hat{\theta}$, all short-term debt is withdrawn and the bank becomes illiquid.*
- *For realizations of θ above $\hat{\theta}$ all short-term debt is rolled over and the bank remains liquid.*

¹²See Appendix 1.9.1 for the details of this global game. For a comprehensive discussion of the use of global games since the seminal papers of Carlsson and van Damme (1993a,b) see Morris and Shin (2003).

The simple structure of the equilibrium highlights the key determinants of a bank's rollover risk before the uncertainty about θ and ℓ is resolved. This ex-ante rollover risk, i.e. the probability that the bank will suffer a run in the interim period is given by

$$\Pr\left[\theta\ell < \frac{\alpha}{b}\right].$$

First, the rollover risk is increasing in the fraction of short-term debt α . Having a balance sheet that relies more heavily on short-term debt makes the bank more vulnerable to runs since it increases the total amount of withdrawals the bank may face. By choosing its debt maturity structure, the bank can therefore directly influence its ex-ante rollover risk.

Second, once the maturity structure is in place, whether the bank suffers a run or not depends on both sources of risk, idiosyncratic and aggregate. A run can be triggered by bad news about the project's expected payoff (low θ), or by bad news about the liquidation value for the project's assets (low ℓ). When deciding whether to roll over, creditors (or their fund managers) are worried about a low θ because it means they will less likely be repaid (or receive their bonus) in the final period. In addition, they are worried about a low ℓ because it means the bank is more likely to become illiquid in the interim period. The first corresponds to a fundamentals-based run while the second corresponds to a market-based run. These two effects are very similar to the ones derived by [He and Xiong \(2010\)](#).¹³

Third, the two sources of risk interact in determining the bank's rollover risk. In particular, the bank is more vulnerable to idiosyncratic risk for a low realization of the liquidation value. The destabilizing effect of a low liquidation value means that the bank suffers runs for idiosyncratic news that would have left it unharmed

¹³The distinction is also similar to the concepts of “funding liquidity” and “market liquidity” in [Brunnermeier and Pedersen \(2009\)](#).

had the liquidation value been higher. If the liquidation value fluctuates with the aggregate state, a bank will be more vulnerable to runs in the low aggregate state than in the high aggregate state, for any ex-ante maturity structure. This effect will play a crucial role in the inefficiency result of this paper.

1.4 Equilibrium without Aggregate Risk

I first analyze the model for the benchmark case without aggregate risk when the distribution of success probabilities is the same across states, $F_H = F_L =: F$. In this case each bank is able to maximize its project's net present value and the competitive equilibrium achieves full efficiency.

1.4.1 Optimal Maturity Structure

In the initial period $t = 0$, short-term and long-term creditors and the bank anticipate what will happen in the following periods. This means that the face values of short-term debt and long-term debt, D^{ST} and D^{LT} , have to guarantee that investors break even. The bank, when choosing its debt maturity structure $(\alpha, 1 - \alpha)$, takes into account the effect of α on the face values D^{ST} and D^{LT} , as well as the effect of α on the creditor coordination in $t = 1$.

Without aggregate risk, the liquidation value ℓ for the bank's assets in $t = 1$ is deterministic. The only uncertainty stems from the project's payoff and this uncertainty is partially resolved in $t = 1$ when the success probability θ is drawn from its distribution F . Depending on the additional information received about the project's expected payoff, it will be efficient to either continue with the project or to abandon it and put the liquidated assets to alternative use. Liquidation is

efficient whenever the project's expected payoff is less than the liquidation value, $\theta X \leq \ell$.

To set up the bank's maximization problem it is instructive to first derive the endogenous face values D^{ST} and D^{LT} . Since the liquidation value is deterministic, so is the threshold determining the outcome of the creditor coordination in $t = 1$:

$$\hat{\theta} = \frac{\alpha}{\ell b}$$

For realizations of θ below $\hat{\theta}$, there will be a creditor run on the bank. In this case, each short-term creditor receives an equal share of the liquidation proceeds, ℓ/α , while long-term creditors don't receive anything. For realizations of θ above $\hat{\theta}$, all short-term creditors roll over their loans and the bank continues to operate the project. In this case, all creditors receive the face value of their loan in $t = 2$ if the project is successful. Note that we are now dealing with the payoffs of the actual investors whose money is at stake, not the payoffs of the fund managers.¹⁴

For a short-term creditor this implies an ex-ante expected payoff given by

$$F(\hat{\theta}) \frac{\ell}{\alpha} + \int_{\hat{\theta}}^1 \theta D^{ST} dF(\theta).$$

With probability $F(\hat{\theta})$ the realization of the success probability is $\theta \in [0, \hat{\theta}]$ and there is a run on the bank leading to full liquidation; in this case the short-term creditor receives an equal share ℓ/α of the liquidation value. Otherwise the realization of the success probability is $\theta \in (\hat{\theta}, 1]$ and there is no run on the bank; in this

¹⁴It is natural to assume that the payments to the fund manager, the bonus bw and base salary w , have to be paid by the investor. For simplicity I assume that these payments are negligible as a fraction of total investment and focus on the limiting case of $w \rightarrow 0$ but holding b constant.

case the short-term creditor receives the face value of his loan D^{ST} if the project is successful which happens with probability θ .

A long-term creditor, on the other hand, only receives a payment if (i) there is no run in the interim period and (ii) the project is successful in the final period. The ex-ante expected payoff of a long-term creditor therefore is

$$\int_{\hat{\theta}}^1 \theta D^{LT} dF(\theta).$$

Since all creditors have to break even at the risk-free rate of zero, their expected payoff has to equal their investment of 1 so the endogenous face values for short-term debt and long-term debt are given by

$$D^{ST} = \frac{1 - F(\hat{\theta})\frac{\ell}{\alpha}}{\int_{\hat{\theta}}^1 \theta dF(\theta)} \quad \text{and} \quad D^{LT} = \frac{1}{\int_{\hat{\theta}}^1 \theta dF(\theta)}. \quad (1.2)$$

Due to the effective seniority of short-term debt in the interim period, the face values satisfy $D^{ST} < D^{LT}$, i.e. the interest rate on short-term debt is lower than the interest rate on long-term debt – an upward sloping yield curve.

Given the rollover risk and the face values for a given maturity structure $(\alpha, 1 - \alpha)$, it remains to derive the bank's ex-ante payoff. For realizations $\theta \leq \hat{\theta}$ there is a run by short-term creditors in the interim period and the bank's payoff is zero. For realizations $\theta > \hat{\theta}$ there is no run in $t = 1$ and with probability θ the project is successful in $t = 2$. In this case the bank receives the project's cash flow X and has to repay the face value of its liabilities $\alpha D^{ST} + (1 - \alpha) D^{LT}$. The ex-ante expected payoff of the bank therefore is

$$\int_{\hat{\theta}}^1 \theta [X - \alpha D^{ST} - (1 - \alpha) D^{LT}] dF(\theta).$$

Substituting in the face values from (1.2) and rearranging, the bank's ex-ante expected payoff becomes

$$F(\hat{\theta})\ell + \int_{\hat{\theta}}^1 \theta X dF(\theta) - 1. \quad (1.3)$$

Due to the rational expectations and the competitive creditors, the bank receives the entire economic surplus of its investment opportunity, given the rollover-risk threshold $\hat{\theta}$. The first term in (1.3) is the economic value realized in the states where the project is liquidated. The second term is the expected economic value realized in the states where the project is continued. The third term is the initial cost of investment. Since it receives the entire economic surplus, the bank fully internalizes the effect of its maturity structure choice on the efficiency of the rollover outcome.

Recalling the expression for $\hat{\theta}$ from the creditor coordination game in (1.1), the bank chooses α to solve the following problem:

$$\max \left\{ F(\hat{\theta})\ell + \int_{\hat{\theta}}^1 \theta X dF(\theta) - 1 \right\} \quad \text{subject to} \quad \hat{\theta} = \frac{\alpha}{\ell b}$$

In choosing its maturity structure α , the bank effectively chooses a rollover-risk threshold $\hat{\theta}$ since the creditor coordination results in a one-to-one mapping from maturity structure to rollover risk. The first order condition to the bank's problem is

$$f(\hat{\theta}) \frac{1}{\ell b} (\ell - \hat{\theta}X) = 0,$$

which implies the following result.

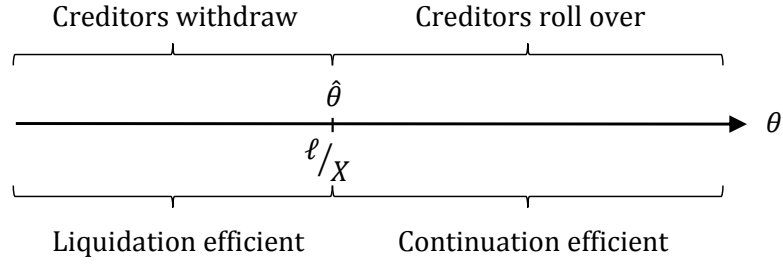


Figure 1.2: Implemented and efficient liquidation policy without aggregate risk

Proposition 1.2 *Without aggregate risk, a bank chooses its maturity structure to implement the efficient liquidation policy:*

$$\alpha^* = \frac{\ell^2 b}{X} \quad \text{resulting in} \quad \hat{\theta}^* = \frac{\ell}{X}.$$

The bank uses short-term debt as a disciplining device to implement a liquidation threshold $\hat{\theta}$ maximizing its payoff. Optimally exposing itself to rollover risk allows the bank to fully eliminate its incentive problem and maximize the project's economic surplus. The perfect match between implemented and efficient liquidation policy is illustrated in Figure 1.2. Efficiency requires that projects with success probabilities $\theta \leq \ell/X$ be liquidated and that projects with success probabilities $\theta > \ell/X$ be continued. Since the bank has full control over its rollover-risk threshold $\hat{\theta}$ it chooses a maturity structure so that creditors withdraw and force liquidation for $\theta \leq \ell/X$ and that they roll over and allow continuation for $\theta > \ell/X$.

This result has important implications for the comparative statics of the bank's rollover risk. While the rollover-risk threshold $\hat{\theta}$ for a given maturity structure α is decreasing in the liquidation value ℓ , the efficient liquidation threshold ℓ/X is increasing in the liquidation value ℓ . As discussed in Section 1.3 above, for a given maturity structure, a higher liquidation value has a stabilizing effect on the

bank and therefore reduces rollover risk. In terms of efficiency, however, a higher liquidation value means that there are better alternative uses for the project's assets which raises the bar in terms of expected project payoff to justify continuing. Since the bank is able to implement the optimal liquidation policy, a higher liquidation value will cause it to increase rollover risk by choosing a maturity structure more reliant on short-term debt. This is reflected in the fact that α^* is increasing in ℓ .

1.4.2 Competitive Equilibrium

In the previous section I derive a bank's optimal maturity structure, taking the equilibrium liquidation value as given. In this section I derive the competitive equilibrium with an endogenous liquidation value. Making the dependence of the liquidation value on the aggregate asset sales ϕ explicit, the optimization of the previous section results in

$$\alpha^*(\phi) = \frac{\ell(\phi)^2 b}{X} \quad \text{and} \quad \hat{\theta}^*(\phi) = \frac{\ell(\phi)}{X}$$

Since all banks are identical ex ante, the competitive equilibrium is symmetric with $\alpha_i^* = \alpha_j^*$ and $\hat{\theta}_i^* = \hat{\theta}_j^*$ for all i, j .

With a continuum of banks $i \in [0, 1]$ and the success probabilities $\{\theta_i\}$ i.i.d. with distribution F , the total mass ϕ of assets sold off in $t = 1$ is equal to the fraction of banks with $\theta_i \leq \hat{\theta}^*(\phi)$ who experience a run by their short-term creditors and have to liquidate their assets. The competitive equilibrium value ϕ^{CE} is therefore the solution to the fixed point equation

$$\phi^{CE} = F\left(\hat{\theta}^*(\phi^{CE})\right). \tag{1.4}$$

This is a fixed point condition for a continuous function mapping the unit interval onto itself so by Brouwer's fixed point theorem there exists a solution. Since the right-hand side of the condition (1.4) is decreasing in ϕ , the fixed point is unique.

Proposition 1.3 *The competitive equilibrium without aggregate risk is characterized by a mass ϕ^{CE} of assets liquidated, implicitly defined by*

$$\phi^{CE} = F\left(\frac{\ell(\phi^{CE})}{X}\right),$$

as well as optimal maturity structures $\{\alpha_i^{CE}\}$ and resulting liquidation thresholds $\{\hat{\theta}_i^{CE}\}$ given by

$$\alpha_i^{CE} = \frac{\ell(\phi^{CE})^2 b}{X} \quad \text{and} \quad \hat{\theta}_i^{CE} = \frac{\ell(\phi^{CE})}{X} \quad \text{for all } i \in [0, 1].$$

Note that the competitive equilibrium is efficient since it equalizes the marginal productivity of assets used in the banking sector to the marginal productivity of assets used in the secondary sector:

$$\hat{\theta}^{CE} X = \ell(\phi^{CE})$$

This efficiency breaks down in the case with aggregate risk discussed next.

1.5 Equilibrium with Aggregate Risk

I now analyze the model with aggregate risk. The state is either high, $s = H$ with probability p , in which case each project's success probability is drawn from the distribution F_H or the state is low, $s = L$, with distribution F_L . Banks are no longer

able to implement the efficient liquidation policy, as a two-sided inefficiency arises: Negative NPV projects are continued in the high state and positive NPV projects are liquidated in the low state.

1.5.1 Optimal Maturity Structure

The additional source of risk with the resulting uncertainty in liquidation values has two main implications from the point of view of an individual bank. The first implication is that the optimal project continuation decision is affected by the realization of ℓ . While in the case without aggregate risk there was a single critical value for the project's expected payoff, there are now two. For the low liquidation value ℓ_L the project should only be continued if $\theta X > \ell_L$, while for the high liquidation value ℓ_H the condition is $\theta X > \ell_H$. In particular, for realizations of the project's success probability θ in the interval $[\ell_L/X, \ell_H/X]$, efficiency calls for liquidation if the assets have a high liquidation value and for continuation if the assets have a low liquidation value.

The second implication of aggregate risk is that the creditor coordination game is different depending on the aggregate state. There are now two equilibrium switching points, $\hat{\theta}_H$ and $\hat{\theta}_L$, one for each realization of ℓ :

$$\hat{\theta}_H = \frac{\alpha}{\ell_H b} \quad \text{and} \quad \hat{\theta}_L = \frac{\alpha}{\ell_L b}$$

If the liquidation value turns out to be high, each creditor is less concerned about the other creditors withdrawing their loans and therefore more willing to roll over his own loan than when the liquidation value turns out to be low. Therefore the bank will be more stable and less likely to suffer a run by its short-term creditors if the liquidation value is high, which is reflected in the rollover-risk threshold being

lower:

$$\hat{\theta}_H < \hat{\theta}_L$$

As in the case without aggregate risk, the bank receives the entire economic surplus of its project, given the liquidation resulting from its maturity structure.

The bank therefore chooses α to solve the following problem

$$\begin{aligned} \max \left\{ p \left(F_H(\hat{\theta}_H)\ell_H + \int_{\hat{\theta}_H}^1 \theta X dF_H(\theta) \right) + (1-p) \left(F_L(\hat{\theta}_L)\ell_L + \int_{\hat{\theta}_L}^1 \theta X dF_L(\theta) \right) \right\} \\ \text{subject to } \hat{\theta}_H = \frac{\alpha}{\ell_H b} \quad \text{and} \quad \hat{\theta}_L = \frac{\alpha}{\ell_L b}, \end{aligned}$$

which gives a first order condition

$$p \left(f_H(\hat{\theta}_H) \frac{1}{\ell_H b} (\ell_H - \hat{\theta}_H X) \right) + (1-p) \left(f_L(\hat{\theta}_L) \frac{1}{\ell_L b} (\ell_L - \hat{\theta}_L X) \right) = 0. \quad (1.5)$$

Although it cannot be solved explicitly for the optimal maturity structure, with $\ell_H > \ell_L$ the first order condition implies $\ell_H - \hat{\theta}_H X > 0$ and $\ell_L - \hat{\theta}_L X < 0$ which gives the following result.

Proposition 1.4 *With aggregate risk and $\ell_H > \ell_L$ a bank chooses its maturity structure resulting in*

$$\hat{\theta}_H < \frac{\ell_H}{X} \quad \text{and} \quad \hat{\theta}_L > \frac{\ell_L}{X}.$$

There is a two-sided inefficiency:

- *For $s = H$, negative-NPV projects are continued whenever $\theta \in (\hat{\theta}_H, \ell_H/X)$.*
- *For $s = L$, positive-NPV projects are liquidated whenever $\theta \in (\ell_L/X, \hat{\theta}_L)$.*

The key effect of aggregate risk is that it drives a wedge between the efficient liquidation policy and the achievable liquidation policy. The effectiveness of using

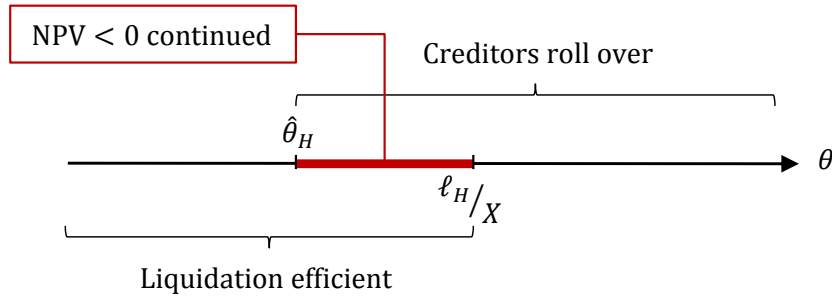


Figure 1.3: Inefficiency in state H

the maturity structure to eliminate the incentive problem and to implement an efficient liquidation policy is undermined when aggregate risk is added to the bank's idiosyncratic risk. It is important to note that there are efficiency losses for both realizations of the liquidation value. When the liquidation value is high, excessively risky projects that should be liquidated because they have negative net present value are continued. This first effect is illustrated in Figure 1.3. When the liquidation value is low on the other hand, valuable projects that should be continued because they have positive net present value are liquidated at fire-sale prices. This second effect is illustrated in Figure 1.4.

