ISSN 0956-8549-632

Rents, learning and risk in the financial sector and other innovative industries

By

Bruno Biais
Jean-Charles Rochet
Paul Woolley

THE PAUL WOOLLEY CENTRE
WORKING PAPER SERIES NO 4
FMG DISCUSSION PAPER 632

DISCUSSION PAPER SERIES

September 2009

Bruno Biais is research professor of economics and finance at the Toulouse School of Economics (IDEI, CNRS). He is currently editor of the Review of Economic Studies. Bruno Biais has been scientific adviser to Euronext, EuroMTS, the London Investment Banking Association and the New York Stock Exchange. He held the Deutsche Bank visiting chair of finance at Oxford University, was Wim Duisenberg Fellow at the European Central Bank, and gave the Rotman Distinguished Lectures at the University of Toronto. Jean - Charles Rochet holds a Ph.D.in Mathematical Economics from Paris – Dauphine University . He has taught in Paris, France (Dauphine University, ENSAE and Ecole Polytechnique) and London, U.K. (B.P. visiting professor, London School of Economics, 2001-02). He is a member of the Institut Universitaire de France and a Fellow of the Econometric Society since 1995. He has also been council member of European Economic Association, and associate editor of Econometrica . He is currently Professor of Economics at the Toulouse School of Economics and Research Director at Institut D'Economie Industrielle. Paul Woolley's career has spanned the private sector, academia and policy-oriented institutions. He has worked at the International Monetary Fund in Washington and as a director of the merchant bank Baring Brothers. For the past 20 years he has run the European arm of GMO, the global fund management business based in Boston, US. Dr Woolley funded The Paul Woolley Centre for the Study of Capital Market Dysfunctionality. He is chairman of the Advisory Board for the Centre and is a full time member of the research team.

Rents, learning and risk in the financial sector and other innovative industries

Bruno Biais[1], Jean–Charles Rochet[2] and Paul Woolley[3]

September 15, 2009

This research was conducted within the Paul Woolley Research Initiative on Capital Market Dysfunctionalities at IDEI, Toulouse. Many thanks to participants in the first conference of the Centre for the Study of Capital Market Dysfunctionality at the London School of Economics, June 2008, the Pompeu Fabra Conference on the Financial Crisis, May 2009, and the third Banco de Portugal conference on Financial Intermediation, June 2009, as well as seminar participants at London Business School, Frankfurt University and Amsterdam University, especially Sudipto Bhattacharya, Arnoud Boot, John Boyd, Markus Brunnermeier, Catherine Casamatta, Zvi Eckstein, Guido Friebel, Alex Gümbel, Philipp Hartmann, Augustin Landier, John Moore, Liliana Pellizzon, Enrico Perotti, Ludovic Phalippou, Guillaume Plantin and Steve Schaefer.

[1] Toulouse School of Economics (CNRS-GREMAQ & PWRI-IDEI)
[2] Toulouse School of Economics (CNRS-GREMAQ & PWRI-IDEI)
[3] London School of Economics (PWC)

Abstract

We study innovative industries subject to two risks. First, it is uncertain whether the innovation is strong or fragile. Second, it is difficult to monitor managers, which creates moral hazard and agency rents. As time goes by and profits are observed, beliefs about the industry are updated. As long as no default occurs, confidence builds up. Initially this spurs growth. But increasingly confident managers end up requesting large rents, curbing the growth of the industry. If rents become too high, investors give up on incentives, and failure rates rise. If the innovation is fragile, eventually there is a crisis. Our model captures stylized facts of the recent financial innovation wave and generates new implications for risks, returns and rents.

Keywords: Innovative industry, financial sector, learning, moral hazard, rents, boom, crisis.

1 Introduction

Innovations can generate booms. Such developments are risky when the fundamental profitability of the innovation is unknown. Furthermore, the actions and performance of managers are particularly difficult to monitor in innovative industries, with which final investors are not familiar. This creates a moral hazard problem. As the industry spurred by the innovation grows, its potential value develops, but its riskiness and the agency problem can worsen. And if the risk eventually materializes, crises can take place. The goal of this paper is to model this evolution in a rational expectations equilibrium framework.

A prominent example of such dynamics is offered by the financial sector at the beginning of this century. Technological advances in information technologies, combined with market liberalization, spurred many innovations, both in terms of techniques, e.g., structured finance, and in terms of strategies and institutions, e.g., hedge funds and private equity funds. Innovation generated remarkable growth, in capital invested, skilled workforce and market valuation.[4] Significant profits were generated. Nevertheless, doubts remained. Were the financial innovations valuable enough to justify the allocation of so much financial and human capital? Did the fund managers and investment bankers really exert the right amount of effort to control risks and did they make the right decisions to ensure that investment would be valuable? Were the level of compensation to financial managers and their growth sustainable? Or did rents absorb all the value created by the industry?

We investigate these issues in the context of an agency model with learning. There are two sectors: the traditional one and the innovative one. There is a continuum of competitive managers, who choose to work in the innovative industry if the compensation it offers exceeds their wage in the traditional sector. For simplicity we assume the traditional sector is well known, so that agency problems and uncertainty are not an issue there. Thus managers working in that sector receive wages equal to their known productivity. The innovative sector differs from the traditional one, both in terms of opportunities and in terms of challenges.

[4] Of course other forces, not captured by our model, contributed to the growth of the financial sector in the early 2000's, such as, e.g., global imbalances or the real estate bubble. For clarity we focus on one aspect only: the link between innovation, learning and information asymmetry.

- First, there is uncertainty.[5] The innovative sector can be strong and potentially very profitable. Alternatively it can be fragile and exposed to the risk of a crisis. Observing the dynamics of realized profits in the innovative sector, agents conduct Bayesian learning about the strength of that industry.

- Second, there is asymmetric information. In the innovative sector it is not easy to monitor how managers use the resources they have been entrusted with. The managers may be tempted to use these resources in ways that are not optimal for the final investors. Thus, this sector is plagued with moral hazard, which we model in the line of Holmstrom and Tirole (1997): when managers fail to exert effort, they earn private benefits from shirking, but the risk of failure is higher than under effort.

Our key assumption is that the nature of the moral hazard problem is related to the fundamental value of the innovation. We assume that if the innovation is strong managerial shirking does not reduce very much the probability of success, while if the innovation is fragile shirking significantly increases the risk of failure. To motivate this assumption consider the example of the recent credit risk transfer innovations. The initial uncertainty was about the extent to which these innovations would enable banks to reduce their exposure to the risk of defaults. A strong innovation is very effective at shedding this risk. A fragile one leaves banks exposed to high risks. In this context, consider the consequences of effort. If bankers exert high effort, they screen out most bad projects. In that case, whether the innovation is strong or fragile, they are exposed to limited default risk. In contrast, if banks shirk and fail to scrutinize loan applicants carefully, they remain relatively safe when the innovation is strong, but are exposed to high default risk when it is fragile.

Ironically, our assumption that the adverse consequences of shirking are greater when the innovation is fragile implies that the moral hazard problem is more severe when the industry is strong. Indeed, when the adverse consequences of shirking on success are limited, shirking is attractive for the manager. Hence agency rents are higher. In that context, our model yields the following equilibrium dynamics:

[5] Uncertainty about the innovation could be modelled in terms of uncertainty aversion, as in Caballero and Krishnamurthy (2008). Our approach, set expected utility – Bayesian learning terms, complements theirs.

- **Boom:** In the early phase of the innovation wave, if there is no crisis, confidence builds up. Correspondingly the innovative sector attracts an increasing fraction of capital and labor.

- **Choking:** If performance continues, managers become very confident. Hence they demand large rents. These limit the net returns left to investors and slow down the growth of the innovative sector.

- **Risk–taking:** At some point, the rents demanded by the managers become so high that investors give up on incentives. Consequently, managers exert less effort, and risk and failure rates go up.

- **Bust:** At any point, if the industry is fragile, there can be a crisis.[6] When it occurs, agents revise downwards their expectations about the profitability of that sector. Consequently, investment, wages and employment in the innovative industry drop.

The boom and bust arising in our model stem from Bayesian learning about the innovation, as in Zeira (1999), Barbarino and Jovanovic (2007) and Pastor and Veronesi (2006). In contrast, the choking and risk–taking regimes are due to the interaction between the moral hazard and learning problems we analyze, and are unique to our analysis.

The equilibrium dynamics our model generates are in line with some of the major stylized facts of the financial innovation wave of the early 2000s and its eventual bust:

- **Uncertainty, learning and growth:** Before the crisis, some analysts claimed that new risk management techniques had very significantly reduced (if not eliminated) the macro–risk associated with the financial sector. For example, in a speech delivered in 2005, A. Greenspan claimed that: "As is generally acknowledged, the development of credit derivatives has contributed to the stability of the banking system by allowing banks, especially the largest systematically important banks, to measure and manage their credit risk more effectively". In the language of our model, this expresses the belief that the innovative industry was strong. As in our model, such optimistic beliefs were motivated by a series of years with high profits and no large financial distress. As in our equilibrium, these developments were accompanied by an increase in the

[6] While the Boom, Choking and Risk–Taking regimes arise both for solid and for fragile innovations, crises take place only when the innovation is fragile.

size of the financial sector. But some uncertainty remained about the strength of the financial innovation wave. For example Rajan (2006) wrote that the "increase in the risk bearing capacity of economies, as well as in actual risk taking, has led to a range of financial transactions that hitherto were not possible. As a result, under some conditions, economies may be more exposed to financial-sector-induced turmoil than in the past." In the language of our model, this corresponds to the awareness that, with some probability, the innovative industry was fragile and prone to crisis.

- **Managers' rents and investors' net returns:** Our theoretical analysis implies that during the successful stages of financial innovation waves, managers operating in the innovative industry earn increasing rents. This is consistent with the empirical results of Philippon and Reshef (2008). They find that, during the recent financial innovation wave, the compensation of managers in the financial sector increased relative to other sectors. Furthermore, they estimate that part of that increase reflected rents. Our model also implies that, as long as there is no crisis, the relationship between the amount of money invested in the innovative sector and its past performance should be inverse U–shaped, which is consistent with the empirical results of Ramadorai (2008). The economic intuition of our theoretical result is that, initially, the increase in confidence due to good records attracts funds, while, later, confidence inflates managers' rents, undermining the net returns of investors.

The next section discusses the literature on innovative industries, learning and moral hazard and the contribution of our analysis relative to these papers. Section 3 presents our model. Section 4 presents the benchmark case where effort is observable. It shows how booms and crashes can arise, because of uncertainty and learning, without moral hazard. Section 5 turns to the case where managerial effort is unobservable. It shows how moral hazard interacts with learning and gives rise to choking and risk–taking. Section 6 discusses empirical implications of our model and confronts them with evidence on the financial sector. Section 7 discusses the policy implications of our theoretical analysis. Section 8 concludes. Proofs omitted in the text are in the appendix.

