The London School of Economics and Political Science

Essays on the Impact of Competition on Financial Intermediaries

Pragyan Deb

A thesis submitted to the Department of Finance of the London School of Economics for the degree of Doctor of Philosophy, London, July 2012

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of approximately 38,000 words.

Statement of conjoint work

- I confirm that the chapter Credit Rating and Competition was jointly co-authored with Nelson Camanho and Zijun Liu.
- I confirm that the chapter Competition, Premature Trading and Excess Volatility was jointly co-authored with Bonsoo Koo and Zijun Liu.

Abstract

The aim of my thesis is to investigate the effect of competition on financial intermediaries in light of the conflicts of interest and perverse incentive structures that exist in the financial system.

The first chapter of my thesis, ***Credit Rating and Competition*** investigates the conflict of interest arising from the issuer pay compensation model of the credit rating industry using a theoretical model of competitive interaction. Rating agencies balance the benefits of maintaining reputation (to increase profits in the future) and inflating ratings today (to increase current profits). Our results suggest that, unless new entrants have a higher reputation vis-à-vis incumbents, rating agencies are more likely to inflate ratings under competition relative to monopoly, resulting in lower expected welfare.

The second chapter, ***Market Frictions, Interbank Linkages and Excessive Interconnections***, studies banks' decision to form financial interconnections. I develop a model of financial contagion that explicitly takes into account the possibility of crisis. This allows me to model the network formation decision as optimising behaviour of competitive banks. I show that regulatory intervention in the form of deposit insurance and more implicit too big to fail type perceptions of government guarantees creates a wedge between social and private optimality. In the presence of these implicit and explicit guarantees, competitive banks find it optimal to form socially suboptimal interconnections in equilibrium.

The final chapter, ***Competition, Premature Trading and Excess Volatility***, attempts to explain the empirically observed excess asset price volatility as a consequence of competitive interaction between market participants. Our model shows that in the presence of competitive pressures, market participants find it optimal to act prematurely on unverified, noisy information. This premature reaction leads to lower total profits, excess market volatility and spike in volatility at the closing time of the market.

Acknowledgement

My heartfelt gratitude goes out to my supervisor, Amil Dasgupta who is responsible for the successful completion of my dissertation. His encouragement, guidance and support helped me greatly in the understanding and writing of the thesis. In addition, I would like to thank Willem Buiter who was my supervisory during the initial stages of my PhD and helped me formulate my research agenda.

I would also like to thank Bo Becker, Sudipto Bhattacharya, Mike Burkart, Daniel Ferreira, Paolo Fulghieri, Stephane Guibaud, Rainer Haselmann, João Mergulhão, Yves Nosbusch, Filippos Papakonstantinou, Jean-Charles Rochet, Joel Shapiro, Dimitri Vayanos, David Webb and Konstantinos Zachariadis for their helpful comments and discussions.

I am also greateful to conference participants at the 22^{nd} Australasian Finance and Banking Conference (2009), the 3^{rd} Swiss Winter Conference on Financial Intermediation (2010), and the American Finance Association annual meetings in Denver (2011), Bank of England Financial Stability Seminar (2011), 4^{th} International Finance and Banking Society Conference (2012) and various Financial Markets Group workshops for helpful comments and discussions.

Contents

List of Figures

Credit Rating and Competition

Nelson Camanho *Pragyan Deb* *Zijun Liu*

n.c.costa-neto@lse.ac.uk *p.deb@lse.ac.uk* *z.liu@lse.ac.uk*

Financial Markets Group

London School of Economics and Political Science

Abstract

We develop a theoretical model to analyse the effect of competition on the conflict of interest arising from the issuer pay compensation model of the credit rating industry. We find that relative to monopoly, rating agencies are more likely to inflate ratings under competition, resulting in lower expected welfare. These results do not depend on the presence of ratings shopping as in Bolton, Freixas, and Shapiro (2012) and Skreta and Veldkamp (2009), but instead focus on the trade-off between maintaining reputation (to increase profits in the future) and inflating ratings today (to increase current profits). Our results suggest that ongoing regulatory initiatives aimed at increasing competition in the ratings industry may reduce overall welfare, unless new entrants have a higher reputation via-à-vis incumbents.