The two-sided inefficiency comes from the ambivalent role played by the liquidation value of the bank's assets. A high liquidation value in good aggregate states makes the bank less vulnerable to runs by its short-term creditors but at the same

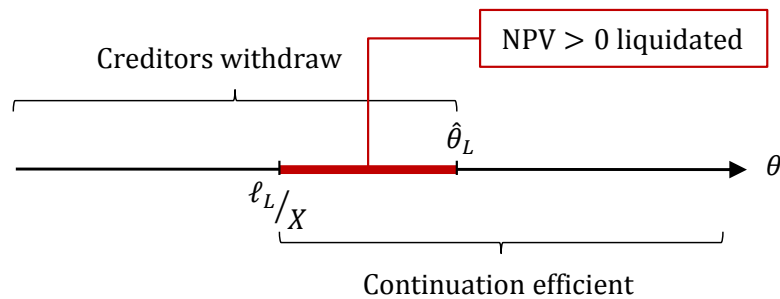


Figure 1.4: Inefficiency in state L

time, the high liquidation value raises the bar in terms of alternate uses for the bank's assets which worsens the incentive problem. Exactly the opposite happens in bad aggregate states where the liquidation value is low. This means that the disciplining effect of short-term debt is weak in the states where it is needed more and is strong in the states where it is needed less.

1.5.2 Competitive Equilibrium

In the competitive equilibrium, each bank $i \in [0, 1]$ chooses its maturity structure optimally according to the first order condition (1.5) of the previous section, taking the liquidation values ℓ_H and ℓ_L as given. As in the case without aggregate risk, the banks' ex-ante symmetry implies that they all choose the same maturity structure. Making the dependence of the liquidation value on the asset sales ϕ explicit I denote the optimal maturity structure from the first order condition (1.5) by $\alpha^*(\phi_H, \phi_L)$. The resulting liquidation thresholds are then given by

$$\hat{\theta}_H(\phi_H, \phi_L) = \frac{\alpha^*(\phi_H, \phi_L)}{\ell(\phi_H)b} \quad \text{and} \quad \hat{\theta}_L(\phi_H, \phi_L) = \frac{\alpha^*(\phi_H, \phi_L)}{\ell(\phi_L)b}.$$

Since the success probabilities $\{\theta_i\}$ are i.i.d. conditional on the aggregate state s , the total mass ϕ_s of assets sold off in state s is equal to the fraction of banks with $\theta_i \leq \hat{\theta}_s(\phi_H, \phi_L)$. As in the case without aggregate risk, the competitive equilibrium is characterized by a fixed point condition for the asset sales (ϕ_H, ϕ_L) , except that it is now two-dimensional:

$$\phi_H^{CE} = F_H\left(\hat{\theta}_H(\phi_H^{CE}, \phi_L^{CE})\right) \quad \text{and} \quad \phi_L^{CE} = F_L\left(\hat{\theta}_L(\phi_H^{CE}, \phi_L^{CE})\right)$$

This is a fixed point condition for a continuous function mapping the unit square onto itself so by Brouwer's fixed point theorem there exists a solution.

Proposition 1.5 *The competitive equilibrium with aggregate risk is characterized by asset sales $(\phi_H^{CE}, \phi_L^{CE})$ implicitly defined by*

$$\phi_H^{CE} = F_H \left(\frac{\alpha^*(\phi_H^{CE}, \phi_L^{CE})}{\ell(\phi_H^{CE}) b} \right) \quad \text{and} \quad \phi_L^{CE} = F_L \left(\frac{\alpha^*(\phi_H^{CE}, \phi_L^{CE})}{\ell(\phi_L^{CE}) b} \right),$$

where $\alpha^*(\phi_H^{CE}, \phi_L^{CE})$ is the optimal maturity structure defined by equation (1.5).

Due to the nesting of endogenous variables with the exogenous functions it is hard to specify general conditions to guarantee that the competitive equilibrium is unique and satisfies $\phi_H^{CE} < \phi_L^{CE}$ and therefore $\ell_H > \ell_L$. For the purposes of this paper I restrict attention to cases where this is true. The functional forms and parameters in the numerical simulation of Section 1.6 show that such cases exist and don't require unreasonable parameter assumptions.

To highlight the equilibrium effect of the two-sided inefficiency it is instructive to compare the competitive equilibrium to the first-best allocation with efficient liquidation thresholds:

$$\hat{\theta}_H^{FB} = \frac{\ell_H}{X} \quad \text{and} \quad \hat{\theta}_L^{FB} = \frac{\ell_L}{X}$$

Using the efficient liquidation thresholds, the first-best allocation is characterized by asset sales $(\phi_H^{FB}, \phi_L^{FB})$ implicitly defined by

$$\phi_H^{FB} = F_H \left(\frac{\ell(\phi_H^{FB})}{X} \right) \quad \text{and} \quad \phi_L^{FB} = F_L \left(\frac{\ell(\phi_L^{FB})}{X} \right).$$

Since by strict first order stochastic dominance $F_H(\theta) < F_L(\theta)$ for any $\theta \in (0, 1)$, the first-best allocation satisfies $\phi_H^{FB} < \phi_L^{FB}$ and therefore $\ell_H^{FB} > \ell_L^{FB}$. This means

that even in the first-best allocation the liquidation values vary across aggregate states.

From Proposition 1.4 we know that $\hat{\theta}_H^{CE} < \hat{\theta}_H^{FB}$ and $\hat{\theta}_L^{CE} > \hat{\theta}_L^{FB}$ which means that compared to the first-best allocation the competitive equilibrium has less liquidation in the high state and more liquidation in the low state. This implies the following result.

Proposition 1.6 *In the competitive equilibrium with aggregate risk and $\ell_H > \ell_L$ the two-sided inefficiency has a self-reinforcing effect by amplifying the volatility in liquidation values:*

$$\ell_H^{CE} > \ell_H^{FB} \quad \text{and} \quad \ell_L^{CE} < \ell_L^{FB}.$$

The two-sided inefficiency originates in the fact that the liquidation values vary across aggregate states which is true even in the first-best allocation. Then the inefficiency causes too little liquidation in the high state and too much liquidation in the low state. This further increases the volatility of liquidation values, reinforcing the inefficiency in a feedback effect.

Nevertheless, it is important to note that the competitive equilibrium is still constrained efficient. Since each bank maximizes the economic surplus of its investment opportunity it has the same objective function as a social planner who is constrained to choosing a debt-maturity structure. The banks fully internalize the effect of their maturity structure when trading off the inefficiencies in the two aggregate states. Therefore, a policy intervention such as a tax on the use of short-term debt would reduce efficiency. It would lead to an increase of the inefficiency due to excessive risk-taking which more than outweighs the reduction of the inefficiency due to excessive liquidation.

1.6 Numerical Simulation

To quantify the two-sided inefficiency of this paper and to illustrate some of the comparative statics I now run a numerical simulation with specific functional forms. I assume a very simple functional form for the distributions of success probabilities in the two states:

$$F_H(\theta) = \begin{cases} 0 & \text{for } \theta < 0 \\ (1 - q)\theta & \text{for } 0 \leq \theta < 1 \\ 1 & \text{for } \theta \geq 1 \end{cases}, \quad F_L(\theta) = \begin{cases} 0 & \text{for } \theta < 0 \\ q + (1 - q)\theta & \text{for } 0 \leq \theta < 1 \\ 1 & \text{for } \theta \geq 1 \end{cases}$$

The distributions F_H and F_L are both uniform on the interval $(0, 1)$ but have probability mass q at one of the endpoints, F_H at $\theta = 1$ and F_L at $\theta = 0$.

This boils down the nature of aggregate risk to two parameters, the probability p of the high state and the probability mass q in the state-contingent distributions of success probabilities. The two parameters have straightforward interpretations. The probability p is a proxy for the negative skew of aggregate risk since for higher values of p the low aggregate state is less likely. The probability mass q is a proxy for the correlation of the banks' projects.¹⁵ I assume that the liquidation value is given by $\ell(\phi) = 1 - \phi$ and that the payoff of a successful project in the final period $t = 2$ is given by $X = 2$.

¹⁵The actual correlation coefficient between any two success probabilities θ_i, θ_j is given by $\rho = 3q^2 / (4q(3p - 1) + 1)$ which is strictly increasing in q for all $p, q \in [0, 1]$.

Without Aggregate Risk In the case without aggregate risk, i.e. $q = 0$, the equation characterizing the competitive equilibrium simplifies to¹⁶

$$\phi^{CE} = F\left(\frac{\ell(\phi^{CE})}{X}\right) = \frac{1 - \phi^{CE}}{2},$$

which implies $\phi^{CE} = \hat{\theta}^{CE} = \frac{1}{3}$ and an equilibrium liquidation value of $\ell^{CE} = \frac{2}{3}$. The expected profit of an individual bank given these values is

$$F(\hat{\theta}^{CE})\ell^{CE} + \int_{\hat{\theta}^{CE}}^1 \theta X dF(\theta) - 1 = \frac{1}{9},$$

which corresponds to a return of 11.1% on the initial investment of 1.

With Aggregate Risk In the case with aggregate risk, i.e. $q > 0$, the optimal maturity structure for given liquidation values ℓ_H, ℓ_L implies liquidation thresholds

$$\hat{\theta}_H = \frac{1}{2} \frac{\ell_H \ell_L^2}{p \ell_L^2 + (1-p) \ell_H^2} \quad \text{and} \quad \hat{\theta}_L = \frac{1}{2} \frac{\ell_H^2 \ell_L}{p \ell_L^2 + (1-p) \ell_H^2}.$$

Substituting these into the fixed point condition characterizing the competitive equilibrium and using the functional forms for F_H, F_L and ℓ yields

$$\begin{aligned} \phi_H &= (1-q) \frac{1}{2} \frac{(1-\phi_H)(1-\phi_L)^2}{p(1-\phi_L)^2 + (1-p)(1-\phi_H)^2} \\ \text{and } \phi_L &= q + (1-q) \frac{1}{2} \frac{(1-\phi_H)^2(1-\phi_L)}{p(1-\phi_L)^2 + (1-p)(1-\phi_H)^2}. \end{aligned} \quad (1.6)$$

I computationally derive the competitive equilibrium for five different values each of p and q . The probability p of the high state varies from 0.5 to 0.9 and the probability

¹⁶Note that for $q = 0$, there is no difference between the states H and L so the probability p doesn't matter.

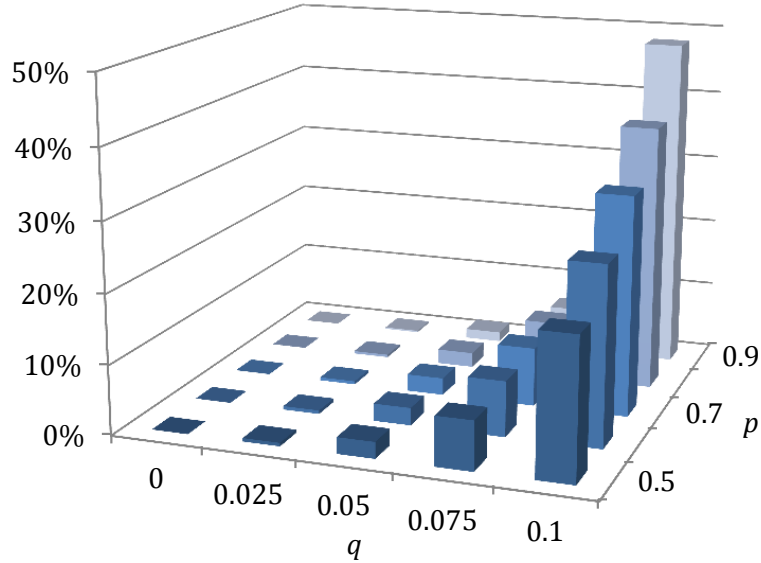


Figure 1.5: Percentage of ex-ante surplus destroyed by the inefficiency

mass q varies from 0 to 0.1.¹⁷ The fixed point condition (1.6) has a unique solution (ϕ_H, ϕ_L) with $\phi_H < \phi_L$ for each of the combinations (p, q) I consider, with ϕ_H ranging from 0.25 to 0.33 and ϕ_L ranging from 0.33 to 0.54.

Figure 1.5 shows the impact of the two-sided inefficiency on economic surplus for the different combinations of p and q . The figure displays the percentage of expected economic surplus lost in the competitive equilibrium relative to the first-best allocation. We see that the inefficiency cost is exponentially increasing in the correlation of projects as captured by q . In addition, the effect is strongest if aggregate risk is negatively skewed, i.e. the probability p of the good state is high and the low state is unlikely to occur. In the worst case almost 50% of ex-ante economic surplus is lost due to the inefficient liquidation policy.

The amplification effect of the two-sided inefficiency is illustrated in Figure 1.6 for the intermediate case of $p = 0.7$ and $q = 0.05$. The figure starts at the liquidation values in the first-best allocation which are $\ell_H^{FB} = 0.678$ and $\ell_L^{FB} = 0.644$. It then

¹⁷These values for p and q span a range of correlation coefficients between two success probabilities θ_i, θ_j from 0% to 2.5%.

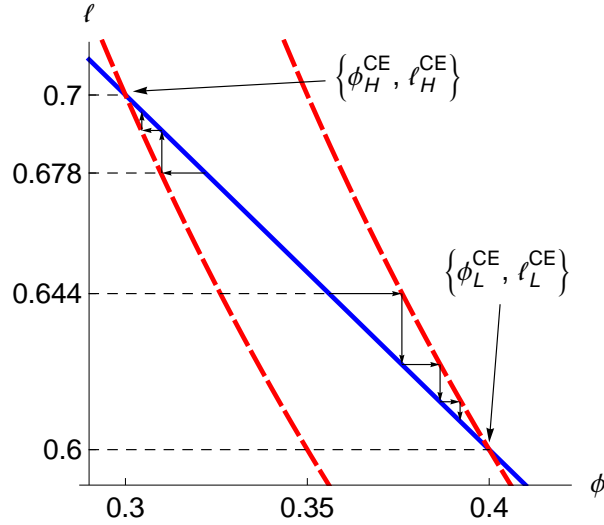


Figure 1.6: Amplification effect for $p = 0.7$ and $q = 0.05$

iterates between (i) the dashed curves, representing asset sales implied by the banks' optimal reaction for given liquidation values and (ii) the solid curve, representing liquidation values implied by given asset sales. The iteration ends at the competitive equilibrium, where $\ell_H^{CE} = 0.700$ and $\ell_L^{CE} = 0.600$. Due to the amplification, the standard deviation of liquidation values increases by a factor of three from 0.015 in the first-best allocation to 0.046 in the competitive equilibrium while the mean changes by only 0.002. We see that the two-sided inefficiency has a strong self-reinforcing effect, significantly amplifying the magnitude of aggregate risk faced by each bank.

1.7 Discussion

For purposes of exposition this paper presents a very stylized model. The main results, however, are robust. The first building block of the model is that the incentive problem of a bank's equity holders is worse when asset liquidity is high. This applies whenever there is still an upside possible in the bank's project and this up-

side is greater than the liquidation value. In this case equity holders stand to gain more by keeping the project running instead of liquidating. If the liquidation value reflects, at least in part, the assets' value in alternative uses, it can be higher than the expected payoff in current use which implies that the equity holders' decision is inefficient. This incentive problem is more severe, the higher the liquidation value, i.e. the more valuable the assets are in alternative uses.

The second building block is the fact that a bank's rollover risk is decreasing in its asset liquidity. This is a very basic comparative static with a strong intuition: When creditors decide whether to roll over their loans in a situation where illiquidity is possible, their decision will depend on how vulnerable the bank is. The key factor determining the bank's vulnerability is how many withdrawals it can satisfy given the liquidation value of its assets before it runs out of funds. Therefore higher asset liquidity means less jittery creditors means lower rollover risk. If asset liquidity varies across aggregate states, so will the rollover risk a bank faces.

If the liquidation value is deterministic as in the case without aggregate risk, a bank can expose itself to exactly so much rollover risk as to implement the optimal liquidation policy. However, if the liquidation value is random as in the case with aggregate risk, then the incentive problem is worse when the disciplining device is weaker and vice versa, causing the two-sided inefficiency.

Two potential solutions to address the inefficiency problem come to mind. The first potential solution is equity financing. As discussed in the paper, a fully equity financed bank will not face an incentive problem so it doesn't fall victim to the inefficiency. In addition, if all banks were fully equity financed, the amplification effect would be eliminated and the volatility of liquidation values would be much lower. I want to highlight that besides the usual argument that equity financing

may not be used because it is more expensive, the amplification effect implies that banks do not internalize the social cost of using debt financing. Therefore, even if I allowed the banks in the model to choose equity financing, their decision would be distorted toward debt financing.