2 Literature

Our analysis on the role of uncertainty and learning in booms and crashes in a Bayesian expected utility framework is in line with the contributions of Zeira (1987, 1999), Rob (1991), Pastor and Veronesi (2006), and Barbarino and Jovanovic (2007). These papers offer rational expectations models where, under learning, an industry can develop as long as good news are observed and then crash with bad news. Zeira (1999) shows that his theoretical model applies to the booms and crashes of 1929 and 1987. Pastor and Veronesi (2006), and Barbarino and Jovanovic (2007) show that their models offer insights into the boom and crash of the telecommunications and internet sector in the late 1990s.

One contribution of our analysis relative to these papers is to introduce moral hazard and study how it affects the dynamics of the innovative industry.[7] This enables us to generate new results relative to the previous literature: We study the interaction between moral hazard and learning. We show how rents arise in equilibrium. We explain how the evolution of these rents affects the dynamics of the division of the value between managers and investors. We also show how moral–hazard can lead to excessive risk–taking at a systemic level.[8]

Our paper is also related to the literature analyzing moral hazard and innovation, in particular Bergemann and Hege (1998, 2005). Their focus is quite different from ours, however. Bergemann and Hege (1998, 2005) consider long lived managers, which is beyond the scope of our one–period contracting framework. But, while Bergemann and Hege (2008) study the interaction between two players only, we consider populations of investors and managers and characterize the allocation of human and physical capital to the innovative and traditional sectors. Also, in Bergemann and Hege (1998, 2005) good past performance alleviates the moral hazard problem, while in our analysis it can increase agency rents. This gives rise to the choking and risk–taking regimes, which are unique to our paper.

He and Krishnamurthy (2008) also study long lived agents and moral hazard, but they do not analyze learning. A common theme between their paper and ours is that two possible regimes arise in

[7] In that respect, while our analysis can shed some light on several innovation waves, including the internet boom, our emphasis on moral hazard makes our model particularly relevant for the recent evolution of the financial industry where agency problems were key.

[8] Risk–taking in our equilibrium is systemic because all bankers engage in it simultaneously, and the corresponding failures affect a large fraction of the economy. But, in our model, there are no externalities or informational contagion.

equilibrium: one in which the incentive constraint does not bind and the other where it does, giving rise to rationing. In He and Krishnamurthy (2008) the market subject to these quantity constraints is the securities markets, while in our analysis it is the labour market for managers.

Our work is also in line with Philippon (2008) who analyzes the interaction between the financial and non–financial sectors. But, in his model, agency problems plague the real sector, and monitoring by financial intermediaries mitigates the moral hazard problem facing manufacturing firms by monitoring them. In contrast, our analysis assumes moral hazard in the innovative industry, of which the financial sector at the turn of this century offers a prominent example.

Finally, our analysis is also related to Lorenzoni (2008), who offers a model of credit booms with agency problems and aggregate shocks, where credit expansion can be followed by contractions. His focus on pecuniary externalities is similar to that of Brunermeier and Sannikov (2009), who analyze moral hazard in the financial sector. They study the amplification dynamics by which bad shocks lead to drops in asset prices which worsen the consequences of the shock.

3 The model

3.1 Agents and goods

Consider an infinite horizon economy, operating in discrete time at periods $t = 1, 2, ...$ There is a mass one continuum of competitive managers and a mass one continuum of competitive investors. All players are risk neutral. The managers have limited liability and no initial wealth. At the beginning of each period, each investor is endowed with one unit of nonstorable investment good. At the end of each period, all agents consume the consumption good, produced, as explained below, by labor and capital. For simplicity, we focus on the simplest possible case, where managers live and contract only for one period.[9] Yet, the model is dynamic, to the extent that agents progressively learn about the strength of the innovative industry. Thus, the link between generations runs through the evolution of beliefs.

[9] When players interact for several periods, shirking can create a wedge between the beliefs of investors and managers. Bergemann and Hege (1998, 2005) and DeMarzo and Sannikov (2008) offer insightful analyses of this problem.

3.2 The two sectors

3.2.1 Managers and investors

There are two sectors, the traditional sector and the innovative one. Managers and investors are heterogeneous. Their types are denoted by ν and ρ respectively. The types of managers are distributed over $[0, \bar{\nu}]$. Their cumulative distribution function is denoted by G. The types of investors are distributed over $[0, \bar{\rho}]$. Their cumulative distribution function is denoted by F. Managers and investors choose in which sector to operate. For simplicity we assume that, in the traditional sector, agents generate output equal to their type. Thus, when a type ν manager operates in the traditional sector, he obtains, at the end of the period, ν units of the consumption good. Similarly, when investor ρ allocates her unit endowment of investment good to the traditional sector, she obtains ρ units of the consumption good at the end of the period.

Operating the innovative technology requires one unit of investment as well as one manager. Capital is provided by the investors, who are endowed with the investment good at the beginning of the period. Managers, who can't work on more than one project, are hired by investors. For example the agent could be an entrepreneur undertaking a venture in the information and communication technology sector. Or the manager could be an investment banker, using the capital to fund innovative financial engineering operations, e.g., in structured finance.

3.2.2 Uncertainty

There is uncertainty about the innovative industry. Consider for example the resecuritization of subprime mortgages obtained through pooling and tranching of Residential Mortgage Backed Securities into CDOs and other structured finance products. It was not initially clear whether or not structured finance techniques could reduce exposure to defaults and improve the allocation of risk without generating perverse effects. Analogous techniques had worked well previously for other types of assets, such as credit card receivables. But this new financial innovation wave had novel and yet untested characteristics, and there was uncertainty about how it could perform in the long run. While it was not obvious ex–ante, it eventually turned out that the innovation involved intrinsic flaws, due to the interdependence between financial institutions, counterparty risk and its interaction with the real

estate bubble.

We model this uncertainty by assuming that the innovative industry can either be strong, which is denoted by $\theta = 1$, or fragile, which is denoted by $\theta = 0$. If the industry is strong, it is immune from crises. But, if the industry is fragile, at each period, with probability $1 - p$, it can be hit by a crisis. In that case, all projects in the innovative sector generate 0 output. Thus, fragility implies that the industry is exposed to the risk of a large crisis eliminating its profits.

3.2.3 Moral hazard

When there is no crisis, the innovative technology generates Y or 0. The distribution of output in that case depends on the effort of the managers. Managerial effort leads to an improvement in the distribution of output in the sense of first order stochastic dominance. This can be interpreted as an increase in the probability of success, e.g., in high–tech industries the manager can exert effort to increase the probability that a new patent will be obtained and a new product or service will be commercialized. Effort can also be interpreted in terms of risk prevention, e.g., in the financial sector managers can exert effort to screen investment opportunities, and avoid those with a large risk of default.

When there is no crisis, for each firm in the innovative sector, if the manager exerts effort, output Y is generated for sure. But if he exerts no effort, there is a risk that output will be 0. When the innovative industry is strong, the probability of default under shirking, denoted by $\underline{\Delta}$, is relatively small. When the industry is fragile, the default rate under shirking can be higher: with probability λ it is the same as for the solid industry (i.e. $\underline{\Delta}$), but with probability $1 - \lambda$ it is $\bar{\Delta} > \underline{\Delta}$. The expected default rate under shirking in the fragile industry ($\lambda\underline{\Delta} + (1 - \lambda)\bar{\Delta}$) is denoted by $\hat{\Delta}$.[10]

To illustrate this, consider the resecuritization of loans. Our assumption means that when the technique is fragile (say in the case of resecuritization of RMBSs issued from subprime mortgages), there is a probability $1-\lambda$ that shirking (i.e. not exerting due diligence in the screening and monitoring of borrowers) will result in higher default rates than for a strong industry, even in the absence of a

[10] One could set λ to zero, so that the probability of default under shirking would always be $\bar{\Delta}$, without major alterations to our results. As will be seen below, $\lambda > 0$ implies that, even when all managers shirk the fragility of the industry may not be discovered immediately.

crisis. By contrast if the industry is strong shirking is less damaging. Thus, when the industry is fragile not only is it exposed to crises (with probability $1-p$ there is a crisis and all firms obtain 0 return), but also it is exposed to smaller shocks (with probability $1-\lambda$ the adverse consequences of shirking are greater than they would have been for the strong industry).

Managerial effort is not observable by investors, and, when the manager does not exert effort, she obtains an unobservable private benefit from shirking, denoted by B. Hence, since managers have limited liability, there is a moral hazard problem.[11] The investor is the principal and the manager the agent. The unfolding of uncertainty within one period is represented in Figure 1.

We now introduce our two key maintained assumptions. First, we assume

$$A1 : p\hat{\Delta} > \underline{\Delta},$$

which means that the expected negative consequences of shirking on output are greater when the industry is fragile than when it is solid.[12] In other words, strong industries are more robust to shirking than fragile ones. Second, we assume

$$A2 : \underline{\Delta}Y > B.$$

which means that effort is socially optimal when the industry is known to be strong. Under A1 this also implies that effort is socially optimal when the industry is fragile.

For simplicity we don't model the details of the activities of the innovative industry. Hence, although our modelling of that industry can be interpreted as a reduced form of the financial sector, we don't study explicitly financial intermediation or services to the traditional sector. This simplification enables us to concentrate on what is the main focus of our analysis: the agency relationship between investors and managers, the allocation of resources between the traditional sector and the innovative one and the dysfunctionalities arising because of moral hazard.

[11] As in Holmström and Tirole (1997) we cast the problem in terms of private benefits foregone by the manager when exerting effort. One could equivalently consider a model without private benefits but where effort would be costly.

[12] Note that A1 implies that $\bar{\Delta} > \underline{\Delta}$.

3.3 Learning

All the agents in the economy observe all returns realizations, and use these to conduct rational Bayesian learning about θ. At the first period ($t = 1$), agents start with the prior probability, π_1 that $\theta = 1$. For $t > 1$, denote by π_t the updated probability that the innovative industry is strong, given the returns realized in the innovative sector at times $\{1, ..., t-1\}$. By the law of large numbers, aggregate default rates may reveal a lot of information. For example, Figure 1 shows that an aggregate default rate of either $\bar{\Delta}$ (large number of individual losses) or 1 can only come from a fragile industry. Thus the fragility of an industry is detected for sure (and π_t goes to 0) as soon as the average default rate strictly exceeds $\underline{\Delta}$.[13] Until this date, π_t remains strictly between 0 and 1. Thus if all projects succeed at date t, it implies that managers have exerted effort and the probability that the industry is strong has to be revised upward:

$$\pi_{t+1} = \frac{\pi_t}{\pi_t + p(1-\pi_t)} > \pi_t. \tag{1}$$

By contrast if the default rate is $\underline{\Delta}$,this shows that managers have shirked. Yet, the probability that the industry is strong increases even further:

$$\pi_{t+1} = \frac{\pi_t}{\pi_t + \lambda p(1-\pi_t)} > \pi_t, \tag{2}$$

since, in spite of shirking, the aggregate default rate has been limited. This is because, when the industry is fragile, shirking implies a positive probability of default rates in excess of $\underline{\Delta}$.