Keywords: Rating agencies, competition, reputation, repeated games, financial regulation

JEL Classifications: C73, D43, D82, D83, G24

1 Introduction

The credit rating industry aims to offer investors valuable information about issuers in need of financing. Due to the asymmetric information between the issuers and the investors, credit ratings often have pivotal impacts on the issuers' financing outcomes. Before the 1970s, the rating agencies relied on an investor-pay model wherein investors subscribed to ratings released by the agencies and these subscription revenues were the main source of income for the rating agencies. However owing to the 'public good' nature of ratings[1] and the increase in free riding, rating agencies switched to the current issuer-pay model and started charging issuers for ratings. As things stand today, the largest source of income for the rating agencies[2] are the fees paid by the issuers the rating agencies are supposed to impartially rate.[3] This tempts rating agencies to rate better than what fundamentals suggest.

Such behaviour has been criticised heavily since the onset of the recent financial crisis, in particularly over the AAA ratings that have been issued to complex structured products. Rating agencies played a crucial role in the rapid growth of structured finance. According to Fitch Ratings (2007), around 60% of all global structured products were AAA-rated, compared to less than 1% for corporate and financial issues. Following a subsequent jump in default rates, rating agencies lowered the credit ratings on structured products widely, indicating that the initial ratings were likely inaccurate.

A number of empirical papers find that the conflicts of interest problem play an important role in rating agencies' decisions. Griffin and Tang (2011) give striking empirical evidence of ratings inflation by rating agencies. They compare the CDO assumptions made by the ratings department and by the surveillance department within the same rating agency, and find the former uses more favorable assumptions. Moreover, it appears that the signals from the surveillance department were ignored and the CDOs favored by the ratings department were subsequently downgraded. Xia and Strobl (2012) provide

[1]This was officially recognised by the Securities and Exchange Commission (SEC) in the 1970s when the big three rating agencies – Standard & Poor's, Moody's and Fitch were designated self-regulatory entities. See Lowenstein (2008).

[2]It is also interesting to note that rating agencies are some of the most profitable businesses. Moody's was the third most profitable company in the S&P 500-stock index from 2002 to 2007, based on pretax margins (ahead of both Microsoft and Google).

[3]Summary Report of Issues Identified in the Commission Staff's Examinations of Select Credit Rating Agencies by the Staff of the Securities and Exchange Commission, 2008, p.9.

further evidence of ratings inflation as a result of the issuer-pay model. They compare the ratings issued by Standard & Poor's Ratings Services (S&P) which follows the issuer-pay model to those issued by the Egan-Jones Rating Company (EJR) which adopts the investor-pay model. They find that S&P inflates more relatively to EJR when S&P's conflict of interest is more acute.

It is often suggested that introducing more competition between rating agencies may help alleviate the conflicts of interest problem. However, a growing body of academic literature suggests that this may not be the case. Skreta and Veldkamp (2009) show that, in the presence of asset complexity and ratings shopping, competition leads to lower welfare in equilibrium. Bolton, Freixas, and Shapiro (2012) also find that competition leads to more ratings inflation as issuers are able to more easily shop for ratings and that this effect is particular acute in boom times, when investors are more trusting. The contribution of our paper is to show that even in the absence of ratings shopping and asset complexity, and with rational investors, competition delivers lower welfare than monopoly. Our results stem from the fact that enhanced competition in the form of a new entrant reduces the incumbent's market share for ratings. This market sharing effect reduces the rent that rating agencies can derive from maintaining their reputation, encouraging ratings inflation even in the absence of ratings shopping. Our results suggest that current regulatory attempts to reduce ratings shopping[4] may not eliminate ratings inflation due to the underlying conflicts of interest problem.