The second potential solution is state-contingent debt. Clearly, if the face value of short-term debt could be made contingent on the liquidation value, the inefficiency problem would be solved. By specifying different face values for different aggregate states, the bank gains an additional degree of freedom in its self-disciplining device. It can then tailor its exposure to rollover risk in exactly the right way, as is the case without aggregate risk. However, this will make short-term debt more risky for the creditors which runs against one of the main reasons for short maturities in the supply of credit.¹⁸

The source of aggregate risk in my model is the correlation of banks' projects which is assumed to be exogenous. This raises the question if banks would choose correlated projects if this decision was endogenous.¹⁹ In my model a bank suffers inefficiency costs in both aggregate states for intermediate realizations of its expected project payoff. This creates an incentive to choose projects with probability mass concentrated away from intermediate realizations. Given that other banks are choosing correlated projects, an individual bank would therefore want to choose a project that is either positively or negatively correlated with other banks. While the details of the parameterization will determine which way the decision goes, it

¹⁸In my model creditors are risk neutral and have no particular preference for the maturity of their loans. This focuses my model on the demand for short-term financing.

¹⁹See [Acharya \(2009\)](#) and [Farhi and Tirole \(2010\)](#) for models of banks choosing the correlation of their portfolios with the prospect of a government bailout.

is possible that banks would want to choose positively correlated projects, even absent considerations such as government bailouts.

1.8 Conclusion

In this paper I present a new model of debt-maturity structure choice and highlight an important inefficiency in the use of short-term debt. The benchmark model of banks facing only idiosyncratic risk establishes the mechanism of using the debt-maturity structure to overcome an incentive problem and implement a liquidation policy. By anticipating the coordination problem among short-term creditors, a bank can choose the right amount of rollover risk to maximize its economic surplus. The competitive equilibrium achieves the first-best allocation.

The addition of aggregate risk, however, severely undermines the disciplining mechanism of short-term debt and drives a wedge between desired and achievable liquidation policy. This implies a two-sided inefficiency where in good aggregate states there is excessive risk-taking, while in bad aggregate states there is excessive liquidation. The reason is that the disciplining effect of short-term debt is weakened in the states where it is needed more and strengthened in the states where it is needed less. The competitive equilibrium shows that this inefficiency causes sizable losses of economic surplus and is self-reinforcing with a significant amplification of aggregate volatility.

1.9 Appendix

1.9.1 Global Game

To apply the standard global games results summarized by [Morris and Shin \(2003\)](#) the payoffs have to satisfy certain properties. Using the fund manager payoffs of [Section 1.3](#), the payoff difference between withdrawing and rolling over is:

$$\Delta(\lambda, \theta) = \begin{cases} w - \theta bw & \text{for } \lambda \leq \frac{\ell}{\alpha} \\ w & \text{for } \lambda > \frac{\ell}{\alpha} \end{cases}$$

This payoff difference is monotone in θ (state monotonicity) and there is a unique θ^* that solves $\int_0^1 \Delta(\lambda, \theta) d\lambda = 0$ (strict Laplacian state monotonicity). In terms of limit dominance, for $\theta < 1/b$ we have $\Delta(\lambda, \theta) > 0$ for all λ (lower dominance region). Taking the approach of [Goldstein and Pauzner \(2005\)](#) I assume that for sufficiently high θ the bank cannot become illiquid, e.g. because the project matures early and pays off X for sure. This implies an upper dominance region $(\bar{\theta}, 1]$ such that for $\theta > \bar{\theta}$ we have $\Delta(\lambda, \theta) < 0$ for all λ . The payoff difference $\Delta(\lambda, \theta)$ is not monotone in λ but it satisfies the following single-crossing property: For each θ there exists a $\lambda^* \in \mathbb{R} \cup \{-\infty, +\infty\}$ such that $\Delta(\lambda, \theta) < 0$ for all $\lambda < \lambda^*$ and $\Delta(\lambda, \theta) > 0$ for all $\lambda > \lambda^*$. In addition, the signal about θ with uniform noise satisfies the monotone likelihood ratio property. Given all these properties, there is a unique equilibrium and it is in symmetric switching strategies around a critical value $\hat{\theta}$.²⁰

²⁰See [Lemma 2.3](#) and the following discussion in [Morris and Shin \(2003\)](#) and [Theorem 1](#) in [Goldstein and Pauzner \(2005\)](#) for details.

In equilibrium, a fund manager with signal $\tilde{\theta}_j = \hat{\theta}$ has to be indifferent between rolling over and withdrawing:

$$\Pr \left[\lambda \leq \frac{\ell}{\alpha} \mid \hat{\theta} \right] E[\theta | \hat{\theta}] bw = w$$

Given the signal structure, for a particular realization θ the distribution of signals is uniform on $[\theta - \varepsilon, \theta + \varepsilon]$ and for a particular signal realization $\tilde{\theta}$ the conditional distribution of θ is

$$f(\theta | \tilde{\theta}) = \begin{cases} \frac{f(\theta)}{F(\tilde{\theta} + \varepsilon) - F(\tilde{\theta} - \varepsilon)} & \text{for } \theta \in [\tilde{\theta} - \varepsilon, \tilde{\theta} + \varepsilon], \\ 0 & \text{otherwise.} \end{cases}$$

With the distribution of $\theta | \hat{\theta}$ we have an expression for $E[\theta | \hat{\theta}]$ so it remains to derive the distribution of $\lambda | \hat{\theta}$. We can derive the corresponding c.d.f. $G(\lambda | \hat{\theta})$ as follows: The probability that a fraction less than λ receives a signal less than $\hat{\theta}$ (and therefore withdraws) equals the probability that θ is greater than θ' defined by

$$\begin{aligned} \frac{\hat{\theta} - (\theta' - \varepsilon)}{2\varepsilon} &= \lambda \\ \Rightarrow \theta' &= \hat{\theta} + \varepsilon - 2\varepsilon\lambda \end{aligned}$$

We therefore have

$$\begin{aligned} G(\lambda | \hat{\theta}) &= 1 - F(\hat{\theta} + \varepsilon - 2\varepsilon\lambda | \hat{\theta}) \\ &= 1 - \int_{\hat{\theta} - \varepsilon}^{\hat{\theta} + \varepsilon - 2\varepsilon\lambda} \frac{f(\theta)}{F(\hat{\theta} + \varepsilon) - F(\hat{\theta} - \varepsilon)} d\theta \\ &= \frac{F(\hat{\theta} + \varepsilon) - F(\hat{\theta} + \varepsilon - 2\varepsilon\lambda)}{F(\hat{\theta} + \varepsilon) - F(\hat{\theta} - \varepsilon)}. \end{aligned}$$

Finally we have to derive the limits as the signal noise ε goes to zero. First, we have that $\lim_{\varepsilon \rightarrow 0} E[\theta|\hat{\theta}] = \hat{\theta}$. Second, we have that

$$\begin{aligned}
\lim_{\varepsilon \rightarrow 0} G(\lambda|\hat{\theta}) &= \lim_{\varepsilon \rightarrow 0} \frac{F(\hat{\theta} + \varepsilon) - F(\hat{\theta} + \varepsilon - 2\varepsilon\lambda)}{F(\hat{\theta} + \varepsilon) - F(\hat{\theta} - \varepsilon)} \\
&= \lim_{\varepsilon \rightarrow 0} \frac{f(\hat{\theta} + \varepsilon) - f(\hat{\theta} + \varepsilon - 2\varepsilon\lambda)(1 - 2\lambda)}{f(\hat{\theta} + \varepsilon) + f(\hat{\theta} - \varepsilon)} \quad \text{by l'Hôpital's rule} \\
&= \lim_{\varepsilon \rightarrow 0} \frac{f(\hat{\theta} + \varepsilon) - f(\hat{\theta} + \varepsilon - 2\varepsilon\lambda)}{f(\hat{\theta} + \varepsilon) + f(\hat{\theta} - \varepsilon)} + \lim_{\varepsilon \rightarrow 0} \frac{2\lambda f(\hat{\theta} + \varepsilon - 2\varepsilon\lambda)}{f(\hat{\theta} + \varepsilon) + f(\hat{\theta} - \varepsilon)} \\
&= \frac{0}{2f(\hat{\theta})} + \frac{2\lambda f(\hat{\theta})}{2f(\hat{\theta})} \\
&= \lambda
\end{aligned}$$

So the distribution of λ conditional on being at the switching point becomes uniform as the signal noise goes to zero.

1.9.2 Model without Fund Managers

Instead of using fund manager payoffs, we can work with the real creditor payoffs. If enough short-term creditors roll over and the bank remains liquid, a creditor who rolls over receives an expected payoff of θD^{ST} while a creditor who withdraws receives 1. If too many short-term creditors withdraw and the bank becomes illiquid, a creditor who rolls over receives zero while a creditor who withdraws receives ℓ/α . With the assumption of an upper dominance region, these payoffs satisfy the global game conditions of Appendix 1.9.1 guaranteeing that the equilibrium is unique and in switching strategies.

Indifference at the switching point between rolling over and withdrawing requires

$$\frac{\ell}{\alpha}\theta D^{ST} = \frac{\ell}{\alpha} + \left(1 - \frac{\ell}{\alpha}\right) \frac{\ell}{\alpha}$$

so the critical value is given by

$$\hat{\theta} = \frac{1}{D^{ST}} \left(2 - \frac{\ell}{\alpha}\right). \quad (1.7)$$

As in the case with fund managers, the liquidation threshold from the creditor coordination game is decreasing in ℓ . This implies that for a given maturity structure α and a given face value D^{ST} the bank is more vulnerable to runs for lower liquidation values.

The main difference to the case with fund managers is that the critical value now depends on the face value D^{ST} . Through the ex-ante break-even condition, the face value D^{ST} is endogenous and depends on $\hat{\theta}$:

$$F(\hat{\theta})\frac{\ell}{\alpha} + \int_{\hat{\theta}}^1 \theta D^{ST} dF(\theta) = 1. \quad (1.8)$$

We see that equations (1.7) and (1.8) jointly determine $\hat{\theta}$ and D^{ST} for any given α .

Combining the two equations gives us an implicit definition of $\hat{\theta}$

$$\hat{\theta} \left(1 - F(\hat{\theta})\frac{\ell}{\alpha}\right) - \int_{\hat{\theta}}^1 \theta dF(\theta) \left(2 - \frac{\ell}{\alpha}\right) = 0 \quad (1.9)$$

For $\alpha > \ell$ the left hand side of (1.9) is strictly increasing in $\hat{\theta}$ which implies that there is a one-to-one mapping between α and $\hat{\theta}$.

Now the bank chooses α to solve the following problem:

$$\max \left\{ F(\hat{\theta})\ell + \int_{\hat{\theta}}^1 \theta X dF(\theta) - 1 \right\} \quad \text{subject to} \quad (1.9)$$

As before, the bank maximizes the project's economic surplus subject to a constraint which defines $\hat{\theta}$ as a function of α and the exogenous parameters.

1.9.3 Case $\alpha \leq \ell$

This section considers the case where the mass of short-term creditors is small enough so they cannot cause the bank to fail. This corresponds to values of $\alpha \leq \ell$, such that withdrawals from all short-term creditors can be satisfied in $t = 1$ without liquidating the entire project. I assume that in $t = 2$ short-term debt is senior to long-term debt. As in the main part I start by deriving the endogenous face values and the bank's expected payoff without aggregate risk.

Without rollover risk, the expected payoff to a fund manager from rolling over is θbw regardless of the number of others withdrawing and the payoff to withdrawing is w as before. The critical value for θ is therefore independent of α and given by

$$\tilde{\theta} = \frac{1}{b}.$$

This implies that for realizations $\theta \leq \tilde{\theta}$ the bank has to liquidate a fraction α/ℓ of its assets at a liquidation value of ℓ which raises a total of α . The remaining fraction of assets $1 - \frac{\alpha}{\ell}$ remains in place. For realizations $\theta > \tilde{\theta}$ the bank doesn't face withdrawals. Combining this with the bank's expected payoff for $\alpha > \ell$ derived

in Section 1.4.1 the complete expected payoff of the bank choosing $\alpha \in [0, 1]$ is

$$\begin{cases} F(\tilde{\theta})\alpha + \int_0^{\tilde{\theta}} \left(1 - \frac{\alpha}{\ell}\right) \theta X dF(\theta) + \int_{\tilde{\theta}}^1 \theta X dF(\theta) - 1 & \text{for } \alpha \leq \ell, \\ F(\hat{\theta})\ell + \int_{\hat{\theta}}^1 \theta X dF(\theta) - 1 & \text{for } \alpha > \ell. \end{cases}$$

The payoff is continuous in α since the two expressions are the same for $\alpha = \ell$ but not differentiable at $\alpha = \ell$. It is either monotone or single-peaked. Due to the linearity of the bank's expected payoff for $\alpha \leq \ell$, the optimal solution will be either $\alpha = 0$, $\alpha = \ell$, or we will be in the region $\alpha > \ell$ discussed in the main part of the paper.

To guarantee that the solution falls into the range of $\alpha > \ell$ we have to assume that the derivative of both pieces are positive at $\alpha = \ell$:

$$\begin{cases} F\left(\frac{1}{b}\right) - \frac{X}{\ell} \int_0^{\frac{1}{b}} \theta dF(\theta) > 0 \\ f\left(\frac{1}{b}\right) \frac{1}{\ell b} \left(\ell - \frac{1}{b}X\right) > 0 \end{cases}$$

These conditions involve only exogenous parameters and can be satisfied.

With aggregate risk, the bank's expected payoff is more complicated

$$\begin{cases} p \left[F_H(\tilde{\theta})\alpha + \int_0^{\tilde{\theta}} \left(1 - \frac{\alpha}{\ell_H}\right) \theta X dF_H(\theta) + \int_{\tilde{\theta}}^1 \theta X dF_H(\theta) - 1 \right] \\ + (1-p) \left[F_L(\tilde{\theta})\alpha + \int_0^{\tilde{\theta}} \left(1 - \frac{\alpha}{\ell_L}\right) \theta X dF_L(\theta) + \int_{\tilde{\theta}}^1 \theta X dF_L(\theta) - 1 \right] & \text{for } \alpha < \ell_L \\ p \left[F_H(\tilde{\theta})\alpha + \int_0^{\tilde{\theta}} \left(1 - \frac{\alpha}{\ell_H}\right) \theta X dF_H(\theta) + \int_{\tilde{\theta}}^1 \theta X dF_H(\theta) - 1 \right] \\ + (1-p) \left[F_L(\hat{\theta}_L)\ell_L + \int_{\hat{\theta}_L}^1 \theta X dF_L(\theta) - 1 \right] & \text{for } \alpha \in (\ell_L, \ell_H) \\ p \left[F_H(\hat{\theta}_H)\ell_H + \int_{\hat{\theta}_H}^1 \theta X dF_H(\theta) - 1 \right] \\ + (1-p) \left[F_L(\hat{\theta}_L)\ell_L + \int_{\hat{\theta}_L}^1 \theta X dF_L(\theta) - 1 \right] & \text{for } \alpha > \ell_H \end{cases}$$

Again the payoff is continuous in α but not differentiable at $\alpha = \ell_L, \ell_H$. It is either monotone or single-peaked. For the optimal α to be in the region $\alpha > \ell_H$ we need

$$\begin{cases} p \left[F_H\left(\frac{1}{b}\right) - \frac{1}{\ell_H} X \int_0^{\frac{1}{b}} \theta dF_H(\theta) \right] + (1-p) f_L\left(\frac{\ell_H}{\ell_L b}\right) \frac{1}{\ell_L b} \left(\ell_L - \frac{\ell_H}{\ell_L b} X \right) > 0 \\ p f_H\left(\frac{1}{b}\right) \frac{1}{\ell_H b} \left(\ell_H - \frac{1}{b} X \right) + (1-p) f_L\left(\frac{\ell_H}{\ell_L b}\right) \frac{1}{\ell_L b} \left(\ell_L - \frac{\ell_H}{\ell_L b} X \right) > 0 \end{cases}$$

Again these are conditions involving only exogenous parameters and can be satisfied.

Chapter 2

Anxiety in the Face of Risk

2.1 Introduction

Economists have extensively investigated dynamically inconsistent preferences. The literature has, however, focused on inconsistency of *time* preferences, while neglecting implications for risk preferences. We study a particular case of dynamically inconsistent *risk* preferences.

We define an anxiety-prone decision maker as more risk averse for imminent than for distant risk. As the resolution of uncertainty draws close, such an agent wants to pull back from gambles he previously decided to take, although there is no new information, and despite his beliefs not having changed for any other reason.

Such behavior is the result of dynamically inconsistent preferences with respect to risk trade-offs. This is markedly different from an agent having time-changing risk preferences. There is no intra-personal disagreement about risk preferences (and the price of risky assets, for that matter) for an agent who simply values risks differently at different points in time. It is also distinct from a preference for the

timing of the resolution of uncertainty, as an anxiety prone decision maker violates the axiom in [Kreps and Porteus \(1978\)](#) that assumes temporal consistency.

If an anxiety-prone agent trades in a financial market, in which dates of resolution of uncertainty approach and pass, he will trade excessively. In particular, he will sell risky securities just before information about these securities' payoffs is revealed, and buy back his position after the resolution of the risks. His trading causes a predictable price dip before announcement dates, and price increases in the period the risk gets resolved. Indeed, the empirical literature has found such an anomaly, and discussed it as the 'earnings announcement premium' ([Bernard and Thomas \(1989\)](#)). [Lamont and Frazzini \(2007\)](#) confirm that the selling pressure before the announcement as well as the buy pressure after the event stems from small investors, with large and presumably sophisticated investors taking the other side of the trades. Our theory predicts both of these features.¹

We also predict investor returns associated with such behavior. Overtrading due to anxiety is costly for two reasons. First, trading costs eat up returns even if trading per se does not lead to losses. [Odean \(1999\)](#) famously documents this. Second, anxious investors sell before announcements when prices tend low, and buy back at higher prices after the resolution of uncertainty, thus losing with each round of trading in expectation, even absent trading costs. The sum of transaction costs and systematic trading losses may explain why retail investors shun equity exposure at prices neo-classical theory would predict. This gives rise to the equity premium puzzle. Our theory thus views (i) overtrading (ii) price anomalies around announcements and (iii) the equity risk premium as stemming from a single behavioral distortion – anxiety.