Thus, as long as the average return of firms operating in the innovative sector is sufficiently good, beliefs about θ improve. But, when it is too low, confidence melts down and π_t suddenly falls down to 0.

Note that assumptions A1 and A2 imply that effort is always socially optimal, and would take place in the first best, i.e.,

$$(\pi_t + (1-\pi_t)p)Y > (\pi_t(1-\underline{\Delta}) + (1-\pi_t)[1-\hat{\Delta}]p)Y + B, \forall \pi_t.$$

[13]If a crisis could also occur when $\theta = 1$, while remaining less likely than for $\theta = 0$, the learning process would be smoother. But it would have the same major qualitative feature, i.e., π_t would increase after good returns and decrease after bad returns.

3.4 Contracts

At time t, newly born investors and managers interact for one period. We assume however that they have access to the complete performance history of the industry, which implies that they share the updated belief π_t.

First consider the case where the contract stipulates that managers are to exert effort. If there is a crisis all projects fail, while otherwise all projects should succeed, and obtain Y. If the output of the project run by manager i is equal to 0, while others succeeded, it must be that manager i shirked. It is optimal to set his compensation to 0 in that case. Also, it is weakly optimal to set managerial compensation to 0 when there is a crisis and all firms obtain 0 output.

Second, consider the case where the contract stipulates managers are not to exert effort. In that case, only the expected compensation of managers matter, not its allocation across states. Hence, it is weakly optimal to compensate the manager only when the output of his firm is Y.

Accordingly, we consider contracts where managers are compensated only when their firm generates output Y. We denote the value of this compensation at time t by m_t, which can be interpreted as the bonus, received by managers if and only if they succeed.

4 Equilibrium when effort is observable

Consider first the benchmark case where effort is observable. In that case, since effort is efficient, managers are simply instructed to exert it, and their compensation is set to clear the labour market. At time t, the expected output of firms operating in the innovative sector is:

$$S_t = [\pi_t + (1 - \pi_t)p]Y, \tag{3}$$

and the expected compensation of the managers is:

$$M_t = [\pi_t + (1 - \pi_t)p]m_t,$$

while investors obtain $S_t - M_t$ in expectation. Managers with opportunities in the traditional sector below M_t prefer to operate in the innovative sector. Hence, the supply of managers for that sector is:

$$G(M_t). \tag{4}$$

Investors whose return opportunities in the traditional sector are below $S_t - M_t$ prefer to invest in the innovative one, and thus need to hire a manager. Hence, the demand for managers in the innovative sector is:

$$F(S_t - M_t). \tag{5}$$

Equating (4) and (5) the labour market–clearing condition is:

$$G(M_t) = F(S_t - M_t). \tag{6}$$

The supply of managers is continuous and increases from $G(0) = 0$ to $G(\bar{\nu}) = 1$, while the demand for managers is continuous and decreases from $F(S_t) > 0$ to 0. Consequently there exists a unique solution M_t^* to (6). In our simple model a natural measure of the size of the innovative sector is the number of managers, or equivalently the number of firms operating in that sector. Based on the discussion above, we state our first proposition (illustrated in Figure 2, Panel A):

Proposition 1: *When effort is observable, the equilibrium expected compensation of managers in the innovative sector is M_t^* (the solution of (6)) and the size of that sector is: $G(M_t^*)$.*

Suppose that, up to $t-1$, there has been no crisis and the updated probability that the industry is strong is π_t. If during period t there is no crisis, then investors become more optimistic and the updated probability that the industry is strong goes up to $\pi_{t+1} > \pi_t$. Thus expected output in the innovative sector increases and the demand curve (5) goes up, while the supply curve (4) stays constant. Consequently, the equilibrium compensation of managers in period $t+1$ is $M_{t+1}^* > M_t^*$. In contrast, after a crisis, the updated probability that the industry is strong goes down to 0 and the expected output in that industry goes down to pY. Consequently the size of the innovative sector shrinks. These remarks are summarized in the following corollary:

Corollary 1: *When effort is observable, as long as there is no crisis the size of the innovative sector goes up, along with the expected compensation of managers employed in that sector. When there is a crisis, the size of the innovative sector and the equilibrium expected compensation of managers suddenly drops.*

Thus, as in Zeira (1999), Barbarino and Jovanovic (2007) and Pastor and Veronesi (2006), rational Bayesian learning leads to dynamics akin to those of speculative booms.

5 Equilibrium under moral hazard

5.1 Incentive compatibility

When effort is not observable but requested, we must impose the incentive compatibility condition that the manager prefers to exert effort rather than consuming private benefits, i.e.,:

$$[\pi_t + (1-\pi_t)p]m_t \geq [\pi_t\{1-\underline{\Delta}\} + (1-\pi_t)\{1-\hat{\Delta}\}p]m_t + B,$$

or

$$m_t \geq \frac{B}{p\hat{\Delta} - \pi_t[p\hat{\Delta} - \underline{\Delta}]}.$$

This condition implies a minimum expected pay-off for the manager. Denote this minimum expected pay-off by R_t. We refer to it as the rent of the manager:

$$R_t = \frac{\pi_t + (1-\pi_t)p}{p\hat{\Delta} - \pi_t[p\hat{\Delta} - \underline{\Delta}]}B. \tag{7}$$

The incentive compatibility condition can be rewritten as,

$$M_t \geq R_t, \tag{8}$$

which means that the compensation promised to the manager must be large enough to entice effort. Note that the rent R_t which must be left to the manager, varies with the beliefs about the strength of the innovative sector. Under $A1$, R_t increases with π_t. This is because, under $A1$, the strong industry is more robust to shirking than the fragile one. Hence, managers find shirking more tempting when they are confident that the industry is robust. Therefore, high rents are needed to provide incentives in that case.

Substracting the rent from the expected output yields the pledgeable income, i.e., the maximum expected revenue that can be pledged to the investors without compromising the incentives of the manager:

$$P_t = S_t - R_t. \tag{9}$$

Since both the expected output and the rent increase with π_t, the pledgeable income can be non–monotonic with the expected strength of the innovative sector. Combining the conditions stated above, we obtain our next proposition.

Proposition 2: *When effort is not observable but requested, the pledgeable income in the innovative sector is a concave function $P(.)$ of π_t . Moreover if*

$$p + (1-p)\frac{Y\underline{\Delta}}{B} < \frac{p\hat{\Delta}}{\underline{\Delta}}, \tag{10}$$

then pledgeable income decreases when π_t is close enough to 1.

Proposition 2 implies the following dynamics for the pledgeable income when effort is requested. As long as high returns are observed, confidence in the innovative sector increases. This boosts expected output, but it also raises rents. Our assumptions imply that the marginal impact of increased confidence on pledgeable income decreases with π_t ($P(\pi_t)$ is concave). Moreover, under condition (10) the increase in rent ultimately dominates the increase in expected surplus, and pledgeable income starts decreasing.

To shed more light on the economic forces at play in Proposition 2, it is useful to define:

$$b(\pi_t) = \frac{R_t}{S_t} = -\frac{B}{Y\left[\pi_t\underline{\Delta} + (1-\pi_t)p\hat{\Delta}\right]},$$

which is the minimum fraction of output that must be given to managers to incite effort. The properties of $b(\pi_t)$ are stated in the next corollary:

Corollary 2: *The minimum fraction of output which must be left to managers to incite effort, $b(\pi_t)$, is increasing in the updated probability that the industry is strong, π_t. But, under A2, this fraction remains lower than 1, i.e., $b(1) < 1$. Furthermore, condition (10) is equivalent to:*

$$b(1) = \frac{B}{\underline{\Delta}Y} > \left(\frac{1}{p} - 1\right) \Big/ \left(\frac{\hat{\Delta}}{\underline{\Delta}} - 1\right). \tag{11}$$

Note that the right-hand side of (11) increases with the probability $(1-p)$ of a crisis, and decreases with the ratio $\frac{\hat{\Delta}}{\underline{\Delta}}$, which measures the incidence of moral hazard in a fragile industry, relative to a strong one. Therefore we can state the following corollary.

Corollary 3: *Condition (10) (under which the pledgeable income $P(\pi_t)$ decreases for π_t close to 1) is more likely to hold when:*

- *moral hazard is severe ($b(1)$ large),*

- *the probability of a crisis is small (p large),*
- *the incidence of moral hazard is much larger when the industry is fragile than when it is strong ($\frac{\hat{\Delta}}{\underline{\Delta}}$ large).*

5.2 Equilibrium with effort

When effort is requested from the agent, there are two possible regimes, depending on whether the incentive compatibility condition binds or not. In the first regime, the market clearing condition determines the equilibrium compensation of the managers, as in the previous section, i.e.,

$$M_t = M_t^* \text{ s.t. } G(M_t^*) = F(S_t - M_t^*)$$

as in (6). For this to be the equilibrium, it must be that the incentive compatibility condition holds for M_t^*, i.e.,

$$M_t^* \geq R_t. \tag{12}$$

As illustrated in Figure 2, Panel A, in this regime the supply and demand curves on the labour market intersect above R_t, so that the incentive compatibility condition does not bind.

In the second regime, as illustrated in Figure 2, Panel B, the supply and demand curves on the labour market intersect below R_t. Thus, the incentive compatibility condition binds, i.e.,

$$M_t^* < R_t. \tag{13}$$

and the expected managerial compensation is

$$M_t = R_t. \tag{14}$$

Since this is above M_t^*, managers employed in the innovative sector earn greater expected compensation than in the observable effort case. Thus, although they are competitive, they earn rents. Such rents make working in the innovative sector very attractive. Indeed the number of managers who want to work in that sector is above the demand for their services, i.e.,

$$G(M_t) = G(R_t) > F(S_t - M_t).$$

Thus there is rationing in the labour market, as in Shapiro and Stiglitz (1984).