We develop an infinite horizon model where rating agencies compete for market share and face a trade-off between reputation and current fees. Competition in our model has two effects - the disciplining effect and the market-sharing effect. Competition decreases ratings inflation through the disciplining effect as rating agencies have incentives to maintain or gain the market leadership. This channel is generally emphasized when it is argued that enhanced competition between rating agencies can resolve the conflict of interest. However, this ignores the other effect of competition - the reward from maintaining reputation is lower because competition implies that the market is shared between a larger number of rating agencies. We call this the market-sharing effect and study the impact of competition on the behaviour of rating agencies by exploring the interaction

[4]See Sangiorgi and Spatt (2011). Note that in a rational expectations setting, ratings inflation might arise due to the possibility of unpublished ratings, which might be countered by regulation.

between these two opposite effects. Our results suggest that on balance the latter effect dominates and higher competition results in greater ratings inflation.

Given the structure of the market - with S&P's and Moody's having 80% of market share,[5] we model competition amongst the rating agencies in a duopolistic setting. In our model, issuers need a good rating to finance their projects. Rating agencies, which can be of two types - *honest* or *strategic*, perfectly observe the quality of the project and can either give the issuer a good rating or refuse rating. An honest rating agency always gives good ratings to good projects and no rating to bad projects while a strategic rating agency acts to maximise its expected profits. Neither investors nor issuers know for sure if a rating agency is honest and they Bayesian update on the reputation of the rating agencies, *i.e.* the probability that a rating agency is honest. The market share of the rating agency is modeled such that rating agencies with higher reputation attract more projects. Hence the rating agencies face a trade-off between current income and reputation which determines their future market share and income.

We compare the behaviour of rating agencies between the duopolistic case and the monopolistic case.[6] We first derive closed-form solutions in a three-period model and show that the lax behaviour of a rating agency increases with the reputation of its competitor, *i.e.* competition leads to more lax behaviour and the market-sharing effect dominates. We then compute numerical solutions under an infinite-period setting, which enables us to relax parameter restrictions and extend the horizon of rating agencies, thereby making reputation more important for them.

Our results show that the market-sharing effect tends to dominate the disciplining effect when the degree of competition is sufficiently high, *i.e.* the reputation of the competitor is high. Moreover, we find that expected welfare is higher in the monopoly case than in the duopoly case as long as the reputation of the entrant rating agency (the competitor) is not greater than that of the incumbent rating agency. In our model, expected welfare rises only when the new entrant has a higher reputation vis-à-vis the incumbent, a situation which appears unlikely. We verify that the results are robust to different parameter specifications and on balance, our results suggest that increasing

[5] The figure stands at 95% if we include the third major player, Fitch.

[6] Although we only focus on competition in a duopolistic setting, our results intuitively extend to situations with higher degrees of competition.

competition is likely to result in more ratings inflation.

The rest of the paper is organised as follows. Section 2 reviews the literature. Section 3 outlines the basic features of our model and Section 4 describes the equilibrium. We present the infinite horizon solution of our model in Section 5 and compare the behaviour of rating agencies under monopoly and duopoly. This section also discusses the expected welfare consequences of enhanced competition. Section 6 concludes. The proofs and additional robustness checks, including a finite horizon analytical solution of our model in a three-period setting, are presented in the Appendix.

2 Literature Review

Mathis, McAndrews, and Rochet (2009) demonstrate that reputational concerns are not enough to solve the conflict of interest problem. In equilibrium, rating agencies are likely to behave laxly, *i.e.* rate bad projects as good and are prone to reputation cycles. Our model innovates by introducing competition through an endogenous market share function and studying how competition affects the behaviour of rating agencies.

Becker and Milbourn (2011) lends support to our results by providing an empirical test of the impact of competition on rating agencies. They measure competition using the growth of Fitch's market share and find three pieces of evidence. First, the overall standards of ratings issued by S&P and Moody's increased (closer to the top AAA rating) with competition, so that ratings became more 'friendly'. Second, the correlation between bond yields and ratings fell as competition increased, implying that ratings became less informative. Third, equity prices started reacting more negatively to rating downgrades, suggesting a lower bar for rating categories. Their findings are consistent with our results that competition will tend to lower the quality of ratings in the market.