¹We predict overtrading by individual unsophisticated agents at this point, not patterns of aggregate trading volume. Excessive aggregate trading can be explained by overconfidence, which is a prediction we derive later in this paper.

It is natural to expect sophisticated agents to come up with strategies to cope with anxiety. Such strategies involve the delegation of investment decisions, which is otherwise puzzling in light of sub-par performance of money managers (Gruber (1996)). Paying an agent to carry out future decisions according to present preferences is a simple but effective way to solve the dynamic inconsistency with respect to risks. Our theory also suggest a demand for particular fee schedules featured in investment funds and brokerage accounts. For example, an anxiety prone decision maker will prefer to have to pay for – or better yet be denied – immediate information about fund performance, because such information may prompt his future self to trade out of a position deemed reasonable presently. This is particularly true for information about increased risks, as we will explain in the section on overconfidence. The timing of investment decisions will be affected as well. Agents will invest in recent winners and pull out funds from recent losers, as Sirri and Tufano (1998) observe. No learning about fund managers' ability is required to explain this pattern.

As another strategy to cope with anxiety, we present a model of endogenous overconfidence. The desire to confine future behavior to present preferences gives rise to a demand for overconfidence. If exposed to a risky environment, the agent finds it beneficial to have overconfident beliefs in the future, as overconfidence helps counterbalance the anxiety he expects his future self to exhibit. Underestimating the risks, his future self will be more likely to take gambles that are favorable according to the manipulating self's preferences, but not according to the anxious self's preferences. We show that the agent can deceive himself to generate such biased beliefs in an intra-personal strategic communication game between his present and future self, despite the future self being a rational Bayesian updater and being

aware of being deceived by its previous self. A comparative statics analysis confirms the intuition that agents more prone to anxiety are more likely to be overconfident, and that they tend to be overconfident to a greater degree. We thus provide a first micro-foundation of a systematic bias of beliefs that has helped explain many puzzles in financial economics that neo-classical theory has left open, such as seemingly excessive amounts of trade.

Moreover, as a result of overconfidence, an anxiety-prone agent may appear to take excessive risks. In our model, overconfidence arises only in high-risk environments. Therefore, we suggest that excessive risk-taking should feature most prominently in inherently risky domains such as securities trading. We conjecture that features of intra-organizational communication patterns can be explained with our theory. Occupational choices and associated cognitive dissonance are other areas we see fit. [Ben-David, Graham, and Harvey \(2010\)](#) confirm that financial top executives are systematically overconfident (realized market returns are within their 80% confidence intervals only 33% of the time). [Ben-David, Graham, and Harvey \(2007\)](#) show that this overconfidence translates into riskier corporate policy.

The rest of this paper is organized as follows. In [Section 2.1.1](#) we relate our work to previous research on anxiety, both in psychology and economics. We present experimental evidence to support our *assumptions*, as well as a short overview on the literature on overconfidence, a *prediction* of our model. [Section 2.2](#) presents our formal setup. [Section 2.3](#) investigates how an anxiety-prone agent behaves in a stylized financial market. We also discuss implied institutional effects in that section. [Section 2.4](#) presents our model of endogenous overconfidence. We conclude and lay out ideas for future research in [Section 2.5](#). All proofs are relegated to the appendix.

2.1.1 Related Literature

Anxiety

People become ‘anxious’ as they approach risky situations. To measure ‘anxiety’ (in a popular sense of the word), psychologists have investigated physiological, emotional, and cognitive responses to anxiety provoking situations. All of them involve being exposed to risks, and immediacy of the risk is found to be a leading determinant for physiological and behavioral reactions to the risk. To illustrate, [Roth, Breivik, Jørgensen, and Hofmann \(1996\)](#) continues a whole series of psychological studies on anxiety of parachutists as the moment of the jump approaches, as well as during the fall (e.g. [Fenz and Epstein \(1967\)](#), [Fenz and Jones \(1972\)](#)). Self-reported anxiety, heart rate and other measures peak right before the jump in novices.² Experienced jumpers learn to inhibit or control their fear, which helps them to perform better in their risky endeavor. [Paterson and Neufeld \(1987\)](#) also find imminence to be a major determinant of the appraisal of a threat in the laboratory. Objectively observable physiological responses besides heart beat and self-reported anxiety include sweating ([Monat and Lazarus \(1991\)](#)).

[Lo and Repin \(2002\)](#) measure the same physical responses of day traders to anxiety provoking situations. In a follow-up paper, [Lo, Repin, and Steenbarger \(2005\)](#) confirm that traders with stronger emotional response generate lower returns. We will argue in this paper that the response to anxiety in the face of risk includes changes of risk preferences, which cause trading losses. Indeed, [Loewenstein, Weber, Hsee, and Welch \(2001\)](#) list changes of risk preferences as emotional reactions

²Fear of flying seems a more commonly experienced situation. Accident statistics rarely change significantly between the time of ticket purchase and the actual flight. Yet, many passengers get more anxious as take-off is imminent. Introspection suggests that the run-up to an academic talk or other forms of public speaking, or performing music, trigger similar feelings.

to the immediacy of risk, despite cognitive evaluations of the risks remaining unchanged.

Economists have used the term anxiety before only in very specific circumstances. Maybe most notably, [Epstein and Kopylov \(2007\)](#) have a model of ‘cold feet’, in which a decision maker becomes more *pessimistic* as risks approach. Besides the prediction that people may pull back from risks previously decided to take, their axiomatization has little in common with our approach.

Experimental Evidence

There are several experimental studies documenting agents who are more risk averse if the resolution of uncertainty is temporally close than when it is distant. We want to highlight three studies which are particularly close to the phenomenon we address in this paper.³

[Jones and Johnson \(1973\)](#) have subjects participate in a simulated medical trial for a new drug where they have to decide on a dose of the drug to be administered. The subjects are told that the probability of experiencing unpleasant side-effects is increasing in the dose administered, as is the monetary compensation. More risk averse subjects should then choose lower doses than less risk averse subjects. In line with the predictions of our theory of anxiety, the study finds that subjects choose higher doses if they are to be administered the next day than when they are to be administered immediately.

In a second, more recent study by [Onculer \(2000\)](#), subjects are asked to state their certainty equivalent for a lottery to be resolved immediately, as well as for the same lottery to be resolved in the future. A lower certainty equivalent corresponds to

³For other studies see [Shelley \(1994\)](#), [Keren and Roelofsma \(1995\)](#), and [Sagristano, Trope, and Liberman \(2002\)](#).

higher risk aversion. The study finds that subjects state significantly lower certainty equivalents for the immediate lottery than for the future lottery.

The third study is by [Noussair and Wu \(2006\)](#). The study presents subjects with a list of choices between two binary lotteries. The first lottery always has prizes (\$10.00, \$8.00) while the second lottery always has prizes (\$19.25, \$0.50). Going down the list, only the respective probabilities of the two prizes change, varying from (0.1, 0.9) to (0.9, 0.1). As probability mass shifts from the second prize to the first prize, the second lottery becomes increasingly attractive compared to the first lottery. Subjects are asked to pick one of two lotteries for each of the probability distributions. The probability distribution at which a subject switches from the “safe” lottery to the “risky” lottery is a proxy for the subject’s risk aversion. One of the chosen lotteries is actually played out, either on the same day or three months later. The study finds that 38.5% of subjects are more risk averse for the present than for the future.⁴

In sum, people react differently to risks as a function of the time to resolution of the uncertainty *without believing the situation to get more risky*.

Overconfidence and its Relation to Forgetting

The previous subsection provided evidence to support the *assumption* of our model – higher risk aversion if the resolution of uncertainty is more imminent. In this section, we review psychological evidence of one of the model’s *predictions*, namely that anxiety-prone agents exhibit overconfidence.⁵

⁴7.7% are more risk averse for the future than the present, risk aversion of the other subjects does not change.

⁵We emphasize the distinction between overconfidence, which refers to holding beliefs with excessively high precision, and over-optimism, which refers to over-estimating the mean of a distribution. Neither is implied by the other, as [Hvide \(2002\)](#) clearly illustrates. Over-optimism is also documented in the psychology lit-

Beginning with [Adams and Adams \(1961\)](#), countless studies in cognitive psychology on the calibration of subjective probabilities have reported that people overestimate the precision of their knowledge (see [Alpert and Raiffa \(1982\)](#), [Kahneman and Tversky \(1973\)](#)). Subjects often answer general knowledge questions incorrectly, yet with high reported confidence or even certainty. Indeed, they are so confident that they are willing to bet on their answers' correctness ([Fischhoff, Slovic, and Lichtenstein \(1977\)](#)). The effect abates, but does not disappear, when subjects are informed about other subjects' overconfidence in the task at hand. Psychologists as subjects are no exception ([Oskamp \(1965\)](#)). More particularly, overconfidence is greatest for difficult tasks, for forecasts with low predictability, and for undertakings lacking fast and clear feedback ([Fischhoff, Slovic, and Lichtenstein \(1977\)](#), [Hoffrage \(2004\)](#)). Financial markets are a prime example of such an environment.

As for the mechanism how overconfidence is generated, in his essay "On the psychological mechanism of forgetting," Freud suggests that anxiety triggering information is prevented from entering memory and gets suppressed ([Freud \(2008\)](#), see also [Guenther \(1988\)](#)). An implication is that forgetting probabilities in anxiety triggering environments should be higher than in subjectively safe situations. [Zeller \(1950\)](#) shows that more anxious people are more forgetful as a result of repression. [Holmes \(1995\)](#) gives a review of other experiments validating the memory manipulation implications of anxiety. Deliberate memory manipulation is also implied in [Pearlin and Radabaugh \(1976\)](#), who find that people "who experienced increased anxiety (...), showed stronger tendencies to endorse drinking as a way of controlling distress" (see also [Morris and Reilly \(1987\)](#)).

erature, albeit less prominently than overconfidence (see [Langer \(1975\)](#), [Weinstein \(1980\)](#)). Applications of overoptimism to economics, such as [Van den Steen \(2004\)](#), are almost exclusively outside of the finance domain.

While overconfidence is a prediction of our model, existing models use overconfidence as an ingredient for finance applications. Agents in those models usually overestimate the precision of signals. Quite naturally, it leads to overreaction to the news associated with the overweighted signal (Daniel, Hirshleifer, and Subrahmanyam (1998)). Other uses of overconfidence are in explaining possibly excessive trade volume (Scheinkman and Xiong (2003)), and pricing of consumer products (Grubb (2009)). We are not aware of prior work that is concerned with overconfidence as a commitment device to take risks.

2.2 Model

Denote a possibly random intertemporal payoff stream from period t to period T by $X_t^T = (x_t, x_{t+1}, \dots, x_T)$. Our anxiety-prone agent evaluates the consumption stream X_t^T according to the utility function

$$U_t(X_t) = E_t [v(x_t) + \delta u(x_{t+1}) + \dots + \delta^{T-t} u(x_T)],$$

where v and u are von Neumann-Morgenstern utility indices, $\delta \leq 1$ is a discount factor and E_t is the expectations operator conditional on the information available at the beginning of period t .⁶

The only difference between our agent and a standard agent is that uncertainty in the current period is evaluated according to the utility function v while un-

⁶Assuming time separable utility inevitably implies marginal rates of intertemporal substitution. Several recent writings find this a desirable trait, and explore the joint effects of non-exponential discounting and implied non-constant risk aversion. See, for example Fudenberg and Levine (2010), and Halevy (2008). The inter-temporal effect is, however, not the focus of our paper. Therefore, for simplicity of exposition, we choose to relegate a treatment with Epstein-Zin preferences to a technical version of this paper, and choose examples in which the inter-temporal implications do not affect the results.

certainty in all future periods is evaluated according to the utility function u . To capture the effect of anxiety affecting imminent uncertainty, we assume that v is more risk averse than u .⁷ The key effect of this assumption is that it introduces a time inconsistency in the agent's preferences which implies that he may choose differently from a given set of alternatives depending on the period of choice. The following example illustrates this point.

Example Let $v(x) = \sqrt{x}$ and $u(x) = x$ and let $\delta = 1$. Then the decision maker is risk averse with respect to current uncertainty and risk neutral with respect to future uncertainty. Now consider the following two lotteries:

$$\tilde{x} = \begin{cases} 4 & \text{with prob. } \alpha \\ 0 & \text{with prob. } 1 - \alpha \end{cases} \quad \text{and} \quad \tilde{y} = 1$$

Then v prefers the risky \tilde{x} to the safe \tilde{y} if $\alpha > \frac{1}{2}$ while u prefers \tilde{x} to \tilde{y} if $\alpha > \frac{1}{4}$ and there is disagreement between the two utility functions for all $\alpha \in (\frac{1}{4}, \frac{1}{2})$. In particular, suppose that $\alpha = \frac{1}{3}$ and that the lotteries are resolved and paid out in period t . Then the agent will choose the safe option \tilde{y} in period t but would prefer to commit to the risky option \tilde{x} in all prior periods $t' < t$. He is willing to pay up to $\frac{1}{3}$ to commit to the risky option before period t and is willing to pay up to $\frac{5}{9}$ to avoid the risky option in period t .

⁷Our notion of “more risk averse than” is the standard one going back to [Pratt \(1964\)](#).

2.3 Finance Applications

2.3.1 Announcement Effects

We consider a standard asset pricing setup in discrete time with two periods $t = 0, 1$. There is a stock with net supply of 1 and a random payoff d which is realized at the end of period 1. No uncertainty is resolved between period 0 and period 1. The uncertainty about the stock's payoff is meant to represent a scheduled earnings announcement which provides information about the stock's dividend. It can also be interpreted more generally as the resolution of payoff-relevant information for holders of the stock – the key element is that the timing of the resolution is fixed and known in advance.

The price of the stock in period t is denoted by p_t and borrowing and lending is possible at a risk-free rate of zero. At the beginning of each period t , the agent has to form a portfolio (ϕ_t, ξ_t) of stock holdings and borrowing/lending, given beginning-of-period wealth w_t .

We solve backwards. In period 1, the uncertainty of the stock's payoff is imminent so the anxious agent chooses a portfolio (ϕ_1, ξ_1) to solve

$$\begin{aligned} & \max_{(\phi_1, \xi_1)} E[v(x_1)] \\ \text{s.t. } & x_1 = \phi_1 d + \xi_1 \\ & \phi_1 p_1 + \xi_1 \leq w_1 \end{aligned}$$

The first-order condition for an interior solution is

$$E[v'(\phi_1 d + w_1 - \phi_1 p_1)(d - p_1)] = 0. \quad (2.1)$$

If the agent already makes the portfolio decision in period 0, the non-anxious preferences u apply, and the first-order condition is

$$E [u' (\phi_0 d + w_0 - \phi_0 p_0) (d - p_0)] = 0. \quad (2.2)$$

Overtrading

Consider our anxiety-prone agent in an asset market dominated by standard agents with dynamically consistent risk aversion. Since there is no additional information revealed between period 0 and period 1, there is no reason for the price to change between the periods and we have $p_1 = p_0 =: p$. In addition, assume that the agent's wealth does not change so we have $w_1 = w_2 =: w$. Then, the first order conditions (2.1) and (2.2) simplify to

$$E [v' (\phi_1 (d - p) + w) (d - p)] = 0$$

and $E [u' (\phi_0 (d - p) + w) (d - p)] = 0.$

This gives us the following result adapted from [Wang and Werner \(1994\)](#).

Proposition 2.1 *If v is more risk averse than u , we have $\phi_0 > \phi_1$.*

This result shows that our agent wants to hold more of the risky asset in period 0, with some distance to the risk, than in period 1, when the resolution of uncertainty is imminent. The implications of this result depend on the degree of sophistication of the agent. A sophisticated agent anticipates in period 0 that he will want to change his portfolio in period 1. If the agent has no way of preventing his future self from rebalancing, he may already choose the anticipated portfolio ϕ_1 in period 0 to avoid trading costs.

The more interesting case is that of a naive agent. In period 0, he will choose a portfolio ϕ_0 but once the resolution of uncertainty is imminent in period 1, he sells some of the risky asset to attain the portfolio $\phi_1 < \phi_0$. When we view the asset market as a sequence of periods with and without news about the asset, the agent overtrades, selling some of the stock before announcements and buying it back afterwards. [Lamont and Frazzini \(2007\)](#) find evidence that selling pressure before announcements indeed stems from small and supposedly unsophisticated traders, as does the buy pressure after announcements. Large and supposedly sophisticated traders take the other side of these trades.

Notably, in the presence of transaction costs, an anxious investor will earn lower returns than a buy-and-hold investor due to overtrading, as in [Odean \(1999\)](#). We examine other factors affecting individual investors' returns in the following sections.

Price Dip

To derive pricing implications, we now model an economy with an anxiety-prone representative agent. This implies that he has to hold the entire net supply of the stock, $\phi_t = 1$, consumes the entire payoff, $x_1 = d$, and cannot borrow or lend, $\xi_t = 0$. Substituting these values into the first order conditions [\(2.1\)](#) and [\(2.2\)](#), they simplify to

$$E [v' (d) (d - p_1)] = 0 \tag{2.3}$$

$$E [u' (d) (d - p_0)] = 0 \tag{2.4}$$

and we have the following result.