For simplicity consider the case where ν is uniformly distributed over $[0,\bar{\nu}]$ and ρ is uniform over $[0,\bar{\rho}]$, i.e.,

$$G(\nu) = \frac{\nu}{\bar{\nu}}, F(\rho) = \frac{\rho}{\bar{\rho}}.$$

In that case the market clearing condition (6) defining M_t^* becomes:

$$\frac{M_t^*}{\bar{\nu}} = \frac{S_t - M_t^*}{\bar{\rho}}.$$

Thus:

$$M_t^* = \beta S, \tag{15}$$

where:

$$\beta = \frac{\bar{\nu}}{\bar{\nu}+\bar{\rho}} \in [0,1].$$

The compensation of the manager is equal to a fraction (β) of the value created by the firm. This fraction reflects the relative values of the outside opportunities of managers and investors in the traditional sector. When the outside opportunities of managers, measured by their skills in the traditional sector, are high and the opportunities of the investors in that sector are poor, the market clearing compensation of the managers in the innovative sector is high.

The condition under which the incentive compatibility condition does not bind, (12), is equivalent to the condition that the pledgeable income be greater than the expected income of the investors in the first best

$$S_t - M_t^* \leq P_t. \tag{16}$$

After some manipulations, we obtain our next proposition, which is illustrated in Figure 3.

Proposition 3: *Consider the case where effort is not observable, but is requested. If F and G are uniform, there exists a threshold value $\bar{\pi}$ such that, for $\pi_t \leq \bar{\pi}$ the incentive compatibility condition is not binding and the compensation of managers is set by the market clearing condition (6).*

- *When*

$$b(1) = \frac{B}{\underline{\Delta}Y} < \beta, \tag{17}$$

then $\bar{\pi} \geq 1$ and the incentive compatibility constraint never binds, so that the equilibrium outcome is the same as in the first best.

- *When*

$$\beta < b(1) = \frac{B}{\underline{\Delta} Y}, \tag{18}$$

the threshold value $\bar{\pi}$ less than 1. For $\pi_t > \bar{\pi}$, the compensation of managers is set by the incentive compatibility condition, (14).

Inequality (17) holds when $b(1)$, the share of expected output that has to be left to managers for incentives reason, is small. In that case, the moral hazard problem is not severe and induces no distortion in equilibrium. Hence, the expected net cashflow obtained by investors in innovative sector firms is: $S_t - M_t^*$.

Inequality (18) holds if $b(1)$ is large. In that case, the agency problem is more severe and, when π_t is high enough, rents become so large that the incentive compatibility condition binds. Correspondingly, the expected net cashflow obtained by investors in the innovative sector is: $S_t - R_t = P_t$.

5.3 Equilibrium without effort

Now turn to the case where effort is not requested from the agent in equilibrium. In that case the expected output from the project is:

$$\hat{S}_t = [\pi_t(1 - \underline{\Delta}) + (1 - \pi_t)p(1 - \hat{\Delta})]Y,$$

and the expected wage earned by managers is

$$\hat{M}_t = [\pi_t(1 - \underline{\Delta}) + (1 - \pi_t)p(1 - \hat{\Delta})]m_t.$$

The demand for managers is: $F(\hat{S}_t - \hat{M}_t)$ while he supply of managers is: $G(\hat{M}_t + B)$. The market clearing expected wage without effort is $\hat{M}_t^*$ is such that

$$F(\hat{S}_t - \hat{M}_t^*) = G(\hat{M}_t^* + B).$$

Using this market clearing condition, the next proposition states the equilibrium wage arising in the uniform case when effort is not requested.

Proposition 4: *Assume F and G are uniform. If effort is not requested in equilibrium, then the labour market for managers clears, the expected compensation of managers is*

$$\hat{M}_t^* = \beta \hat{S}_t - (1 - \beta)B,$$

and their total expected utility is $\beta(\hat{S}_t + B)$ *while that of investors is* $(1-\beta)(\hat{S}_t + B)$.

When effort is not exerted, the total expected value created by each firm is $\hat{S}_t + B$. Since the agent does not exert effort no rent is needed and the market clears. Thus the share of the total value created obtained by managers and investors simply reflect their outside options in the traditional sector. Correspondingly, managers get fraction β of $\hat{S}_t + B$ while investors get the complementary fraction.

5.4 Is there effort in equilibrium?

We now investigate if effort prevails in equilibrium. Consider a candidate equilibrium, where effort is requested. Could a pair manager–investor be better off by deviating to a contract which would not request effort? If there is no such profitable deviation, then effort is requested in equilibrium. Symmetrically, consider a candidate equilibrium where effort is not requested. Could a pair manager–investor be better off by deviating to a contract which would request effort? Again, if there is no such profitable deviation, there exists an equilibrium without effort. The following proposition, illustrated in Figure 3, states the conditions on parameter values under which one of the candidate equilibria or the other prevails in the uniform case.

Proposition 5: *Assume* F *and* G *are uniform and* (10) *and* (18) *hold. If*

$$b(1) > 1 - \frac{1}{\underline{\Delta} + \frac{1}{1-\beta}}. \tag{19}$$

there exists a threshold value $\hat{\pi} > \bar{\pi}$ *such that effort is requested in equilibrium for* $\pi_t \leq \hat{\pi}$, *while equilibrium involves no effort for larger values of* π_t.

The intuition behind the proposition is the following: As long as $\pi_t < \hat{\pi}$, the rents which must be left to the managers are sufficiently small that the pledgeable income is greater than the expected income investors would get if effort was not requested. So, for these values of π_t, investors prefer to request effort, and it is implemented in equilibrium. (19) is the condition under which the threshold value $\hat{\pi}$ is lower than 1. This condition states that for $\pi_t = 1$, the pledgeable income is lower than the expected income of the investors when effort is not requested. This situation is more likely to happen

when moral hazard is severe ($b(1)$ is large), shirking has a limited impact on a strong industry ($\underline{\Delta}$ is small) and competition for managers is not too fierce (β small).

5.5 Cycles

The cycle in the innovative sector Denote by $H(\pi_t)$ the fraction of the manager's population employed in the innovative sector. It is a measure of the size of the innovative sector. The above analysis, and in particular Propositions 3, 4 and 5, imply the following proposition.

Proposition 6: *Assume F and G are uniform and* (10) *and* (18) *hold. As long as there is no crisis, the equilibrium size of the innovative sector, measured by the fraction of the managers employed in that sector, $H(\pi_t)$, is characterized as follows:*

- *For $\pi_t < \bar{\pi}$, $H(\pi_t) = F((1-\beta)S_t)$,*
- *for $\bar{\pi} \leq \pi_t < \hat{\pi}$, $H(\pi_t) = F(P_t)$,*
- *and for $\hat{\pi} \leq \pi_t$, $H(\pi_t) = F((1-\beta)(\hat{S}_t + B))$.*

Proposition 6 implies the following pattern for the lifecycle of the financial sector, illustrated in Figure 4.

- First, as long as there is no crisis, π_t grows (from π_1 to $\bar{\pi}$). This increase in the confidence placed in the innovative sector attracts managers and capital. Correspondingly the size of the sector increases. This is the initial boom period following the introduction of the innovation.
- Second, if there is still no crisis, π_t continues to grow (from $\bar{\pi}$ to $\hat{\pi}$) and managers become so confident that investors must leave them rents to provide them with incentives. As confidence builds up, these rents absorb an increasing fraction of the returns realized in the innovative sector. This deters the allocation of capital to the innovative sector and slows down its growth, which can even become negative. This is the choking period.
- Third, if the crisis still does not occur, confidence continues to build up, as π_t goes from $\hat{\pi}$ towards 1. But the rents needed to incentivize effort have now grown so high that investors

prefer not to request effort in equilibrium. Consequently, managers screen projects less carefully, and failures take place. This is the risk–taking period. As long as the rate of these failures remains as low as $\underline{\Delta}$, confidence still builds up. But as soon as the rate of failure exceeds $\underline{\Delta}$, investors realize the sector is fragile, and there is a crisis.

- At any point in time, if the industry is fragile, this can be revealed, either by a large crisis implying the failure of the entire innovative sector, or (in the risk–taking period) by a large average failure rate $\bar{\Delta}$ in the innovative industry. In both cases, confidence is destroyed and π_t drops to 0, where it remains. At that point, the size of the innovative sector and the wages in that sector undergo a sharp drop.

The cycle in GDP In our analysis, the cycle in the innovative sector drives that of aggegate output. Denote by $\hat{\nu}$ the type of the marginal manager and by h the size of the innovative sector. We have that

$$F(S_t - M_t) = G(\hat{\nu}) = h.$$

As long as there is no crisis, and the agents are incentivized to exert effort, GDP in our model is equal to

$$F(S_t - M_t)Y + \int_{S_t - M_t}^{\bar{\rho}} z dF(z) + \int_{\hat{\nu}}^{\bar{\nu}} s dG(s),$$

where the first term is the output of the innovative sector, while the second and last terms are the values created in the traditional sector by investors and managers, respectively. After a change of variables, this can be rewritten as

$$hY + \int_h^1 F^{-1}(x)dx + \int_h^1 G^{-1}(x)dx. \tag{20}$$

This implies that the derivative of GDP with respect to h is positive, and we obtain the following result.

Proposition 7: *Assume* (10) *holds. As long as there is no crisis and the agents are incentivized to exert effort, GDP increases with the size of the innovative sector.*

The proposition implies that, as confidence in the innovative sector builds up and the size of that sector increases, GDP increases too, as well as the share of the innovative sector in the economy. This

is consistent with the growth in GDP and in the financial sector's share of GDP observed during the innovation wave at the beginning of this century (for evidence on this growth pattern, see, e.g., Philippon, 2008.) The proposition also implies that, after this initial period, as we enter in the choking region, GDP growth slows down. And, if there is a crisis in the innovative sector, GDP suddenly drops.

6 Empirical implications

In this section, we draw the empirical implications of our theoretical analysis for the financial industry. We discuss the relationship between stylized facts and the equilibrium outcomes in our model, in particular those described in Proposition 6 and illustrated in Figure 4. We also offer new empirical implications that have not been tested previously.

Note that our model is not relevant for all parts of the financial industry and all periods. It applies to segments and periods with significant innovations, uncertainty about the strength of the innovation and informational asymmetries between investors and finance sector managers. Thus, our model does not apply to standard banking activities, such as loans to corporations, conducted in traditional ways, during periods with little financial innovation, e.g., the 1950s or the 1960s. In contrast, it is relevant to analyze the flow of financial innovations which took place at the beginning of this century, such as e.g., structured finance, hedge funds and private equity funds. It is also relevant for previous waves of financial innovations, with uncertainty and information asymmetry, such as, e.g., that of the 1920s, the growth of synthetic portfolio insurance in the early 1980s, or that of the junk-bonds markets from late 1970s until they crashed in the 1980s, following the Michael Milken scandals.