A recent paper by Xia (2012) provides some contrasting empirical evidence. The author compares S&P's rating quality before and after the entry of an investor paid rating agency and finds a significant improvement in the quality of S&P's ratings following the entry of the new rating agency. This result however is completely compatible with our model since an investor paid rating agency in our setting would be perfectly honest and our results suggest that in cases in which the incumbent RA has lower reputation than the entrant RA, welfare improvement is possible.

There has been an extensive literature that studies competition through reputation. For example, Horner (2002) shows that the incentive to maintain good reputation and stay in the market can induce good firms to exert higher effort and try to distinguish themselves from the bad ones. The adverse effects of competition on the building and maintenance of reputation has been studied by Klein and Leffler (1981). They argue that when faced with a choice between supplying high quality products or low quality ones, firms would be induced to supply high quality products only when the expected value of future income given a high reputation outweighs the short-run gain of lying. Bar-Isaac (2003) points out that the overall effect of competition on reputational incentives is ambiguous and may be non-monotonic, since increased competition can reduce the discounted value of maintaining a high reputation on the one hand, but can also lead to a more severe punishment for low reputation on the other. This intuition is very close to ours, except that we use a richer framework in the context of credit rating agencies.

Bouvard and Levy (2009) examine the trade-off between reputation and profits of rating agencies in a competitive setting and find that the threat of entry attenuates reputational effects. Mariano (2012) models how reputational concerns change rating agencies incentives to reveal private information. In a setting in which rating agencies have access to private and public information, her results provide a mechanism in which competition between rating agencies might inflate the ratings even in the absence of conflicts of interest. Compared to the above, the innovation of our paper is to endogenise the market share of rating agencies and to explore the welfare implications of competition.

Damiano, Hao, and Suen (2008) study how the rating scheme may affect the strategic behaviour of rating agencies. They compare ratings inflation among centralised (all firms are rated together) and decentralised (firms are rated separately) rating schemes. When the quality of projects is weakly correlated, centralised rating dominates because decentralised rating leads to lower ratings inflation. The reverse holds when the correlation is strong. Sangiorgi, Sokobin, and Chester (2009) model and analyse the equilibrium structure of ratings reflected by ratings shopping. They interpret how the correlation between different rating agencies' models influence ratings shopping and bias. They also use selection as an equilibrium interpretation for notching by a rival rating agency. Moreover, they show that a higher cost of obtaining indicative ratings lead to inflation in published ratings, as they are obtained less frequently.

Ashcraft, Goldsmith-Pinkham, and Vickery (2010) study credit ratings on subprime and Alt-A mortgage-backed-securities (MBS) deals issued between 2001 and 2007. Although they find that the fraction of highly rated securities in each deal is decreasing in mortgage credit risk, their results suggest a progressive decline in standards around the MBS market peak between the start of 2005 and mid-2007.

White (2010) gives a historic overview of the market of the credit rating agencies and suggest that the regulatory framework contributed to the subprime mortgage debacle and associated financial crisis. They highlight how the major reliance of regulators on major rating agencies propelled them to the centre of US bond markets and led the mistakes by those rating agencies to have serious consequences for the financial sector.

Bar-Isaac and Shapiro (2011) explore how the labour market for analysts and their incentives influence ratings accuracy. Motivated by the fact that rating analysts were fleeing the rating agencies for better paid investment bank jobs, they build a 2 period model in which analysts work for rating agencies in period 1 and can leave them to a better paid investment bank in period 2. They show that ratings accuracy increases with monitoring and also with investment bank profitability (as analysts train harder in period 1), but it is non-monotonic in the probability of the analyst getting a job in the investment bank.

Bar-Isaac and Shapiro (2012) analyse how reputational concerns of rating agencies vary over the business cycle. A rating agency is more likely to issue less-accurate ratings in boom times, when income from fees is high, competition in the labour market for analysts is tough, and default probabilities for the securities rated are low. They also show that competition among the rating agencies delivers similar qualitative results. However, competition is not the main focus of their paper and is modelled through an exogenous function between the degree of competition and the fees received by rating agencies.