Proposition 2.2 *If v is more risk averse than u , we have $p_0 > p_1$.*

This result shows that the price at which the agent is willing to hold the stock is lower when the resolution of uncertainty is imminent than when it is still distant. If the agent is naive about his anxiety, he will be happy to hold the stock at a price of p_0 in period 0, irrationally expecting the price not to change in period 1. Once the earnings announcement is imminent, the agent becomes anxious and the price drops to p_1 . Note that the price jumps after the announcement (albeit not as much) also also in a model with a standard risk averse agent. However, the price dip before the announcement is uniquely produced by anxiety.

Rewriting the expectations in conditions (2.3) and (2.4) allows us to write the prices explicitly:

$$p_0 = E[d] + \frac{Cov(u'(d), d)}{E[u'(d)]} \quad \text{and} \quad p_1 = E[d] + \frac{Cov(v'(d), d)}{E[v'(d)]}$$

The second term in the two price equations is the risk premium. It discounts expected dividends more strongly at $t = 1$ than at $t = 0$, as shown in Proposition 2.2. In particular, the covariances are negative and the expectations positive, as both u' and v' are positive but decreasing. As one should expect, increasing but risk averse utility functions imply a price discount of the risky asset, relative to expected value. More risk aversion makes for heavier discounting, and vice versa. In the case of risk neutrality, $u'(d) = c$, the covariance term is zero as $u''(d) = 0$. Then, the asset trades at expected dividends.⁸

In a market populated by both anxious and standard agents, there will be a price drop before any scheduled announcement but not as large as in a market with only

⁸The same pricing equations result if the representative agent maximizes $u(x_0) + v(x_1)$ in period $t = 0$ and consumes out of wealth.

anxious agents. Accompanying the price drop we should expect to see anxious agents selling part of their stocks to standard agents. Right after the announcement, prices should on average appreciate as anxiety-prone agents buy back their positions.

Our theory thus combines predictions about both asset price movements *and* trade volume around announcement dates, which is a crucial feature of announcement anomalies, as [Lamont and Frazzini \(2007\)](#) explain. These authors also confirm that institutional investors lean against the individual investors' trades. A strategy of buying before announcement dates and selling thereafter yields excess returns of 7% to 18%, which they call the announcement premium. While their paper focuses on explaining the price and volume patterns with the 'attention grabbing hypothesis' (see also [Lee \(1992\)](#), [Hou, Peng, and Xiong \(2009\)](#), [Barber and Odean \(2008\)](#)), their empirical results provide equal support for our theory. Our theory shifts the focus to the other side of the same medal that [Lamont and Frazzini \(2007\)](#) examine: we ask why prices tend relatively *lower before* the announcement, which is depicted by [Bernard and Thomas \(1989\)](#). We call this the 'pre-earnings announcement dip'. We thereby offer a possible "common underlying cause for both volume and the premium" that [Lamont and Frazzini \(2007\)](#) have called for, as an alternative to the 'attention-grabbing hypothesis'.

Realized Returns

The stylized model above is not yet suited to be calibrated with data. However, the analysis in [Bernard and Thomas \(1989\)](#) suggests a pre-earnings announcement dip on the order of -0.5% (smaller for large firms than for small-caps). With four scheduled earnings announcements per year, a naive agent as depicted above stands to lose about 2% per year by overtrading in the face of scheduled quarterly earnings

announcements alone. This loss comes on top of the transaction costs of overtrading. This squares nicely with the empirical result by [Lo, Repin, and Steenbarger \(2005\)](#), who confirm that more anxious agents generate lower returns. Our model predicts a similar price effect for scheduled news events relevant to the equity market as a whole, such as the publication of unemployment figures.

A naive anxiety-prone investor's actual equity returns, i.e. the returns he enjoys from investing in equity after accounting for the losses imposed by anxious behavior, are lower than buy-and-hold returns derived from market data may suggest. This helps explain the equity premium puzzle. The Equity Premium Puzzle (EPP) states that equity returns are too high relative to bond returns than can be explained by reasonable levels of risk aversion and discount rates. If risk aversion were as high as implied by the difference between equity and bond returns, bond returns would have to be much higher than they actually are. The latter part of the problem is known as the "risk-free rate puzzle." Hence, models attempting to explain the EPP with agents, who are, effectively, very risk averse, can only explain the difference in returns between bonds and equity, but fail to explain the ensuing risk-free rate puzzle. For example, models assuming ambiguity aversion typically run into that problem. In contrast, anxiety explains part of the EPP without running into the risk-free rate puzzle by showing that effective equity returns to an anxiety-prone investor are not as high as they appear in the data, while bond returns are unaffected by anxiety. Dynamic inconsistency with respect to risks only distorts the price of the locally risky equity, but not the price of locally risk-free bonds.

Our predictions stem from the analysis of a naive anxiety-prone agent. As we will discuss in the next section, a sophisticated anxiety-prone agent may find ways to behave in a dynamically consistent way and thus suffer to a lesser extent from the

costs of overtrading. Yet, the disutility implied by the use of a commitment device needs to be subtracted from the utility from equity returns of such an agent. For example, the following section shows how overconfidence can let an anxiety-prone decision maker make more dynamically consistent decisions. But then, the disutility from overconfidence, stemming from ‘excessive risk-taking’, needs to be subtracted from the now higher utility from holding equity without overtrading. Consequently, even a sophisticated anxiety-prone agent will find equity a worthwhile investment only at returns that are higher than the ones a standard consumption-based asset pricing model yields.

Note that most firms’ equity prices may also be depressed, since institutional counterparties may find it more profitable to use their capital to exploit the behavioral distortions of retail investors trading in stocks that have imminent earnings announcements, instead of pushing up equity prices across the board. Moreover, anxiety-prone agents’ counterparties may anticipate the selling pressure by anxious agents before earnings announcements. If (they know that) they can not absorb the sales at the same price level, they will demand higher premia already ahead of the announcement date.

In sum, our theory of anxiety in the face of risk thus links the equity premium, price reactions to earnings announcements, and overtrading, and square nicely with the results of [Lo, Repin, and Steenbarger \(2005\)](#) on the relation of anxiety and trading performance.

2.3.2 Institutional Effects

An agent who plans according to preferences u , but is afraid his future self will disagree with these plans (because of having preferences v), may try to find ways

to commit his future actions to his presently chosen plan of action. While Schelling (1984) and others have discussed the ethical aspects such a possibility brings about, the present discussion is only concerned with *that*, and how, the agent can restrict his future self's behavior – simply by virtue of having a first-mover advantage. Indeed, dynamic inconsistency with respect to risks gives a strong economic rationale for doing so. As sketched out above, an anxiety-prone agent faces losses that are not compensated by higher consumption at any time (as is the case for a hyperbolic discounter).

Delegation Hiring an agent to carry out risk-taking decisions in the *future* according to the *current* self's preferences is one way to prevent future selves' preferences from conflicting with the current self's plans. In an investment setting, it may be the case that the anxious self is too risk averse to invest in equity, although the agent realizes this has long-run benefits. In this situation it makes sense for the agent to delegate investment decisions to a portfolio manager. The manager can still react to news about particular assets, but has to stick to a predetermined split of asset classes.

As is the case for commitment devices for hyperbolic discounters, it is clear that having them is desirable, but it is less clear when an agent would start using them. The delegation of investment decisions provides a nice exemption to that rule. An agent prone to anxiety differs from a standard agent only in his evaluation of immediate risks. Thus, we expect to see greater inflows to money managers when immediate risks seem to be low, relative to the associated returns, even if such a temporary calm does not carry information about future performance. This may help to reinterpret respective evidence from the mutual fund industry. As falling prices increase risk estimates, low returns should be associated with low inflows

to money managers. Indeed, [Sirri and Tufano \(1998\)](#) find that high returns trigger fund inflows, and vice versa.

Of course, effort costs of managing one's portfolio may also lead to delegation of investment management. However, effort costs can not justify hiring an agent that underperforms the index on average, as buying index funds is virtually costless and free of effort. Yet, the mutual funds industry is huge, and actual fund managers still tend to underperform the market [Gruber \(1996\)](#). While buying the index is free of effort, it is not free of anxiety. Self 0 may thus correctly anticipate that the anxious self 1 will underperform the market even more than a random portfolio manager by failing to invest in equity at all. Self 0 will therefore be willing to pay an investment manager, even if he expects him to underperform the market. The obvious solution would be to hire an agent to simply buy the index, but that may be infeasible in a model of career concerns.

Fees A redemption fee is another feature of investment funds that sophisticated anxiety-prone decision-makers will demand. This may be one explanation why management and other fees are being competed away in the mutual funds industry, while lock-in fees continue to feature prominently. Variations of punishments for pulling out of risks an investor previously decided to take include fees for changing the equity/bonds ratio of one's investment in mutual funds, as well as fees imposed if the total exposure to a certain asset class falls below a threshold.

Timing of Orders The widespread practice of retail investors to submit overnight limit orders can be viewed as another costly way of coping with anxiety. Submitting overnight limit orders deprives the investor from the possibility to react to news in the time between submission of the order and execution, and furthermore rep-

resents a positive externality to other market participants: it represents an option to buy/sell at the quoted price. See [Harris \(2003\)](#) for a discussion. Writing such an option to trade, as well as foregoing the option to react to overnight news, would never be optimal for a standard agent. However, it helps overcome commitment problems imposed by anxiety. Instead of waiting to see his future self pull out from the decision to invest in the stock, the current self preempts the decision before going to bed, when the uncertainty is not yet imminent.

Demand for Delayed Resolution of Uncertainty and Costs of Information

Self 1's risk preferences about future gambles are identical to self 0's preferences about the same gambles if there is no immediate resolution of uncertainty at $t = 1$. This implies a disutility for resolution of uncertainty, i.e. a disutility for information, in period $t = 1$. Self 0 will therefore be willing to pay for delaying the resolution of uncertainty from $t = 1$ to a later date in order to harmonize self 1's behavior with self 0's preferences. To be sure, this is not driven by a preference for the timing of the resolution of uncertainty, which *requires* temporal consistency [Kreps and Porteus \(1978\)](#). Hedge funds impose pull-out restrictions and publish performance reports at low frequencies, although the information is available continuously and creating a report is a largely automatable task. Note that the cost of having to provide liquidity does not explain such clauses. Imposing costs on deposits with short maturities will compensate the fund for the cost of liquidity provision, but putting a temporal distance between the investor's decision to pull out and the payout of the funds does neither protect the fund from withdrawals nor compensate for the implied costs. Concealing present risks, however, prevents anxious investors from pulling out.

2.4 Overconfidence

If commitment devices are not available, an anxiety-prone agent has an incentive to distort his future self's beliefs. In particular, the present self would like to convince his future self that risks are lower than they actually are. This would lead the future self to take riskier decisions which are more in line with the current self's preferences. However, if the future self has access to additional information, the distorted beliefs may lead to decisions that are excessively risky, even from the current self's point of view. In this section we analyze such a situation similar to the model of [Bénabou and Tirole \(2002\)](#).

For the sake of simplicity, we again restrict ourselves to two time periods, $t = 0, 1$, and set the discount factor to $\delta = 1$. In period 1 the agent has to choose between a risky or a safe alternative. The risky alternative is given by a lottery with random payoff x . The lottery is characterized by its distribution function G_θ where $\theta \in \{H, L\}$ denotes a state of the world that determines how risky the lottery is. We assume that G_H is a mean-preserving spread of G_L so the risky alternative is unambiguously riskier in state H than in state L . The prior probability of the high-risk state H is given by π . The safe alternative, on the other hand, is given by a constant payoff a .

The anxious agent in period 1 wants to take the risky alternative whenever

$$E_\theta[v(x)] > v(a),$$

where E_θ denotes the expectation with respect to G_θ . Denoting the certainty equivalent of G_θ given the utility function v by c_v^θ , this condition can be rewritten as

$$c_v^\theta > a.$$

The agent wants to take the risky alternative, whenever its certainty equivalent c_v^θ is greater than the safe alternative a .

The agent in period 0, when the risk is not imminent, wants to take the risky alternative whenever

$$E_\theta[u(x)] > u(a) \Leftrightarrow c_u^\theta > a.$$

Since v is more risk averse, we have $c_u^\theta > c_v^\theta$ for both $\theta \in \{H, L\}$ so the agent in period 0 (self 0) and the agent in period 1 (self 1) will disagree about the course of action if $a \in [c_v^\theta, c_u^\theta]$.

To make this problem interesting, we assume that the payoff of the safe alternative a is not known to the agent until period 1. Self 0 only knows the prior distribution F on $[\underline{a}, \bar{a}]$ but self 1 observes the realized value of a . The state of the world θ , on the other hand, is revealed to the agent at the beginning of period 0 in form of a perfectly informative “red flag” warning signal s if the state is high-risk

$$s = \begin{cases} R & \text{if } \theta = H \\ \emptyset & \text{if } \theta = L \end{cases}$$

If he receives a red flag, self 0 can choose the probability $\lambda \in [0, 1]$ with which he will remember the signal, i.e.,

$$\lambda = \Pr[\hat{s} = R | s = R],$$

where \hat{s} is self 1's recollection of the signal. We assume that self 1 is fully aware of his prior incentive to forget warning signals, so if he expects a memory probability λ^e and doesn't remember seeing a red flag he uses a Bayesian posterior

$$\pi(\lambda^e) = \frac{\pi(1 - \lambda^e)}{\pi(1 - \lambda^e) + 1 - \pi}.$$

Given this setup, self 0 and self 1 are playing a kind of Stackelberg game. First self 0 chooses the memory probability λ taking into account self 1's behavior and then self 0 decides between the risky and the safe alternative taking into account self 0's behavior. We are interested in the perfect Bayesian equilibria of this intrapersonal game.

First, we derive self 1's best response in $t = 1$, taking as given an expected memory probability λ^e . If self 1 remembers seeing a red flag, $\hat{s} = R$, he knows that the state of the world is high-risk and chooses the risky alternative if $c_v^H > a$. If self 1 doesn't remember seeing a red flag, $\hat{s} = \emptyset$, he uses the Bayesian posterior $\pi(\lambda^e)$ and chooses the risky alternative if $c_v(\lambda^e) > a$ where $c_v(\lambda^e)$ is the certainty equivalent of the risky alternative given λ^e defined by

$$E[v(x) | \pi(\lambda^e)] = v(c_v(\lambda^e)).$$

Second, we derive self 0's best response in $t = 0$, taking as given self 1's behavior to an expected λ^e . If self 0 receives a warning signal and chooses a memory

probability λ , his expected utility is

$$\lambda \left[\int_{\underline{a}}^{c_v^H} E_H[u(x)] dF(a) + \int_{c_v^H}^{\bar{a}} u(a) dF(a) \right] \\ + (1 - \lambda) \left[\int_{\underline{a}}^{c_v(\lambda^e)} E_H[u(x)] dF(a) + \int_{c_v(\lambda^e)}^{\bar{a}} u(a) dF(a) \right].$$

With probability λ the agent remembers the warning signal in period 1 and uses the certainty equivalent c_v^H as the threshold, choosing the risky alternative for payoffs of the safe alternative below the threshold and choosing the safe alternative for payoffs above the threshold. With probability $1 - \lambda$ the agent forgets the warning signal and uses the certainty equivalent $c_v(\lambda^e)$ as the threshold.

We denote the derivative of self 0's expected utility with respect to λ by

$$D(\lambda^e|v) := \int_{c_v^H}^{c_v(\lambda^e)} (u(a) - E_H[u(x)]) dF(a).$$

This expression has a very natural interpretation. The warning signal changes self 1's decision only for values of $a \in [c_v^H, c_v(\lambda^e)]$. In this interval, self 1 chooses the risky alternative whenever he remembers seeing a red flag and the safe alternative otherwise. The effect on self 0's expected utility of remembering the warning signal more often is exactly the difference in utility from the safe action compared to the risky action for the values of a where the decision is affected.

There are three possibilities for perfect Bayesian equilibria in this setting:

- **Honesty Equilibrium:** If $D(1|v) \geq 0$, there is an equilibrium with $\lambda^* = 1$. In this equilibrium the agent never ignores red flags and doesn't influence his future self's beliefs.

- Overconfidence Equilibrium: If $D(0|v) \leq 0$, there is an equilibrium with $\lambda^* = 0$. In this equilibrium the agent always ignores red flags and makes his future self maximally overconfident.
- Mixed Equilibrium: If $D(\bar{\lambda}|v) = 0$ for some $\bar{\lambda} \in (0, 1)$, there is an equilibrium with $\lambda^* = \bar{\lambda}$. In this equilibrium the agent plays a mixed strategy, ignoring the red flag with probability $1 - \bar{\lambda}$, and makes his future self partially overconfident.

Proposition 2.3 *One of the extreme equilibria always exists, either the honesty equilibrium or the overconfidence equilibrium or both. If both extreme equilibria exist, a mixed equilibrium also exists.*

The existence of each kind of equilibrium depends on the degree of anxiety of the agent, i.e., how big the difference in risk aversion is for risks that are imminent compared to risks that are distant. In particular, we can say that an agent i is more prone to anxiety than an agent j , if u_i and u_j are equally risk averse but v_i is more risk averse than v_j . This enables us to state the following result.

Proposition 2.4 *For an agent that is more prone to anxiety, (i) the honesty equilibrium is less likely to exist, (ii) the overconfidence equilibrium is more likely to exist, and (iii) if the mixed equilibrium exists, then it is associated with more overconfidence.*

Somewhat counterintuitively, people who are most prone to anxiety in the face of risk are the same ones that are most likely to exhibit overconfidence. Note further that a risky environment is necessary for overconfidence to arise, and to show effects in decision making. Financial markets are a prime example of such an environment.