Note also that neither in our model nor in practice do all financial innovations necessarily result in crashes. For example the securitization of standard (not subprime) bank loans is commonly viewed as a successful financial innovation. It started in 1968, when the GNMA created the passthrough security for mortgages guaranteed by the US government. It was later extended to broader classes of loans, including those guaranteed by Government Sponsored Entities (Fannie Mae and Freddie Mac) as well as more risky loans such as auto, credit card, commercial mortgage, student and business loans. So far, none of these loan categories has experienced any crash.

6.1 Size and risk

The subprime mortgages industry offers a good example of an innovative industry that eventually proved fragile and crashed. It relied on new techniques i) to attract new categories of borrowers previously denied access to credit (e.g., "low document loans" or "option mortgages") and to ii) transfer risks (e.g., tranching). There was uncertainty about the strength of that industry, related in particular to uncertainty about the probability of a nationwide drop in real–estate prices. And there was some information asymmetry regarding the exact structure and content of the portfolios resulting from tranching and the creation of "special investment vehicles."

The dynamics of that industry are consistent with our model, in particular Proposition 6. Until 2006 it performed well and expanded. Then, in 2007, the risk realized, there was a surge in defaults and a strong decline in activity. To illustrate this pattern, Figure 5 (which we borrow from Jaffee, 2008) plots the annual volume of subprime originations from 1994 to 2007.

Furthermore, Proposition 6 states that, after the initial boom period, and before the possible eventual crisis, risk–taking can prevail in equilibrium. This arises when confidence in the strength of the industry has become high and the rents that must be conceded to the agents to spur efforts have grown large. At that point investors choose to give up on incentives. Consequently, agents fail to exert effort, and failure rates rise. Interpreting effort as the care taken by agents to avoid investment in poorly performing assets, this pattern is consistent with that observed in the subprime industry. While the crisis materialized only in the summer and fall of 2007, the lending standards deteriorated earlier, as illustrated by Figure 6 (which we borrow from Jaffee, 2008). This figure compares the deliquency rates of subprime loans by age and year of origination. There were no marked differences across years from 2000 to 2005. But loans made in 2006 and 2007 performed more poorly.

More in–depth empirical analyses show that the decline in performance for loans made in 2006-2007 cannot be fully explained by observable composition effects (i.e. changes in the types of subprime loans and in observable borrower characteristics) nor by economic conditions. Demyanyk and Van Hemert (2008) regressed deliquency rates on borrower/loan characteristics and economic variables such as house price indices for different vintages of subprime loans. The explanatory power of the regressors is lower in 2006 and 2007. In the context of our model, this increase in the role of the residuals in 2006 and 2007 can be interpreted as reflecting the shift to a regime with unobservable agents' shirking, with

higher failure rates.

The investors who purchased CDO tranches from the pools of subprime loans were rather sophisticated institutional investors. One might think that such professional investors should have anticipated or found out any decline in effort and the resulting increase in failure rates. This, again, is consistent with our model, where systemic risk–taking arises in equilibrium and is fully anticipated by investors.

6.2 Rents

Our model implies that during innovation waves, after the industry has acquired a good performance record, managers earn rents, i.e., their pay exceeds the market clearing wage they would receive in a frictionless market. These rents arise in equilibrium, in spite of competition between managers, because of incentive constraints, in line with Shapiro and Stiglitz (1984).

This theoretical result is consistent with the empirical findings of Philippon and Reshef (2008). They observe that during the 1920s and after the 1980s, there was a burst of financial innovation accompanied by an increase in the pay of managers in the finance sector. Philippon and Reshef (2008) estimate that, in the recent period, rents accounted for 30% to 50% of the wage differential between the financial sector and the rest of the economy. They also find that this induced an increase in involuntary unemployment in the financial industry. Our theoretical analyses delivers results which are exactly in line with these facts.

Another implication of our model is that, as confidence in the financial innovation builds up, the share of the value created in the innovative financial sector should be increasingly allocated to the agents at the expense of investors. This new implication has not yet been explicitly tested. But it is consistent with anecdotal evidence. For example, in the case of private equity funds at the beginning of this century, on top of the 2 percent annual management fee and the carried interest incentive fee, extra fees were progressively added, sometimes labeled portfolio fees (see Phalippou, 2008).

6.3 Net returns and size

In our equilibrium, as long as there is no crisis, realized returns *net of management fees*, start at a high level and then go down with past performance. While realized net returns are non–increasing, *expected* net returns initially increase. This is because the increase in the probability of success outweighs the

increase in fees. The corresponding increase in expected net returns attracts capital. Hence the size of the sector (measured, e.g., by assets under management or total open positions) initially increases with past performance. Thus, during that period, an increase in size coexists with a decrease in realized net returns, which can be interpreted in terms of decreasing marginal returns to size.[14] But, at some point the increase in agency rents can outweigh the increase in confidence and expected net returns themselves start decreasing. When that point is reached, investment in the innovative sector declines. Thus, overall, the relationship between the amount of money invested in the innovative sector and its past performance will be inverse U shaped. This implication from our theoretical analysis is in line with the empirical results of Ramadorai (2008). He studies the closed–fund premium for hedge funds, which were among the major institutional arrangements in the financial innovation wave at the beginning of this century. The closed–end fund premium offers a proxy for the demand for fund management services. When it is large, it suggests that many investors are eager to delegate the management of their money to that innovative segment of the financial industry. Consistent with our model, Ramadorai (2008) finds that this premium is initially increasing and then decreasing with past performance.

6.4 Managers' skills

In our model, as the size of the innovative sector increases, it attracts better and better managers. This implication of our analysis is in line with the empirical findings of Philippon and Reshef (2008) that the average skill level of managers in the financial services industry rose during the financial innovation wave at the beginning of this century.

6.5 Time– and money– weighted average returns

During the ascending phase of the cycle the size of the innovative sector goes up. Hence time–weighted average returns give more weight to early returns than do money–weighted averages. Now,

[14] This is reminiscent of the results of Berk and Green (2005), but obtains for different reasons. In the present model, decreasing returns to scale are not directly driven by the technology used by managers. They stem from the combination of learning, moral hazard and equilibrium effects. Also, managers in our model are identical, while they are heterogeneous in Berk and Green (2005).

as mentioned above, even if there is no crisis, realized net returns tend to decrease. Hence the time–weighted average net return is greater than the money–weighted one. And the former overestimates further returns more than the latter. Thus our model warns that investors should view past time–weighted average net returns with caution, and should not extrapolate them, since this would neglect the growth in agency rents.

6.6 Estimating risk

Our learning model is a good description of those industries, such as the financial one at the turn of the twenty first century, where until some rare event occurs there is no strong change in beliefs, and then when the rare event happens, there is a strong decline in optimism. In such an environment, frequentist estimations of risk would be very misleading. Consider the case where, up to time t, there has been no crisis. The estimated default rate computed using past data is 0. This is lower than the rational Bayesian estimate of the risk of default, $1 - (\pi_t + (1 - \pi_t)p)$, which is strictly positive as $\pi_t < 1$. This is in line with the analysis of the recent crisis offered by Brunnermeier (2009) and Rajan et al (2008). Brunnermeier (2009) notes that: "the statistical models of many professional investors and credit-rating agencies provided overly optimistic forecasts about structured finance products. One reason is that these models were based on historically low mortgage default and delinquency rates."

In our simple framework, it takes only one negative shock for the confidence in the innovative sector to be entirely lost. It would be straightforward to extend our model to allow for a more gradual process. Suppose that the probability of a crisis is p when the industry is fragile and ϵ when it is strong, with $p > \epsilon > 0$. After the observation of a first shock, π_t would go down but would not immediately reach 0. Then, after each crisis episode, market participants would become more and more pessimistic and π_t would progressively go to 0.

7 Policy implications

As stated in Propositions 6 and 7, cycles arise in our model: Innovation waves spur periods of rapid growth, which can be followed by increases in risk taking and then crises. In the case of the financial sector, the risk–taking equilibrium regime, characterized in Proposition 6, can be interpreted as

financial system instability.

To avoid such instability, should regulators step in and curb the increase in valuations during periods of rapid growth? In the context of our model, the answer to this question is negative. Policies curbing the growth of the innovative sector, such as taxes or quantity restrictions, would not enhance stability and would decrease welfare. This is because, in our analysis, the growth of the financial sector is not an irrational bubble, based on misguided beliefs and also there are no externalities between agents.

This negative answer does not imply that policy actions are not called for. At the origin of the imperfections in our equilibrium lies the unobservability of the actions of the agents. Final investors cannot observe if managers are making the necessary efforts to screen out bad investments and avoid high default risk, or if they shirk and consume private benefits. An important policy measure in this context would be to increase transparency and disclosure. This could be achieved, for example, by switching from OTC markets to exchanges with clearing houses. In the former, trades and positions are unobservable, while in the latter they are reported and disseminated. Policy intervention would be useful to foster the use of such market infrastructures to the extent that they involve fixed costs like utilities, and also because switching from one market platform to another one is difficult due to coordination problems and liquidity externalities. Enhancing transparency and disclosure increases the extent to which agents' actions are observable and reduces the ability to shirk, i.e., in our model, reduces B. This, in turn, lowers the rents earned by the agents in the innovative sector and thus increases pledgeable income, P_t. Inspecting Figure 3, one can see that an upward shift in P_t leads to a reduction in the zone where risk–taking arises in equilibrium. Hence raising transparency and disclosure improves stability.

Figure 3 also shows that this increase in P_t is useful only at later stages in the cycle. Initially, as long as the incentive compatibility condition is not binding, to the extent that policies aiming at raising transparency and disclosure would be costly, they would be suboptimal. But, at later stages in the cycle, when the innovative sector has grown large and bankers become very confident, policies reducing information asymmetries are called for. In that sense the optimal policy is countercyclical, since regulatory intervention is needed when the growth cycle has built momentum.

While such policies would raise financial stability and net returns for final investors, they are not

unanimously desired by all the agents in our model. Once they have been recruited in the innovative sector, managers prefer information asymmetries to be large, in order to earn large rents. This suggests that self–regulation is not likely to give rise to optimal policies. The intervention of regulators is called for.

Another important variable in our analysis is the structure of the compensation of the managers in the innovative sector. Following the recent financial crisis, and the severe losses incurred by banks, it has been discussed whether large bonuses should be granted to managers in the financial sector. One of the arguments offered by advocates of such compensation is that it is necessary to retain skilled bankers. To discuss that point, consider a variant of our model where the compensation of managers would be constrained to be greater than a given constant, even after failure. In the region where the incentive compatibility condition is imposed and binds, this would imply that the compensation of managers be raised also after success, in order to maintain the wedge between the reward after success and that following failure. This would lead to an increase in the total expected compensation of the managers. In turn, this would reduce pledgeable income, P_t, and consequently to an increase in the risk–taking zone. Thus, our analysis warns that granting bonuses to managers after failure undermines stability.[15]

8 Conclusion

Our model analyzes the lifecycle of innovative industries, characterized by uncertainty about the fundamental value of the innovation and information asymmetry about the actions of the managers. These assumptions fit particularly well the evolution of the financial sector at the beginning of this century. And our theoretical analysis delivers a rich set of empirical implications, in line with the evolution of that industry and its eventual crisis. In particular our model shows how, in equilibrium, during the initial period of success, the industry grows, attracting human and financial capital, while the rents of the managers increase. This increase in rents undermines the net returns obtained by investors and can eventually make incentives so expensive that excessive risk–taking prevails in the last part of the cycle.

[15] Note however that Edmans and Gabaix (2009) offer interesting arguments showing that, in other setups than ours, such pay can be optimal.

Our model suggests that the appropriate policy response in this context is not to curb the increase in valuations. Policy should instead aim at increasing transparency and disclosure. By reducing information asymmetries and rents, such policies would reduce the occurence of the risk–taking regime, and thus improve stability. Along with this advice to policy makers, our analysis also offers a warning: Even with the maximum possible level of disclosure and transparency, rents, risk–taking and crises are bound to remain prevalent issues in innovative industries.

But, while fragile industries are eventually hit by crises, strong ones undergo sustained development. While this is not modelled in the present paper, it is reasonable to expect that, after some time, as the industry becomes better understood by the population of investors, agency problems become less and less of an issue. Thus, in the long term, strong industries cease to be speculative. A promising avenue of further research would be to explore, along these lines, the long run consequences of successive waves of innovation, and the corresponding cycles and growth dynamics.

References:

Barbarino, A., and B. Jovanovic, 2007, "Shakeouts and market crashes," *International Economic Review*, 385–420.

Berk, J., and R. Green, 2005, "Mutual fund flows and performance in rational markets," *Journal of Political Economy.*

Bergemann, D., and U. Hege, 1998, "Dynamic Venture Capital Financing, Learning and Moral Hazard," *Journal of Banking and Finance*, 703-735.

Bergemann, D., and U. Hege, 2005, "The financing of innovation: learning and stopping," *Rand Journal of Economics*, 719-752.

Brunnermeier, M., 2009, "Deciphering the Liquidity and Credit Crunch 2007–2008," *Journal of Economic Perspectives*, 77–100

Brunnermeier, M., and Y. Sannikov, 2009, "A Macro-Economic Model with a Financial Sector," Working Paper, Princeton University.

Caballero, Ricardo and Arvind Krishnamurthy. 2008. "Collective Risk Management in a Flight to Quality Episode." *Journal of Finance.* 63(5): 2195-2236.

DeMarzo, P., and Y. Sannikov, 2008, "Learning in Dynamic Incentive Contracts" Working paper, Stanford University.

Demyanyk, Y., and O. Van Hemert, 2008, "Understanding the Subprime Mortgage Crisis", Supervisory Policy Analysis Working Paper 2007-05. Federal Reserve Bank of St. Louis. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1020396.

Edmans, A. and X. Gabaix, 2009, "Is CEO Pay Really Inefficient? A Survey of New Optimal Contracting Theories", forthcoming *European Financial Management Journal*.

He, Z. and A. Krishnamurthy, 2008, "A Model of Capital and Crises", Working Paper, Northwestern University.

Holmstrom, B., and J. Tirole, 1997, "Financial intermediation, loanable funds and the real sector," *Quarterly Journal of Economics*, 663–692.

Jaffee, D., 2008, "The US Subprime Mortgage Crisis: Issues Raised and Lessons Learned." Commission on Growth and Development Working Paper n^0 28, Washington DC.

Lorenzoni, G., 2008, "Inefficient credit booms," *Review of Economic Studies*, 809–833.

Pastor, L., and P. Veronesi, 2006, "Was there a Nasdaq bubble in the late 1990's ?," *Journal of Financial Economics.*, 81, 161–100.

Phalippou, L., 2008, "Beware of venturing into private equity," *Journal of Economic Perspectives*, Volume 22, Number 4.

Philippon, T., 2008, "The evolution of the US financial industry from 1860 to 2007: Theory and evidence," Working paper, New York University.

Philippon, T., and A. Reshef, 2008, "Skill biased financial development: education, wages and occupations in the U.S. financial sector", Working paper, New York University.

Rajan, R., 2006, "Has Financial Development Made the World Riskier?", *European Financial Management* 12: 499.

Rajan, U., A. Seru, and V. Vig, 2008, "The Failure of Models that Predict Failure: Distance, Incentives and Defaults," Working paper, University of Michigan at Ann Arbor.

Ramadorai, T., 2008, "The secondary market for hedge funds and the closed hedge fund puzzle.", Working paper, Saïd Business School, Oxford University.

Rob, R., 1991, "Learning and capacity expansion under demand uncertainty." *Review of Economic*

Studies, 655–677.

Shapiro, C. and J. Stiglitz, 1984, "Equilibrium Unemployment as a Worker Discipline Device," *The American Economic Review*, 74, 433 – 444.

Zeira, J., 1987, "Investment as a process of search." *Journal of Political Economy*, 204–210.

Zeira, J., 1999, "Informational overshooting, booms and crashes." *Journal of Monetary Economics*, 237–257.

Appendix: Proofs

Proof of Proposition 2:

Substituting the values of S_t from (3) and R_t from (7), the pledgeable income at time t can be written as a function of π_t:

$$P_t = P(\pi_t) = [p + (1-p)\pi_t][Y + \frac{B}{[p\hat{\Delta} - \underline{\Delta}](\pi_t - \pi^*)}],$$

where $\pi^* = (p\hat{\Delta})(p\hat{\Delta} - \underline{\Delta})^{-1}$.Note that, under $A1$: $\pi^* > 1$. The first derivative of the pledgeable income is:

$$P^{'}(\pi_t) = (1-p)Y - B\frac{p + (1-p)\pi^*}{[p\hat{\Delta} - \underline{\Delta}](\pi_t - \pi^*)^2},$$

while the second derivative is:

$$P^{''}(\pi_t) = 2B\frac{p + (1-p)\pi^*}{[p\hat{\Delta} - \underline{\Delta}](\pi_t - \pi^*)^3}.$$

Thus, $P^{''}(\pi_t) < 0$ for the relevant values of π_t, i.e., $\pi_t \leq 1 < \pi^*$. Thus the pledgeable income as a function of π_t is a hyperbola, concave for $\pi_t \in [0, 1]$. Now,

$$P^{'}(1) < 0 \Leftrightarrow (1-p)Y < B\frac{p + (1-p)\pi^*}{[p\hat{\Delta} - \underline{\Delta}](1 - \pi^*)^2}.$$

That is:

$$p + (1-p)\frac{Y\underline{\Delta}}{B} < p\frac{(\hat{\Delta} - \underline{\Delta})}{\underline{\Delta}}.$$

Hence the proposition.

QED

Proof of Proposition 3:

Substituting the market clearing managerial compensation (15) into (16), we obtain $(1-\beta)S_t \leq P_t$,that is:

$$(1-\beta)[p+(1-p)\pi_t]Y \leq [p+(1-p)\pi_t]Y - \frac{[p+(1-p)\pi_t]B}{p\hat{\Delta} - \pi_t[\hat{\Delta}p - \underline{\Delta}]}.$$

Simplifying both sides by $p+(1-p)\pi_t$ and rearranging terms, we obtain a much simpler condition:

$$\beta Y \geq \frac{B}{p\hat{\Delta} - \pi_t[\hat{\Delta}p - \underline{\Delta}]}. \tag{21}$$

which is satisfied for π_t small enough. Note also that, if $B \leq \beta\underline{\Delta}Y$, condition (21) holds at $\pi_t = 1$, and therefore for all π_t. In contrast, if $\beta\underline{\Delta}Y < B < \beta\hat{\Delta}pY$, condition (21) holds at $\pi_t = 0$, but not at $\pi_t = 1$.

QED

Proof of Proposition 4:

In the uniform distribution case, the market clearing condition without effort amounts to

$$\frac{M_t^* + B}{\bar{\nu}} = \frac{\hat{S}_t - M_t^*}{\bar{\rho}}.$$

That is $\hat{M}_t^* + B = \frac{\bar{\nu}}{\bar{\rho}}(\hat{S}_t - \hat{M}_t^*)$.Or

$$\hat{M}_t^*(1+\frac{\bar{\nu}}{\bar{\rho}}) + B = \frac{\bar{\nu}}{\bar{\rho}}\hat{S}_t.$$

That is

$$\hat{M}_t^* = \frac{\bar{\nu}}{\bar{\rho}+\bar{\nu}}\hat{S}_t - \frac{\bar{\rho}}{\bar{\rho}+\bar{\nu}}B = \beta\hat{S}_t - (1-\beta)B.$$

Hence the expected total gain of the manager is $\hat{M}_t^* + B = \beta\hat{S}_t - (1-\beta)B + B = \beta(\hat{S}_t + B)$. Similarly, the expected profit of the investors is $\hat{S}_t - M_t^* = (1-\beta)(\hat{S}_t + B)$.

QED

Proof of Proposition 5:

The preliminary step of the proof is to compare the following three functions of π_t: $(1-\beta)S_t$, $(1-\beta)(\hat{S}_t + B)$ and P_t. Note the following:

- $(1-\beta)S_t$ and $(1-\beta)(\hat{S}_t + B)$ are linear and increasing in π_t .
- (18) implies that, at $\pi_t = 0$, $P(\pi_t = 0) > (1-\beta)S(\pi_t = 0)$.
- A1 and A2 imply that $(1-\beta)S_t > (1-\beta)(\hat{S}_t + B)$.
- (10) implies that P_t is concave in π_t, increasing at 0 and decreasing at 1.
- By (10) and (18), there exists a threshold value $\bar{\pi} \in [0,1]$ such that: $P(\pi_t) < (1-\beta)S_t$ if and only if $\pi_t > \bar{\pi}$.
- $P(\pi_t) < (1-\beta)(\hat{S}_t + B)$ holds for $\pi_t = 1$ if and only if $Y - \frac{B}{\underline{\Delta}} < (1-\beta)[(1-\underline{\Delta})Y + B]$, which simplifies to condition (19).

This implies that under condition (19) there exists a threshold value $\hat{\pi} \in [0,1]$ such that: $P(\pi_t) < (1-\beta)(S_t + B)$ if and only if $\pi_t > \hat{\pi}$. Also, since $(1-\beta)S_t > (1-\beta)(\hat{S}_t + B)$, we have that $\hat{\pi} > \bar{\pi}$. The functions $(1-\beta)S_t$, $(1-\beta)(\hat{S}_t + B)$ and P_t are plotted in Figure 3. The remainder of the proof consists in three steps, each one considering a candidate equilibrium and spanning the different possible values of π_t.