Ben-David, Graham, and Harvey (2010) confirm that financial top executives are systematically overconfident: realized market returns are within their 80% confidence intervals only 33% of the time. A manifestation of overconfidence that is important in finance, and possibly important to understand individual agents' behavior during the recent financial crisis, is excessive risk-taking.

Excessive Risk-Taking Equilibria with partial or maximal overconfidence can display excessive risk taking. In these equilibria it can be the case that the future self ends up taking risks which even the less risk averse current self would have avoided. To an observer who is unaware of the agent's intra-personal conflict, the agent seems to take risks that are greater than can be explained with 'reasonable' preferences, e.g. u . This can happen if the true state of riskiness is high and the agent forgets the warning signal. In this case, whenever the payoff of the safe alternative is below the cutoff $c_v(\lambda^*)$ self 1 uses but above the cutoff c_u^H self 0 would like him to use, i.e. $a \in (c_u^H, c_v(\lambda^*))$, the agent takes risks in period 1 that that self 0 considers excessive. Analytically, this can arise since the condition for an equilibrium with overconfidence, $D(\lambda^*|v) \geq 0$, does *not* necessarily imply $E_H[u(x)] > u(a)$ for all $a < c_v(\lambda^*)$, where self 1 chooses the risky alternative. Such a situation arises in all equilibria λ^* with $c_u^H < c_v(\lambda^*)$, i.e. the equilibrium cutoff used by self 1 is greater than the cutoff self 0 would use. To an outside observer who knows that the state is H , the anxious agent using the cutoff $c_v(\lambda^*)$ appears as if he were *less* risk averse than the non-anxious preference u .

Proposition 2.5 *In an equilibrium with $\lambda^* < 1$ and $c_u^H < c_v(\lambda^*)$, the agent will be observed to take excessive risks, i.e. he will appear less risk averse than u .*

Ben-David, Graham, and Harvey (2007) confirm empirically that overconfidence, observed in Ben-David, Graham, and Harvey (2010), translates into riskier corporate policy.

2.5 Conclusion

In this paper, we define an anxiety-prone decision maker as an agent, whose risk aversion is higher the closer in time the resolution of uncertainty is. We discuss experimental evidence that is predicted by our model and show in examples and in a financial market model how this leads to dynamically inconsistent behavior. Linking such behavior to established puzzles about price and volume around earnings announcements, we suggest a clean, and arguably more credible way to think about these patterns than existing theories propose. Evidence from the trading floor also confirms our prediction that more anxiety-prone traders perform worse. We explain how sophistication about dynamic inconsistency and the associated costs will trigger institutional responses such as delegation of investment decisions, and the distinct design of brokerage and investment fund fees. We further suggest a connection to optimal patterns of information provision in financial markets. Finally, we show why it may be beneficial to a sophisticated anxiety-prone agent to hold overconfident beliefs, and how this can be accomplished.

Combining the above model of endogenous overconfidence with problems in financial economics seems a fruitful field of future research. We conjecture four possible areas of applications.

First, there should be an equilibrium level of overconfidence in financial markets. The costs of overtrading due to anxiety around news announcements can be mitigated by overconfidence. On the other hand, overconfidence may cause over-

trading independent of news announcements according to [Scheinkman and Xiong \(2003\)](#). Such trading, while not directly causing expected losses, still bears transaction costs. But in addition, an overconfident agent also suffers from excessive risk taking, implying a disutility for the planning-self at $t = 0$. Trading off these costs should yield an optimal amount of overconfidence according to self 0's preferences. The equilibrium level of overconfidence should be increasing in transactions costs and bid-ask spreads, and thus be more pronounced in more illiquid securities. It should be negatively related to the earnings announcement premium, i.e. the predictable price fluctuations between announcement periods and periods without earnings announcements, and positively to the frequency of scheduled announcements.

Second, recent influential works by [Akerlof and Shiller \(2010\)](#) and [Reinhart and Rogoff \(2009\)](#) have strongly suggested that time-changing confidence needs to be part of realistic models of market dynamics and the business cycle. Empirically, confidence is high when leverage is high and maturities are short, and vice versa. This is consistent with our notion that overconfidence arises when risks are high, and (not shown in the above model) under-confidence may arise when risks are low. As overconfident traders have a greater demand for risk than rational types do, overconfidence sustains excessive risk levels. Conversely, under-confidence helps sustain price levels below fundamentals in the crisis. Both outcomes may be possible under the same parameters in a model with multiple equilibria. Extending this static argument to a dynamic model will be more challenging.

Third, the above model of self-delusion is not necessarily to be taken literally, but can be seen as a metaphor for the choice of information systems and communication structures in organizations. Given a preference for a biased posterior, an anxiety-

prone leader will implement information and communication systems that have him misinformed about risks. The scarcity of critical upward feedback, which is often said to be mandated by the head of the organization ('killing the messenger'), may be explained in this way. The more anxiety-prone the leader, the less upward feedback will be provided.⁹ In the investment domain, the 'Ostrich Effect' may serve as an example. [Karlsson, Loewenstein, and Seppi \(2005\)](#) find that investors look up their portfolio performance less often after receiving a signal about increased risks.¹⁰

Fourth, occupational choices and associated cognitive dissonance may be a fruitful domain for applications of the overconfidence model. Nothing in the model prevents that the agent, rather than nature, choose the riskiness of the environment (and the thus implied perfectly informative signal). Parallel to the mechanism in the present model, the agent may choose to forget the information the he based his prior decision upon, i.e. that he chose a risky job over a safe one, and thus render himself overconfident (see [Akerlof and Dickens \(1982\)](#)). This will be beneficial if the agent's job involves risk-taking. Professions such as securities trading should then be particularly likely to feature overconfident agents.

⁹Management publications view the lack of upward feedback as the source of countless corporate disasters and a widespread phenomenon. There are also examples in history, where leaders that were certainly not known for pronounced propensity to anxiety, demanded critique by any means. Queen Elizabeth I is said to have rebuked a jester "for being insufficiently severe with her."

¹⁰The original finding is that investors tend to not look up their portfolio's performance after market-wide declines, about which they are likely to become informed via generic news reports. Note that (i) price drops may be caused by increases in risk levels, but also (ii) falling prices increase volatility estimates. Thus, in any case, falling prices are a signal for increased risk.

2.6 Appendix

Proof of Proposition 2.1. See [Wang and Werner \(1994\)](#). □

Proof of Proposition 2.2. Since v is more risk averse than u we have

$$\begin{aligned} -\frac{v''(x)}{v'(x)} &> -\frac{u''(x)}{u'(x)} \\ -\frac{d}{dx} \log v'(x) &> -\frac{d}{dx} \log u'(x) \end{aligned}$$

Integrating both sides yields

$$\frac{v'(d)}{v'(p)} < \frac{u'(d)}{u'(p)}$$

for $d > p$ and the reverse inequality for $d < p$. For general p, d we then have

$$\left(\frac{u'(d)}{u'(p)} - \frac{v'(d)}{v'(p)} \right) (d - p) > 0$$

Taking expectations we get

$$\frac{E[u'(d)(d-p)]}{u'(p)} > \frac{E[v'(d)(d-p)]}{v'(p)}$$

Substituting in p_1 the RHS is zero and we get

$$E[u'(d)(d-p_1)] > 0,$$

which implies that $p_0 > p_1$. □

Proof of Proposition 2.3. The belief $\pi(\lambda^e)$ is continuous and decreasing in λ^e . Therefore the certainty equivalent $c_v(\lambda^e)$ is continuous and increasing in λ^e . Finally, this implies that $D(\lambda^e|v)$ is continuous and increasing in λ^e . We then have either

$D(1|v) \geq 0$ or $D(0|v) \leq 0$ or both so one of the extreme equilibria $\lambda^* \in \{0, 1\}$ always exists. For the case where $D(1|v) \geq 0$ and $D(0|v) \leq 0$, there exists a $\bar{\lambda} \in (0, 1)$ such that $D(\bar{\lambda}|v) = 0$ so the mixed equilibrium $\lambda^* = \bar{\lambda}$ also exists. \square

Lemma 2.1 *Consider two von Neumann-Morgenstern utility functions v_1 and v_2 . If v_2 is more risk averse than v_1 , then $D(\lambda^e|v_2) < D(\lambda^e|v_1)$ for all λ^e .*

Proof. If v_2 is more risk averse than v_1 , then $c_{v_2}^H < c_{v_1}^H$ and $c_{v_2}(\lambda^e) < c_{v_1}(\lambda^e)$ for all λ^e . This implies that for all λ^e :

$$\begin{aligned} D(\lambda^e|v_2) &= - \int_{c_{v_2}^H}^{c_{v_2}(\lambda^e)} (u(a) - E_H[u(x)]) dF(a) \\ &< - \int_{c_{v_1}^H}^{c_{v_1}(\lambda^e)} (u(a) - E_H[u(x)]) dF(a) \\ &= D(\lambda^e|v_1) \end{aligned}$$

\square

Lemma 2.2 *Consider two von Neumann-Morgenstern utility functions v_1 and v_2 . If v_2 is more risk averse than v_1 and if there are $\bar{\lambda}_1$ and $\bar{\lambda}_2$ such that $D(\bar{\lambda}_1|v_1) = 0$ and $D(\bar{\lambda}_2|v_2) = 0$, then $\bar{\lambda}_1 < \bar{\lambda}_2$.*

Proof. If v_2 is more risk averse than v_1 , then $c_{v_2}^H < c_{v_1}^H$ so the integral in $D(\bar{\lambda}_2|v_2)$ has a smaller lower bound. Since $(u(a) - E_H[u(x)])$ is a strictly increasing function of a , for $D(\bar{\lambda}_1|v_1) = D(\bar{\lambda}_2|v_2) = 0$ it is necessary that $c_{v_2}(\bar{\lambda}_2) > c_{v_1}(\bar{\lambda}_1)$, i.e. that the integral in $D(\bar{\lambda}_2|v_2)$ must have a greater upper bound. Since $c_{v_2}(\lambda) < c_{v_1}(\lambda)$ for a given λ , and $c_v(\lambda)$ is increasing in λ for v_1 and v_2 , this implies $\bar{\lambda}_2 > \bar{\lambda}_1$. \square

Proof of Proposition 2.4. From Lemma 2.1 we know that $D(1|v_2) < D(1|v_1)$ for v_2 more risk averse than v_1 . Therefore an honesty equilibrium exists for v_1 if it

exists for v_2 . Again using Lemma 2.1 we know that $D(0|v_2) < D(0|v_1)$ for v_2 more risk averse than v_1 . Therefore an overconfidence equilibrium exists for v_2 if it exists for v_1 . Finally, if a mixed equilibrium exists for v_1 and v_2 , characterized by $\bar{\lambda}_1$ and $\bar{\lambda}_2$ respectively, then by Lemma 2.2 we have $\bar{\lambda}_1 < \bar{\lambda}_2$. \square

Proof of Proposition 2.5. Follows directly from the derivation in the main text.

\square

Chapter 3

Flexibility as an Instrument in Digital Rights Management¹

3.1 Introduction

The arrival of digital goods came with the promise of easy transferability and portability across various media and devices. In fact, for a user of digital goods, the corresponding flexibility is often an essential aspect of their valuation. Yet, for the provider of these goods, flexibility comes with the risk that unlicensed copies will circulate and undermine revenue-generating sales.

The objective of digital-rights-management (DRM) technologies is to enable the providers of digital goods to control the details of how consumers can use the goods. In many current DRM systems, the provider attempts to control the consumers' use of the good along several dimensions. Typical parameters include how long the

¹An early version of this paper was presented at the Fourth Workshop on the Economics of Information Security at the Kennedy School of Government, Harvard University.

consumer can use the good, how often he can use it, on how many devices he can use it simultaneously, and whether he can copy or alter it in any way.

The current paper aims to analyze the basic design of a DRM system as an optimal trade-off between the increase in the value of a licensed copy and the increase in the number of unlicensed copies. Intuitively, an increase in the allowed flexibility of a digital product increases the value of the product for its user and hence will allow the seller to charge a higher price for a licensed copy. On the other hand, with an increase in the flexibility comes the risk that a non-paying customer will get, legally or not, access to the digital good. Hence an increase in flexibility may undermine sales volume. We explicitly model the choice of flexibility in an environment where perfect security is only possible in the limit when flexibility is severely restricted. This is meant to represent the pervasive view that the Internet will always be a “greynet” without perfect security provisions.²

We begin our analysis with a single content provider who offers a digital good to many consumers. The consumers have to choose between acquiring a licensed copy of the product and hoping to receive an unlicensed copy. The likelihood that the consumer will be able to receive an unlicensed copy is increasing in their permitted flexibility. The policy instruments of the content provider are price and permitted flexibility. An increase in the flexibility increases the revenue per item sold, but it also increases the likelihood that a given consumer will obtain access to an unlicensed copy. The resulting equilibrium policies of the content provider will attempt to find the optimal balance between flexibility and sales. In equilibrium, the consumers will be split into buyers of licensed products and consumers of unlicensed

²We refer to “greynet” here to describe the use of digital files outside the strictly licensed context. This includes both the use of unlicensed copies on a small scale and the possibility of file sharing through peer-to-peer networks. [Biddle, England, Peinado, and Willman \(2002\)](#) used the term “darknet” to describe exclusively peer-to-peer networks.

copies. The equilibrium volume of sales will be determined endogenously by price and flexibility. An important determinant of the equilibrium policies will be the rate at which licensed copies translate into access to unlicensed copies. In reality, this may depend on factors such as bandwidth of internet links, social connectedness, and other technological as well economic determinants.

In the case of online music sales, the most successful example is certainly Apple. It is currently by far the dominant provider of high quality digital-music files with its music store and playback software iTunes. Apple's success in selling music files is closely connected to its introduction of the portable music player iPod. In addition to having a significantly larger storage capacity than the previously common flash memories, the iPod also makes use of DRM technology. Only high quality files bought from Apple and those extracted from a user's own CDs using the iTunes software can be played by an iPod.³ Conversely, the high-quality files from Apple's iTunes store can only be played on its own devices. The software and hardware provided by Apple clearly represent complementary products to the digital good. In the specific case of iTunes and iPods, they represent a platform for the use of the digital good that enhances the value of that good. At the same time, the digital goods sold by Apple can be used only on the platform provided by Apple. The platform thus achieves two objectives for Apple. It enhances the security of the DRM system itself, but it also restricts the use of unlicensed copies. Even the unlicensed copies can essentially only be used on the Apple platform. As a result, Apple as the platform provider can realize revenue from two sources: the sales of the music files and the sale of the platform (i.e. the hardware and associated software).

³The iPod also plays low quality files as MP3 which certainly are no perfect substitutes for high quality files.

We therefore investigate the role of a platform in the context of DRM. We make the assumption that, although the digital good may be acquired in the form of an unlicensed copy, it will still have to run on the platform. This assumption completely removes concern about the security of the platform, but the essential part of the argument only requires that the platform be less susceptible to unlicensed appearance than the digital good itself.

We then compare the outcomes of two extreme cases, assuming first that separate firms sell the digital good and the platform and then second that an integrated firm sells both. The analysis of two separate firms shows that there is a natural conflict between the owner of the rights to the digital good and the owner of the rights to the platform. The owner of the digital good would like to increase the revenue-generating sales of the good. For this reason, the content provider will want to reduce the flexibility and increase the price. On the other hand, the platform provider cares less about the revenue coming from the sales of the digital good and more about the perceived value of the platform. He will therefore want to increase the flexibility of the DRM system, thus increasing the number of circulating copies of the digital good, licensed or not, in order to sustain the market for the platform. We show that the resulting equilibrium will lead to a low level of flexibility, a high price of the digital good and a low price for the platform.

Next we analyze the case of a single provider that sells both a platform for his digital content and the content itself. The products are offered jointly but priced separately. We show that the joint provider who also sells a platform finds it optimal to provide each user with a higher and socially more efficient level of flexibility than the provider who doesn't sell a platform. In addition, the price of the digital good itself will lower than before, even considering the higher level of flexibility.

However, the platform provider is less concerned about the unlicensed segment of the market, because he can recover part of the surplus that arises due to the availability of unlicensed access through revenue from the sale of the platform itself. Consequently, the price of the platform serves the same function as an entrance fee to an amusement park. Because the content provider cannot extract all the surplus in the market for digital goods, he will leave surplus to the consumers. Thus, he can charge a substantial price for the platform that gives the consumers access to the market for digital goods. In fact, we show that the joint provider charges a higher price for the platform than the platform provider in the case of separate firms. Note that this is a novel business model that contrasts with the model employed in other markets of complementary goods in which customers make a one-time purchase of a device and then make recurring purchases of items that complement the device or subscribe to a complementary service. For example, Gillette makes money by selling blades not razors, and integrated communications companies make money by signing up cell-phone subscribers rather than by selling phones.

The development of Apple's use of DRM since its entry into the digital music market strongly resembles the findings of our model. Initially, under the iTunes DRM rules, flexibility was rather limited. Each music file could be played on only five devices at the same time that had to be authorized by the buyer of the file. Playlists, i.e. specific arrangement of several files, could only be burned to CDs seven times. At the time Apple as the provider of the platform was in a relatively weak position when negotiating with the music industry who owned the rights to the digital music files. These early negotiations were characterized by the conflict between separate content provider and platform provider predicted by our model:

An article in the Financial Times on February 2nd 2005 quotes a music industry insider as saying “Our music is not something to be given away to sell iPods.”