First, we consider the case where $\pi_t < \bar{\pi}$. Consider the case where effort is requested. Since $\pi_t < \bar{\pi}$, the incentive compatibility condition does not bind, we are at the first best and there is no scope for deviations to no effort. Consequently, for $\pi_t < \bar{\pi}$, there is an equilibrium with effort.

Second, turn to the case where $\hat{\pi} > \pi_t > \bar{\pi}$. As above consider the candidate equilibrium where effort is requested. The investor receives P_t and the manager R_t. The sum of the two is S_t. Could a manager and an investor both prefer to deviate to a contract with effort? In that deviation, the total value created by the firm would be $\hat{S}_t + B$. Under A1 and A2, this is lower than the total value created in the candidate equilibrium, S_t. Hence, the investor could not both agree to such the deviation. Consequently, for $\hat{\pi} > \pi_t > \bar{\pi}$, there is an equilibrium with effort.

Third, focus on the case where $\pi_t > \hat{\pi} > \bar{\pi}$. Consider a candidate equilibrium with effort. In that candidate equilibrium, in the innovative sector, managers would receive R_t while investors would

obtain P_t. Would an investor be better off deviating to no effort? This deviating investor could hire a manager from the traditional sector. The cheapest of these managers would be the marginal one, with type $\nu = G^{-1}(F(P_t)) = \frac{\beta}{1-\beta} P_t$. Hiring this manager to implement the project without effort, the investor would obtain $\hat{S}_t + B - \frac{\beta}{1-\beta} P_t$. This is profitable for the investor if that leads to greater exepcted profits for her, i.e., if

$$\hat{S}_t + B - \frac{\beta}{1-\beta} P_t \geq P_t.$$

That is

$$(1-\beta)(\hat{S}_t + B) \geq P_t,$$

which holds by construction for $\pi_t \geq \hat{\pi}$. Hence, equilibrium cannot prevail in equilibrium. Now, consider a candidate equilibrium without effort. The investor receives $(1-\beta)\hat{S}_t$ and the manager $\beta\hat{S}_t$. Could a manager and an investor both prefer to deviate to a contract with effort? In that deviation, the investor could at most get P_t. By construction, $P(\pi_t) < (1-\beta)\hat{S}_t$, hence, the investor could not agree to such a deviation. Consequently, for $\pi_t > \hat{\pi}$, there is an equilibrium without effort.

QED

Figure 1: The structure of uncertainty in period t

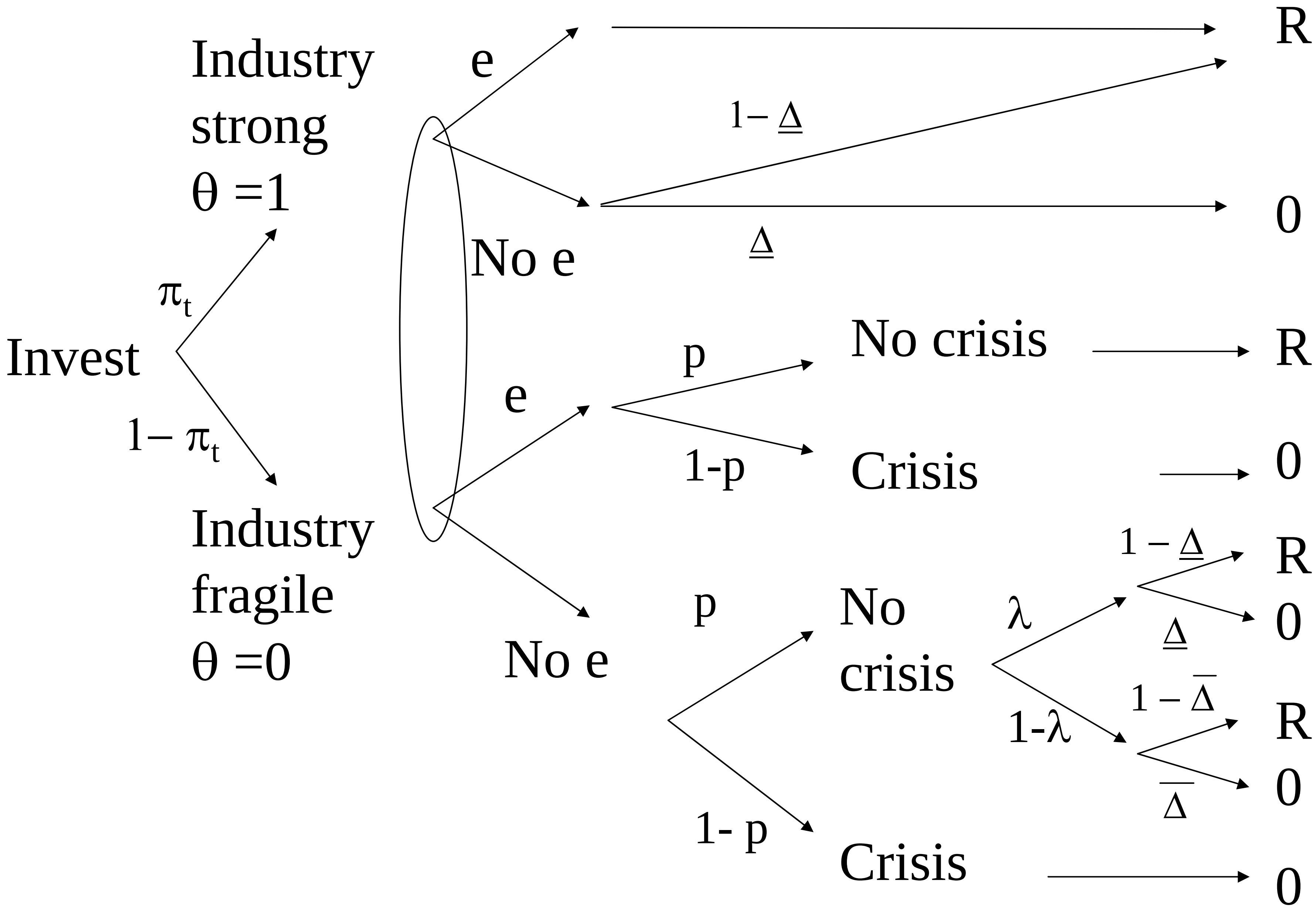


Figure 2: Supply, demand & rents

M_t is the expected compensation of the manager. $G(M_t)$ is the mass of managers who, given this compensation, prefer to work in the speculative sector, and $F(S_t\text{-}M_t)$ the mass of investors who also choose that sector. M_t^* is the market clearing expected compensation and R_t the rent which must be left to managers to incentivize effort.

Panel A:
When $M_t^* > Rt$ there is no rationing.

Panel B:
When $M_t^* < Rt$ there is rationing.

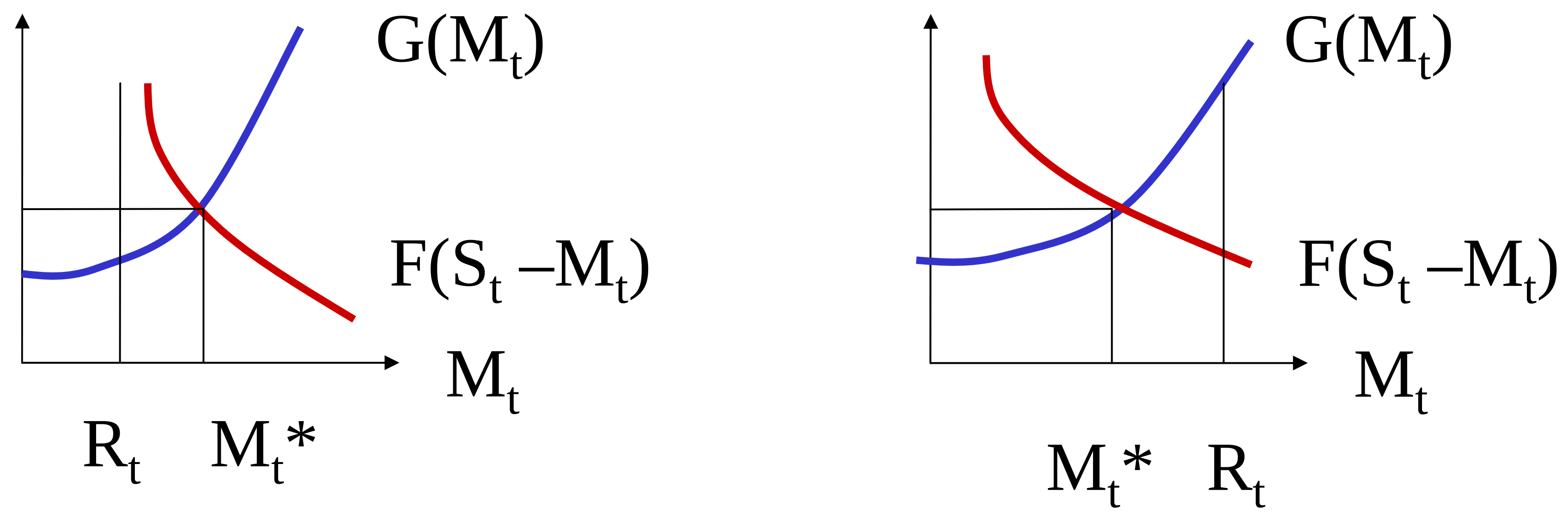


Figure 3: Pledgeable income & expected profit

P_t is the pledgeable income. $(1-\beta)S_t$ is the equilibrium expected profit of investors if IC does not bind and effort is exerted. $(1-\beta)(\hat{S}_t+B)$ is the expected profit of investors without effort.

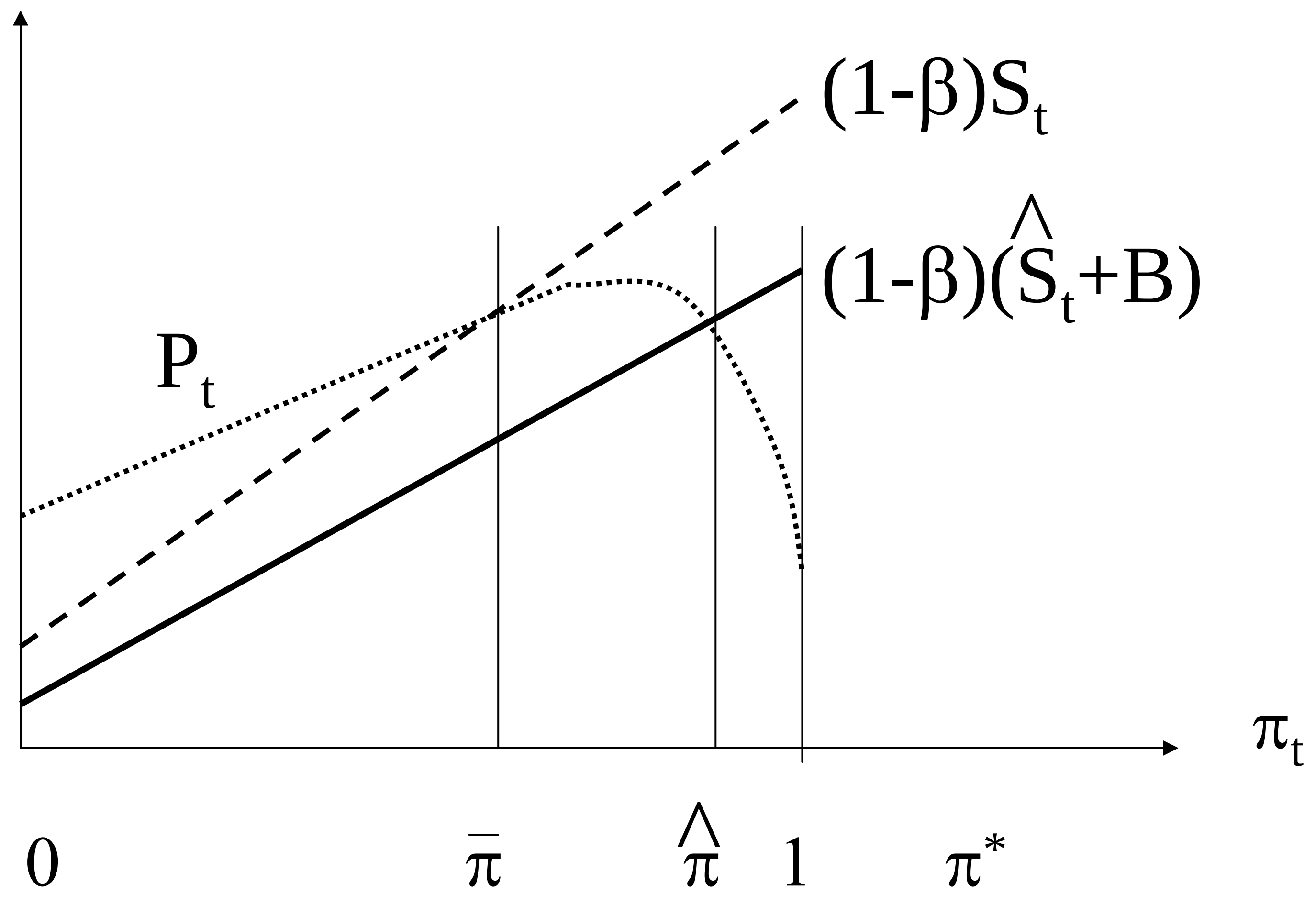
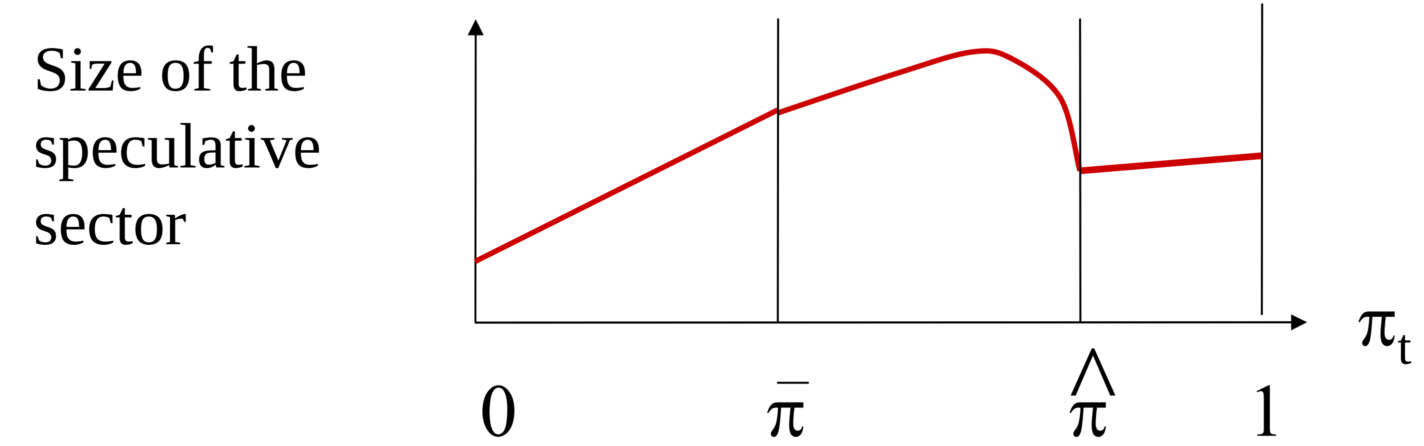
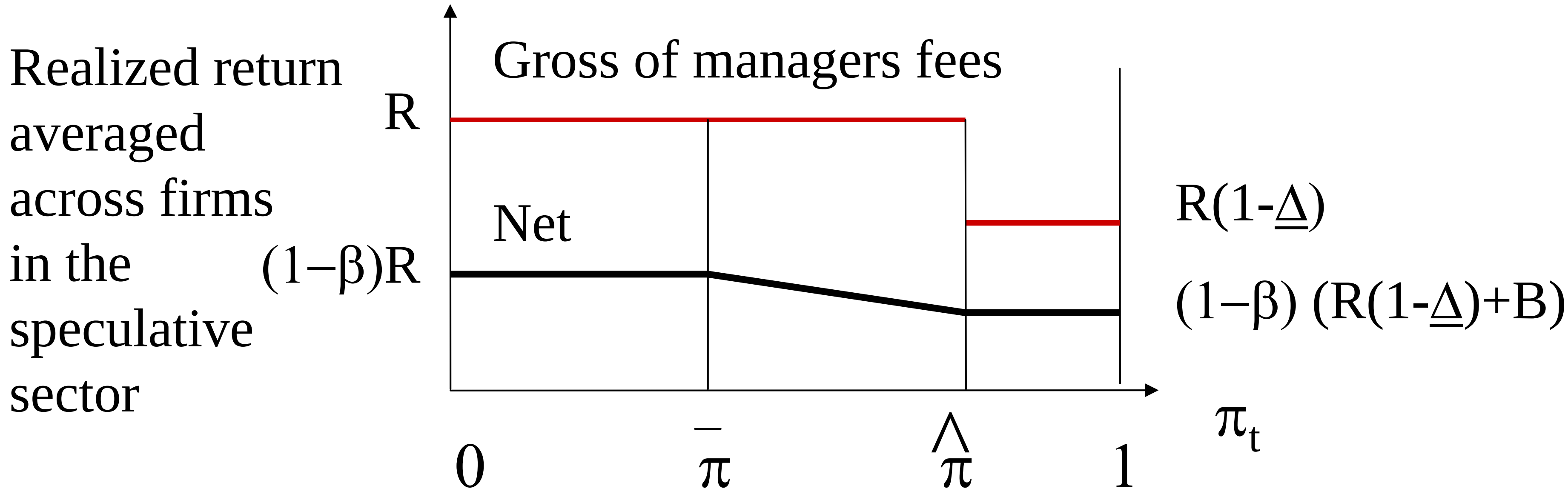


Figure 4: The lifecycle of the speculative sector

When the industry is solid, there is no crisis, π_t tends towards 1 and the figures below illustrate the evolution of the speculative industry. When the industry is fragile, at any point a crisis can occur, with probability p. In that case π_t goes to π_0 and remains at that point.





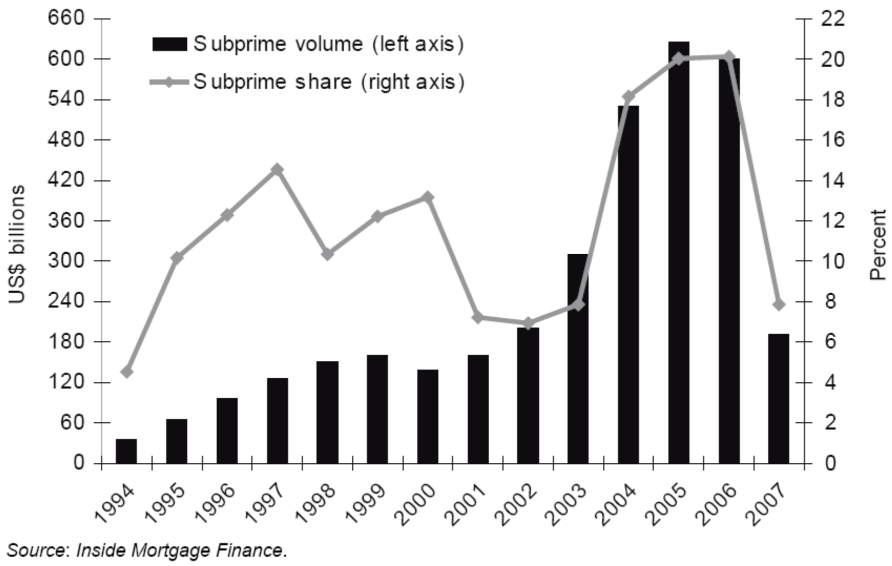


Figure 5: Subprime Mortgage Originations, Annual Volume and Percent of Total

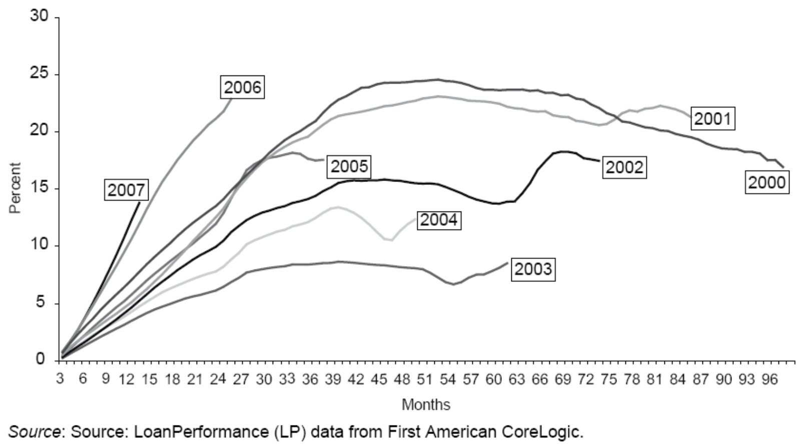


Figure 6: Subprime Delinquency Rate 60+ Days, By Age and Year of Origination