In the time since then, Apple has become the dominant player in the market for digital music which has significantly increased its bargaining position against the music industry, moving the situation closer to our assumption of a joint provider. In 2007 Apple started a public push for the sale of files without DRM restrictions resulting in agreements with some record labels to sell DRM-free files at higher prices. By April of 2009 all music sold on iTunes was available without DRM restrictions. In contrast, in markets for digital goods where Apple does not have a dominant platform such as TV shows and movies, the files are still only sold with severe DRM restrictions.

Related Literature Several authors have put forth arguments about why piracy of easily reproducible goods might be beneficial to providers as well as consumers, thus adding new aspects to the discussion about copyright protection. [Liebowitz \(1985\)](#) was the first to show that, when each good is shared by a defined group of consumers (also called a “club”), the provider can indirectly appropriate revenues from all members of the group by charging a higher price. [Varian \(2000\)](#) finds that piracy in groups can be beneficial to the provider if sharing is cheaper than producing additional units, or if it enables price discrimination based on consumers’ different valuations. [Bakos, Brynjolfsson, and Lichtman \(1999\)](#) emphasize that selling to groups may reduce demand uncertainty (just as bundling reduces it) and thus enable more profitable pricing. [Parker and van Alstyne \(2005\)](#) consider the pricing of complementary products in a model of two-sided markets. In our model, the complementary products, content and platform, are offered in a single market.

Dropping the assumption of sharing in defined groups, [Conner and Rumelt \(1991\)](#) and [Takeyama \(1994\)](#) show that piracy can increase profits if the good exhibits a positive network externality. Because piracy expands the user base, thus increasing the value of the good, the provider can charge buyers higher prices than he could without piracy. [Sundararajan \(2004\)](#) considers the role of digital management to restrict digital piracy in the context of an optimal pricing model. In his model, the possibility of piracy acts as a constraint on the pricing policy, but there is no interaction between the level of flexibility and the implicit cost of piracy in terms of foregone sales.

In an intertemporal setting, [Takeyama \(1997\)](#) finds that piracy among low-valuation consumers can reduce the provider's price-commitment problem if the good is durable over time. The negative effect of piracy on the quality the provider offers for his goods is studied in an early paper by [Novos and Waldman \(1984\)](#); they show that increased copyright protection raises the offered quality.

Regarding illegal online sharing of music, recent empirical studies by [Oberholzer-Gee and Strumpf \(2007\)](#) and [Rob and Waldfogel \(2004\)](#) show a very limited effect of piracy on legal music sales.

3.2 Model

The digital good is demanded by a continuum of consumers on the unit interval $[0, 1]$. The gross utility of consumer i from a digital good is given by

$$\theta_i u(\lambda).$$

The valuation θ_i represents the willingness to pay for the digital good, whereas $\lambda \in [0, 1]$ represents the flexibility with which the digital good can be used by the consumer. The utility for flexibility $u(\lambda)$ is increasing and strictly concave with $u'(\lambda) \rightarrow \infty$ for $\lambda \rightarrow 0$ and $u'(1) = 0$. For simplicity, we shall assume that $\theta_i = i$ and that the consumers are uniformly distributed on the unit interval.

The seller of the digital good determines the price p and the level of flexibility λ at which the digital goods are sold to the consumers. The level of flexibility λ is the key choice variable in the seller's DRM design. For simplicity, we shall assume that the marginal cost of increasing flexibility is constant and equal to zero.⁴ The revenue of the seller is given by the product of the price p and the sold quantity $q \in [0, 1]$. With zero marginal cost, net profit is equal to the revenue, i.e.

$$\Pi(p, q) = pq.$$

Each consumer i can purchase the digital good at the offered price p and flexibility λ . The net utility of a purchase for consumer i is then

$$\theta_i u(\lambda) - p.$$

We refer to the digital good that is purchased from the seller as a *licensed product*.

In the presence of a “greynet,” a potential buyer can alternatively attempt to obtain the digital good unlicensed as a *pirated copy*. However, a consumer who doesn't buy the digital good cannot be certain of receiving a pirated copy. Instead,

⁴In the case of digital goods, the assumption of low marginal costs appears to be rather innocuous. We should point out, however, that, in the presence of DRM technology, there is a sense in which the cost of providing flexibility may not be constant or even monotone increasing. It might be most difficult technically to support intermediate levels of flexibility; very lenient or very strict DRM rules may be easier to implement.

a pirating consumer receives a copy only with a probability $\pi(\alpha, \lambda) \in [0, 1]$ so that the expected utility for consumer i of pirating is

$$\pi(\alpha, \lambda) \theta_i u(\lambda).$$

For simplicity we assume that $\pi(\alpha, \lambda) = \alpha\lambda$. The key idea is that the probability of receiving an pirated copy is increasing in the flexibility λ with which the licensed versions are sold. The parameter $\alpha \in [0, 1]$ represents an exogenous access rate to digital goods and characterizes the permeability of the content-distribution environment, not the good itself. We consider α to capture both technical and non-technical factors, so increased permeability can result, e.g., from factors such as higher internet bandwidth or contact frequency among consumers, or from more lenient copyright law or less vigilant enforcement of existing copyright laws. The probability of obtaining a pirated copy is therefore increasing both in the flexibility λ of the digital good itself as well in the permeability α of the environment. Finally, we assume that flexibility and permeability are complementary since

$$\frac{\partial^2 \pi(\alpha, \lambda)}{\partial \alpha \partial \lambda} > 0,$$

that is a higher level of permeability doesn't reduce the effect of flexibility and vice versa.

Later in the paper, we shall introduce the possibility of a platform in the form of a hardware device, a secure application program, or a secure hardware-software combination that is the only environment in which the content can be consumed. In this case, there will be an additional product that the consumers need to acquire

in order to be able to realize the utility from the digital goods. Yet, this will not affect the basic elements of demand for digital goods presented in the model.

3.3 Optimal Flexibility and Price

For a given flexibility λ and price p set by the provider of the digital good consumer i will decide to purchase a licensed copy if his net utility from a purchase is greater than his expected utility from pirating

$$\theta_i u(\lambda) - p \geq \alpha \lambda \theta_i u(\lambda).$$

The marginal buyer with valuation $\hat{\theta}$ is exactly indifferent between buying and pirating so $\hat{\theta}$ is given by

$$\hat{\theta} = \frac{p}{(1 - \alpha \lambda) u(\lambda)}.$$

Since all consumers with valuation $\theta_i \geq \hat{\theta}$ are buyers, the provider faces a demand function for licensed copies of

$$q(p, \lambda) = 1 - \frac{p}{(1 - \alpha \lambda) u(\lambda)} \quad (3.1)$$

We can see that the demand for the digital good is decreasing in p as would be expected. The more interesting comparative static is the effect of flexibility λ on demand.⁵

Proposition 3.1 (Effect of Flexibility) *The effect of flexibility on the demand for licensed copies of the digital good is ambiguous:*

⁵All proofs are relegated to the appendix.

- *In the absence of any piracy threat, for $\alpha = 0$, the demand is strictly increasing in the level of flexibility λ .*
- *With a threat of piracy, for $\alpha > 0$, demand is single-peaked in the level of flexibility λ , initially increasing but then decreasing.*

The fact that with the possibility of piracy the demand is single-peaked captures a key trade-off facing the seller of digital goods when deciding about the level of flexibility in his DRM design. An increase in flexibility leads to a higher value of the product for the consumers which has a positive effect on demand. Yet, at the same time the increase in flexibility leads to a higher likelihood of obtaining a pirated copy which has a negative effect on demand. Initially the increase in utility more than offsets the piracy threat and demand increases with flexibility, but since the marginal utility of any single consumer for flexibility is decreasing, it will ultimately be dominated by the easy access to pirated copies and lead to lower demand.

The revenue of the provider depends on the charged price p and the allowed flexibility λ , with:

$$\Pi(p, q) = pq(p, \lambda). \quad (3.2)$$

Maximizing this profit over p and λ leads to the following proposition.

Proposition 3.2 (Dealing with Piracy) *The threat of piracy has a strong effect on the optimal policy of the digital-good provider:*

- *For $\alpha = 0$, the provider chooses the efficient level of flexibility $\lambda_0^* = 1$ and sells the digital good at a price of $p_0^* = \frac{1}{2}u(1)$.*
- *For $\alpha > 0$, the provider sets flexibility to $\lambda^* < 1$ implicitly defined by*

$$(1 - \alpha\lambda^*) u'(\lambda^*) - \alpha u(\lambda^*) = 0$$

and sells the digital good at a price of $p^* = \frac{1}{2}(1 - \alpha\lambda^*)u(\lambda^*)$.

- *The optimal level of flexibility λ^* , the optimal price p^* and the provider's profit are decreasing in the threat of piracy α .*

Without the threat of piracy the provider acts like a standard monopolist and sets the flexibility at the highest possible level since it comes at no cost. Once the threat of piracy appears and α increases from zero, flexibility comes at a cost of decreased sales. Therefore the provider cuts back flexibility to reduce the probability of consumers obtaining a pirated copy to the point where the positive effect on demand is offset by the negative effect. Since the lower flexibility also reduces the utility buyers of licensed copies receive from the digital good, the provider has to also reduce the price. The higher the threat of piracy α , the more the provider cuts flexibility and price and the lower are his profits.

3.4 Sales of Platforms for Digital Goods

In the presence of the “greynet,” the provider – even though a monopolist – is constrained in capturing the utility that the consumers derive from the digital good. Because every consumer can always try to obtain unlicensed copies instead of buying licensed ones, the provider is forced by this outside option to leave an extra rent to all consumers. The provider of the digital good therefore faces the problem of recovering the residual surplus from the consumer. A feasible and common strategy in digital-content distribution is the provision of a platform on which to use the digital good. In the current section, we therefore introduce a second product, a platform that is required in order to use the digital good. In the case of digital

audio files, the immediate examples include digital music players such as Apple's iPod.

In economic terms, the platform constitutes a complimentary product to the digital good. In the presence of a platform, even the consumers who own unlicensed copies of the digital good have to buy the platform to consume the digital good. In other words, the platform does not create any additional value for the buyer over and above the consumption of the digital good. It simply represents a gatekeeper to the digital good. The platform owner can now recover some of the rent that the buyers obtained in the market for digital goods.

We denote by r the price of the platform. Now the utility consumer i receives from purchasing a licensed copy is given by

$$\theta_i u(\lambda) - p - r$$

and the utility consumer i receives from pirating the digital good is given by⁶

$$\alpha \lambda \theta_i u(\lambda) - r.$$

Conditional on purchasing the platform, the marginal buyer of the digital good with valuation $\hat{\theta}$ is indifferent between buying and pirating so $\hat{\theta}$ is given as before

$$\hat{\theta} = \frac{p}{(1 - \alpha \lambda) u(\lambda)}.$$

In addition, we now have to specify the marginal buyer of the platform with valuation $\bar{\theta}$. If the marginal buyer of the platform is a consumer who plans to pirate the

⁶We assume that consumers who choose to pirate have to purchase the platform before they know if they will obtain an unlicensed copy.

digital good, he is indifferent between purchasing the platform and not participating in the market at all

$$\alpha\lambda\bar{\theta}u(\lambda) - r = 0.$$

The marginal buyer of the platform is therefore given by

$$\bar{\theta} = \frac{r}{\alpha\lambda u(\lambda)}.$$

All consumers with $\theta_i \geq \bar{\theta}$ will purchase the platform and among these, all consumers with $\theta_i \geq \hat{\theta}$ will purchase licensed copies of the digital good so the demand function for the platform $Q(r, \lambda)$ and the demand function for the digital good are simply

$$Q(r, \lambda) = 1 - \frac{r}{\alpha\lambda u(\lambda)}$$

$$q(p, \lambda) = 1 - \frac{p}{(1 - \alpha\lambda) u(\lambda)}$$

The demand for the digital good is as before, decreasing in its price p and single peaked in its level of flexibility λ . The demand for the platform is also decreasing in its price r . In addition, however, it is strictly increasing in the flexibility λ of the digital good. We see that while a higher level of flexibility has an ambiguous effect on the demand for the digital good itself, it has a strictly positive effect on the demand for the platform.

Separate Firms

We first analyze the role of the platform in the context where the property rights to the platform technology and to the digital good are in the hands of separate firms.

In this case, a classic conflict arises between the platform provider and the content provider. In the case of separate providers, the provider of the digital good chooses his price p and flexibility λ to solve

$$\max_{p,\lambda} pq(p, \lambda),$$

while the provider of the platform chooses only his price r to solve

$$\max_r rQ(r, \lambda).$$

This leads to the following proposition.

Proposition 3.3 (Separate Firms) *If the digital good and the platform are sold by separate firms:*

- *The provider of the digital good behaves as in Proposition 3.2, setting flexibility λ^* and selling the digital good at a price of $p^* = \frac{1}{2}(1 - \alpha\lambda^*)u(\lambda^*)$.*
- *The provider of the platform takes the level of flexibility λ^* as given and sells the platform at a price of $r^* = \frac{1}{2}\alpha\lambda^*u(\lambda^*)$.*

Since the provider of the digital good doesn't take into account the effect his choice of flexibility has on the demand for the platform he behaves in the same way as if there were no platform. He chooses the level of flexibility that maximizes legal demand and then sets the monopolist price. The provider of the platform simply reacts to the level of flexibility chosen by the digital-good provider and sets his own price in accordance. While increasing flexibility has a purely positive effect on the profit of the platform provider (because it increases the value of access to the

digital good), the digital-good provider faces the trade-off between increasing the value of licensed copies and restricting the availability of unlicensed ones.

Integrated Firm

We now analyze the role of the platform in the context of a single firm that sells both the digital good and the platform. In other words, the seller has the property rights and controls the prices of the digital content as well as the platform. In this case the joint provider chooses price p and flexibility λ for the digital good and price r for the platform to solve

$$\max_{p,r,\lambda} \{pq(p, \lambda) + rQ(r, \lambda)\}.$$

This leads to the following proposition.

Proposition 3.4 (Integrated Firm) *If the digital good and the platform are sold by an integrated firm:*

- *The level of flexibility is higher than in the separate case, $\lambda^{**} > \lambda^*$.*
- *The price of the digital good is lower than in the separate case, $p^{**} < p^*$.*
- *The price of the platform is higher than in the separate case, $r^{**} > r^*$.*

The joint provider fully takes into account the effect of the digital good's flexibility on the demand for the good itself and on the demand for the platform. When increasing the level of flexibility beyond λ^* the provider loses sales of licensed copies to easier piracy but on the other hand he gains in platform sales. Since the first effect is initially of second order whereas the second effect is of first order, the

joint provider chooses a strictly greater level of flexibility for the digital good. Since flexibility is socially costless but valuable this implies an increase in welfare.

To recoup the lost sales of licensed copies the joint provider reduces the price below p^* so he ends up offering a higher level of flexibility at a lower price. He can afford to do so because at the same time he sells the platform at a higher price. The higher level of flexibility increases the value of the outside option available to each consumer in the form of piracy. By increasing the price of the platform the joint provider is able to appropriate part of that higher rent, making it indeed optimal to offer a higher level of flexibility.

3.5 Conclusion

In this paper, we provide a simple analysis of the role flexibility and platform play in digital rights management. The basic model shows that the optimal use of flexibility displays an important trade-off between providing a higher value to paying customers and increasing the likelihood of distribution through channels other than legitimate sales. We then show that a platform for the digital goods may lead to a socially beneficial improvement in the design of the flexibility rules if digital good and platform are owned by the same seller. However, if digital good and platform are complementary goods, but offered and priced by different sellers, then a conflict over the optimal flexibility rule emerges.

Our basic model had a number of simplifying features. Clearly, the analysis will have to be extended to better understand the emerging market structure and security provisions for digital goods. In many instances, content is available in many forms. Music, for example, is distributed through radio, TV, CDs, and digital copying. Because the demand for music in each market segment interacts with

the other segments, the distribution and management policies will naturally be dependent on the structure of the other market segments. We began with a single provider and a single platform, and it is logical to ask how DRM would be affected by competing providers and platforms.

On the demand side, it seems natural to think about the intensity of demand for digital goods and the ease with which unlicensed copies can be obtained. The music industry's concern about file sharing by students in college dormitories clearly arises in part from the fact that their best customers in terms of sales volume are the ones that have the best technology for accessing unlicensed copies.

Finally, as soon as flexibility becomes an issue, more sophisticated pricing strategies seem natural. In this paper, we focused on the single-file pricing policy, but other plans are clearly being used or conceived to find an optimal trade-off. For example, monthly fees for limited or unlimited access to databases of music files are alternatives to single-file transactions.

3.6 Appendix

Proof of Proposition 3.1. Differentiating the expression for $q(p, \lambda)$ in (3.1) we get

$$\frac{\partial q(p, \lambda)}{\partial \lambda} = p \frac{(1 - \alpha\lambda) u'(\lambda) - \alpha u(\lambda)}{(1 - \alpha\lambda)^2 u(\lambda)^2}$$

Since u is increasing and concave, for small values of λ the term $(1 - \alpha\lambda) u'(\lambda)$ is big and the term $\alpha u(\lambda)$ is small so that $\partial q(p, \lambda) / \partial \lambda > 0$. As λ increases the first term decreases and the second increases, eventually leading to $\partial q(p, \lambda) / \partial \lambda < 0$. \square

Proof of Proposition 3.2. Maximizing profit given by (3.2), the first order condition with respect to λ is given by

$$\begin{aligned} p^2 \frac{(1 - \alpha\lambda) u'(\lambda) - \alpha u(\lambda)}{(1 - \alpha\lambda^*)^2 u(\lambda^*)^2} &= 0, \\ \Rightarrow (1 - \alpha\lambda^*) u'(\lambda^*) - \alpha u(\lambda^*) &= 0. \end{aligned} \quad (3.3)$$

The first order condition with respect to p yields

$$\begin{aligned} 1 - 2 \frac{p}{(1 - \alpha\lambda) u(\lambda)} &= 0, \\ \Rightarrow p^* &= \frac{1}{2} (1 - \alpha\lambda^*) u(\lambda^*). \end{aligned} \quad (3.4)$$

which results in a demand of $q(p^*, \lambda^*) = 1/2$. Implicit differentiation of (3.3) gives us the comparative static of λ^* with respect to α :

$$\frac{d\lambda^*}{d\alpha} = \frac{\lambda^* u'(\lambda^*) + u(\lambda^*)}{(1 - \alpha\lambda^*) u''(\lambda^*) - 2\alpha u'(\lambda^*)} < 0$$

Differentiating (3.4) using the envelope theorem we get

$$\frac{dp^*}{d\alpha} = -\frac{1}{2} \lambda^* u(\lambda^*) < 0.$$

Finally, this implies for the provider's profit

$$\frac{d\Pi(p^*, \lambda^*)}{d\alpha} = \frac{1}{2} \frac{dp^*}{d\alpha} < 0.$$

□

Proof of Proposition 3.3. The provider of the digital good faces the same problem as before, resulting in the first order conditions (3.3) and (3.4). The provider of the platform has a first order condition with respect to r given by

$$1 - 2\frac{r}{\alpha\lambda u(\lambda)} = 0. \quad (3.5)$$

Given the equilibrium level of flexibility λ^* this results in

$$r^* = \frac{1}{2}\alpha\lambda^*u(\lambda^*).$$

□

Proof of Proposition 3.4. Differentiating the joint provider's profit with respect to λ we get

$$p^2 \frac{(1 - \alpha\lambda) u'(\lambda) - \alpha u(\lambda)}{(1 - \alpha\lambda)^2 u(\lambda)^2} + r^2 \frac{\alpha\lambda u'(\lambda) + \alpha u(\lambda)}{(\alpha\lambda)^2 u(\lambda)^2} \quad (3.6)$$

At the level of flexibility λ^* chosen by the provider of the digital good in the separate case, the first term of (3.6) is zero but the second is positive. Therefore a joint provider will choose a higher level of flexibility $\lambda^{**} > \lambda^*$. The first order conditions with respect to p and r are identical to (3.4) and (3.5) respectively. Note that the resulting expression for p ,

$$p = \frac{1}{2}(1 - \alpha\lambda)u(\lambda),$$

is maximized at λ^* given by (3.3) so it has to be lower for λ^{**} and therefore $p^{**} < p^*$.

The expression for r ,

$$r = \frac{1}{2}\alpha\lambda u(\lambda),$$

is increasing in λ so it has to be higher for λ^{**} and therefore $r^{**} > r^*$.

□

Bibliography

- ACHARYA, V. V. (2009): “A Theory of Systemic Risk and Design of Prudential Bank Regulation,” *Journal of Financial Stability*, 5(3), 224–255.
- ADAMS, J. K., AND P. A. ADAMS (1961): “Realism of Confidence Judgments,” *Psychological Review*, 68(1), 33–45.
- AKERLOF, G. A., AND W. T. DICKENS (1982): “The Economic Consequences of Cognitive Dissonance,” *American Economic Review*, 72(3), 307–319.
- AKERLOF, G. A., AND R. J. SHILLER (2010): *Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism*. Princeton University Press.
- ALPERT, M., AND H. RAIFFA (1982): “A Progress Report on the Training of Probability Assessors,” in *Judgment Under Uncertainty: Heuristics and Biases*, ed. by D. Kahneman, P. Slovic, and A. Tversky. Cambridge University Press.
- ANGELETOS, G.-M., C. HELLWIG, AND A. PAVAN (2007): “Dynamic Global Games of Regime Change: Learning, Multiplicity, and the Timing of Attacks,” *Econometrica*, 75(3), 711–756.
- BAKOS, Y., E. BRYNJOLFSSON, AND D. LICHTMAN (1999): “Shared Information Goods,” *Journal of Law and Economics*, 42(1), 117–156.

- BARBER, B. M., AND T. ODEAN (2008): “All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors,” *Review of Financial Studies*, 21(2), 785–818.
- BEN-DAVID, I., J. R. GRAHAM, AND C. R. HARVEY (2007): “Managerial Overconfidence and Corporate Policies,” Discussion Paper 13711, National Bureau of Economic Research.
- (2010): “Managerial Miscalibration,” Discussion Paper 16215, National Bureau of Economic Research.
- BERNARD, V. L., AND J. K. THOMAS (1989): “Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium?,” *Journal of Accounting Research*, 27, 1–36.
- BIDDLE, P., P. ENGLAND, M. PEINADO, AND B. WILLMAN (2002): “The Darknet and the Future of Content Distribution,” in *Proceedings of the 2nd ACM Workshop on Digital Rights Management*.
- BÉNABOU, R., AND J. TIROLE (2002): “Self-Confidence and Personal Motivation,” *Quarterly Journal of Economics*, 117(3), 871–915.
- BRUNNERMEIER, M. K., AND M. OEHMKE (2010): “The Maturity Rat Race,” Working Paper 16607, National Bureau of Economic Research.
- BRUNNERMEIER, M. K., AND L. H. PEDERSEN (2009): “Market Liquidity and Funding Liquidity,” *Review of Financial Studies*, 22(6), 2201–2238.
- CALOMIRIS, C. W., AND C. M. KAHN (1991): “The Role of Demandable Debt in Structuring Optimal Banking Arrangements,” *American Economic Review*, 81(3), 497–513.

- CARLSSON, H., AND E. VAN DAMME (1993a): "Equilibrium Selection in Stag Hunt Games," in *Frontiers of Game Theory*, ed. by K. Binmore, A. Kirman, and P. Tani. MIT Press.
- (1993b): "Global Games and Equilibrium Selection," *Econometrica*, 61(5), 989–1018.
- CHENG, I.-H., AND K. W. MILBRADT (2010): "The Hazards of Debt: Rollover Freezes, Incentives, and Bailouts," mimeo.
- CONNER, K. R., AND R. P. RUMELT (1991): "Software Piracy: An Analysis of Protection Strategies," *Management Science*, 37(2), 125–139.
- DANIEL, K., D. HIRSHLEIFER, AND A. SUBRAHMANYAM (1998): "Investor Psychology and Security Market Under- and Overreactions," *Journal of Finance*, 53(6), 1839–1885.
- DIAMOND, D. W., AND R. G. RAJAN (2001): "Liquidity Risk, Liquidity Creation, and Financial Fragility: A Theory of Banking," *Journal of Political Economy*, 109(2), 287–327.
- EPSTEIN, L. G., AND I. KOPYLOV (2007): "Cold Feet," *Theoretical Economics*, 2(3), 231–259.
- FARHI, E., AND J. TIROLE (2010): "Collective Moral Hazard, Maturity Mismatch and Systemic Bailouts," *American Economic Review*, forthcoming.
- FENZ, W. D., AND S. EPSTEIN (1967): "Gradients of Physiological Arousal in Parachutists as a Function of an Approaching Jump," *Psychosomatic Medicine*, 29(1), 33–51.

- FENZ, W. D., AND G. B. JONES (1972): "Individual Differences in Physiologic Arousal and Performance in Sport Parachutists," *Psychosomatic Medicine*, 34(1), 1–8.
- FISCHHOFF, B., P. SLOVIC, AND S. LICHTENSTEIN (1977): "Knowing with Certainty: The Appropriateness of Extreme Confidence," *Journal of Experimental Psychology: Human Perception and Performance*, 3(4), 552–564.
- FREUD, S. (2008): *Zur Psychopathologie des Alltagslebens*. Echo Library.
- FUDENBERG, D., AND D. K. LEVINE (2010): "Risk, Delay, and Convex Self-Control Costs," mimeo.
- GOLDSTEIN, I., AND A. PAUZNER (2005): "Demand-Deposit Contracts and the Probability of Bank Runs," *Journal of Finance*, 60(3), 1293–1327.
- GRUBB, M. D. (2009): "Selling to Overconfident Consumers," *American Economic Review*, 99(5), 1770–1807.
- GRUBER, M. J. (1996): "Another Puzzle: The Growth in Actively Managed Mutual Funds," *Journal of Finance*, 51(3), 783–810.
- GUENTHER, R. K. (1988): "Mood and Memory," in *Memory in Context: Context in Memory*, ed. by G. M. Davies, and D. M. Thomson. Wiley.
- HALEVY, Y. (2008): "Strotz Meets Allais: Diminishing Impatience and the Certainty Effect," *American Economic Review*, 98(3), 1145–62.
- HARRIS, L. (2003): *Trading and Exchanges: Market Microstructure for Practitioners*. Oxford University Press.
- HE, Z., AND W. XIONG (2010): "Dynamic Debt Runs," mimeo.

- HOFFRAGE, U. (2004): "Overconfidence," in *Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgement and Memory*, ed. by R. Pohl. Psychology Press.
- HOLMES, D. S. (1995): "The Evidence for Repression: An Examination of Sixty Years of Research," in *Repression and Dissociation: Implications for Personality Theory, Psychopathology and Health*, ed. by J. L. Singer. University of Chicago Press.
- HOU, K., L. PENG, AND W. XIONG (2009): "A Tale of Two Anomalies: The Implications of Investor Attention for Price and Earnings Momentum," mimeo.
- HVIDE, H. K. (2002): "Pragmatic Beliefs and Overconfidence," *Journal of Economic Behavior & Organization*, 48(1), 15–28.
- JENSEN, M. C., AND W. H. MECKLING (1976): "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics*, 3(4), 305–360.
- JONES, E. E., AND C. A. JOHNSON (1973): "Delay of Consequences and the Riskiness of Decisions," *Journal of Personality*, 41(4), 613–637.
- KAHNEMAN, D., AND A. TVERSKY (1973): "On the Psychology of Prediction," *Psychological Review*, 80(4), 237–251.
- KARLSSON, N., G. LOEWENSTEIN, AND D. SEPPI (2005): "The 'Ostrich Effect': Selective Attention to Information about Investments," mimeo.
- KEREN, G., AND P. ROELOFSMA (1995): "Immediacy and Certainty in Intertemporal Choice," *Organizational Behavior and Human Decision Processes*, 63(3), 287–297.

- KREPS, D. M., AND E. L. PORTEUS (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46(1), 185–200.
- KRISHNAMURTHY, A. (2010): "How Debt Markets Have Malfunctioned in the Crisis," *Journal of Economic Perspectives*, 24(1), 3–28.
- LAMONT, O., AND A. FRAZZINI (2007): "The Earnings Announcement Premium and Trading Volume," Working Paper 13090, National Bureau of Economic Research.
- LANGER, E. (1975): "The Illusion of Control," *Journal of Personality and Social Psychology*, 32(2), 311–328.
- LEE, C. M. C. (1992): "Earnings News and Small Traders: An Intraday Analysis," *Journal of Accounting and Economics*, 15(2–3), 265–302.
- LELAND, H. E., AND K. B. TOFT (1996): "Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads," *Journal of Finance*, 51(3), 987–1019.
- LIEBOWITZ, S. J. (1985): "Copying and Indirect Appropriability: Photocopying of Journals," *Journal of Political Economy*, 93(5), 945–957.
- LO, A. W., AND D. V. REPIN (2002): "The Psychophysiology of Real-Time Financial Risk Processing," *Journal of Cognitive Neuroscience*, 14(3), 323–339.
- LO, A. W., D. V. REPIN, AND B. N. STEENBARGER (2005): "Fear and Greed in Financial Markets: A Clinical Study of Day-Traders," *American Economic Review*, 95(2), 352–359.

- LOEWENSTEIN, G. F., E. U. WEBER, C. K. HSEE, AND N. WELCH (2001): "Risk as Feelings," *Psychological Bulletin*, 127(2), 267–286.
- MONAT, A., AND R. S. LAZARUS (1991): *Stress and Coping: An Anthology*. Columbia University Press.
- MORRIS, S., AND H. S. SHIN (2003): "Global Games: Theory and Applications," in *Advances in Economics and Econometrics (Proceedings of the Eighth World Congress of the Econometric Society)*, ed. by M. Dewatripont, L. P. Hansen, and S. J. Turnovsky. Cambridge University Press.
- (2004): "Coordination Risk and the Price of Debt," *European Economic Review*, 48(1), 133–153.
- MORRIS, W. N., AND N. P. REILLY (1987): "Toward the Self-Regulation of Mood: Theory and Research," *Motivation and Emotion*, 11(3), 215–249.
- NOUSSAIR, C., AND P. WU (2006): "Risk Tolerance in the Present and the Future: An Experimental Study," *Managerial and Decision Economics*, 27(6), 401–412.
- NOVOS, I. E., AND M. WALDMAN (1984): "The Effects of Increased Copyright Protection: An Analytic Approach," *Journal of Political Economy*, 92(2), 236–246.
- OBERHOLZER-GEE, F., AND K. STRUMPF (2007): "The Effect of File Sharing on Record Sales: An Empirical Analysis," *Journal of Political Economy*, 115(1), 1–42.
- ODEAN, T. (1999): "Do Investors Trade Too Much?," *American Economic Review*, 89(5), 1279–1298.

- ONCULER, A. (2000): "Intertemporal Choice under Uncertainty: A Behavioral Perspective," Working Paper 2000/37/TM, INSEAD.
- OSKAMP, S. (1965): "Overconfidence in Case-Study Judgements," *Journal of Consulting Psychology*, 29(3), 261–265.
- PARKER, G. G., AND M. W. VAN ALSTYNE (2005): "Two-Sided Network Effects: A Theory of Information Product Design," *Management Science*, 51(10), 1494–1504.
- PATERSON, R. J., AND R. W. J. NEUFELD (1987): "Clear Danger: Situational Determinants of the Appraisal of Threat," *Psychological Bulletin*, 101(3), 404–416.
- PEARLIN, L. I., AND C. W. RADABAUGH (1976): "Economic Strains and the Coping Function of Alcohol," *American Journal of Sociology*, 82(3), 652–663.
- PRATT, J. W. (1964): "Risk Aversion in the Small and in the Large," *Econometrica*, 32(1/2), 122–136.
- RAJAN, R. G. (1992): "Insiders and Outsiders: The Choice between Informed and Arm's-Length Debt," *Journal of Finance*, 47(4), 1367–1400.
- REINHART, C. M., AND K. S. ROGOFF (2009): *This Time is Different: Eight Centuries of Financial Folly*. Princeton University Press.
- ROB, R., AND J. WALDFOGEL (2004): "Piracy on the High C's: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students," Working Paper 10874, National Bureau of Economic Research.

- ROCHET, J.-C., AND X. VIVES (2004): "Coordination Failures and the Lender of Last Resort: Was Bagehot Right After All?," *Journal of the European Economic Association*, 2(6), 1116–1147.
- ROTH, W. T., G. BREIVIK, P. E. JØRGENSEN, AND S. HOFMANN (1996): "Activation in Novice and Expert Parachutists While Jumping," *Psychophysiology*, 33(1), 63–72.
- SAGRISTANO, M. D., Y. TROPE, AND N. LIBERMAN (2002): "Time-Dependent Gambling: Odds Now, Money Later," *Journal of Experimental Psychology: General*, 131(3), 364–376.
- SCHEINKMAN, J. A., AND W. XIONG (2003): "Overconfidence and Speculative Bubbles," *Journal of Political Economy*, 111(6), 1183–1219.
- SCHELLING, T. C. (1984): "Self-Command in Practice, in Policy, and in a Theory of Rational Choice," *American Economic Review*, 74(2), 1–11.
- SHELLEY, M. K. (1994): "Gain/Loss Asymmetry in Risky Intertemporal Choice," *Organizational Behavior and Human Decision Processes*, 59(1), 124–159.
- SHLEIFER, A., AND R. W. VISHNY (1992): "Liquidation Values and Debt Capacity: A Market Equilibrium Approach," *Journal of Finance*, 47(4), 1343–1366.
- SIRRI, E. R., AND P. TUFANO (1998): "Costly Search and Mutual Fund Flows," *Journal of Finance*, 53(5), 1589–1622.
- SUNDARARAJAN, A. (2004): "Managing Digital Piracy: Pricing and Protection," *Information Systems Research*, 15(3), 287–308.

- TAKEYAMA, L. N. (1994): "The Welfare Implications of Unauthorized Reproduction of Intellectual Property in the Presence of Demand Network Externalities," *Journal of Industrial Economics*, 42(2), 155–166.
- (1997): "The Intertemporal Consequences of Unauthorized Reproduction of Intellectual Property," *Journal of Law and Economics*, 40(2), 511–522.
- VAN DEN STEEN, E. (2004): "Rational Overoptimism (and Other Biases)," *American Economic Review*, 94(4), 1141–1151.
- VARIAN, H. R. (2000): "Buying, Sharing and Renting Information Goods," *Journal of Industrial Economics*, 48(4), 473–488.
- WANG, Z., AND J. WERNER (1994): "Portfolio Characterization of Risk Aversion," *Economics Letters*, 45(2), 259–265.
- WEINSTEIN, N. D. (1980): "Unrealistic Optimism About Future Life Events," *Journal of Personality and Social Psychology*, 39(5), 806–820.
- ZELLER, A. F. (1950): "An Experimental Analogue of Repression," *Psychological Bulletin*, 47(1), 39–51.