3 Model Setup

We consider a discrete time setting with 3 types of agents – the issuers, the rating agencies (RA) and the investors. Each period, we have a *new issuer*[7] with a project that requires financing. We assume that issuers do not have funds of their own and need to obtain outside financing. The investors have funds and are willing to invest in the project provided they are convinced that it is profitable to do so. The role of the RA in this setting is to issue ratings that convince investors to provide financing.

More formally, each period we have one issuer that has a project which lasts for one period. All projects have a fixed pay-off Φ if successful and 0 otherwise and require an investment of X. We assume that the required investment X is uniformly distributed over (a,b) and its realisation is observed by all agents. This ensures that we have a range of projects with different returns - projects that require low investment have high return and vice versa. We can get similar results if we assume fixed investment with uncertain pay-off. The project is *good* with probability λ and *bad* with probability $1 - \lambda$, and λ is independent of X. Good projects succeed with probability p_G and fail with probability $(1 - p_G)$. Bad projects always fail.

We assume that *a-priori* projects are not worth financing without rating, *i.e.* $\lambda p_G \Phi \leq X$. Further, the RAs can perfectly observe the type of project at no cost. After observing the type, the RA can either issue a good rating (GR) or no rating (NR). Note that we do not distinguish between bad rating and NR and abstract away from a ratings scale. In our setup, a good rating is one that allows the issuer to borrow from investors. It does not matter if this rating is AAA or A or BBB or even C. As long as the rating allows the firm to get financing, we consider it to be a GR. A bad rating in this setting will be a rating which does not enable a project to get financing. This is the same outcome as a NR and thus, a bad rating and NR are equivalent in our model.

The rating agency receives income I if it issues GR, and 0 otherwise.[8] This assumption arises from the conflict of interest in the ratings industry. Given the *non-transparent*

[7] *New Issuer* implies that it is a one shot game for the issuer and we rule out the possibility that issuers try to maximise profits over multiple periods. This assumption also ensures that issuers have the same belief as the investors about the reputation of the RAs. If we allow the same issuers to approach the rating agencies in subsequent periods, then issuers will have more information than investors.

[8] This is a standard simplifying assumption in the literature. See Mathis, McAndrews, and Rochet (2009) and Skreta and Veldkamp (2009).

nature of the market and the widespread use of *negotiated ratings*, issuers and RAs routinely have negotiations and consultations before an official rating is issued. RAs, as part of their day-to-day operations, give their clients 'creative suggestions' on how to repackage their portfolios or projects in order to get better ratings. To quote former chief of Moody's, Tom McGuire[9]

> "The banks pay only if [the rating agency] delivers the desired rating... If Moody's and a client bank don't see eye-to-eye, the bank can either tweak the numbers or try its luck with a competitor..."

Following Mathis, McAndrews, and Rochet (2009), we assume that only the RA knows the type of project. This simplyfying assumption can be motivated by the market for structured products, for which RAs played a vital advisory role. For structured products, RAs routinely advise issuers about the level of credit credit enhancment needed to obtain a good rating.

We model two types of RAs - *honest* and *strategic* . An honest RA always issues a GR to a good project and NR to a bad project while a strategic RA behaves strategically to maximise its expected future profits. The strategic RA faces the following trade-off :

1. (**Truthful**) It can either be truthful and maintain its reputation, thus ensuring profits in the future

2. (**Lie**) It can inflate ratings (give a good rating to a bad project) and get fees now, at the cost of future profits

We consider a duopolistic setting of rating agencies.[10] The type of the RA is chosen *ex ante* by nature and is known only to the rating agency itself. The *reputation* of the rating agency is defined as the probability that it is honest, denoted by $q_i, i \in \{1, 2\}$. The reputation evolves over time depending on the ratings and outcome of the projects. The *strategy* of the RA is x_i, the probability the RA issues a GR to a bad project.[11]

[9] New York Times Magazine, Triple-A-Failure, April 27, 2008.

[10] Given the structure of the market, with Moody's and S&P controlling nearly 80% of the market, we believe that this is a reasonable approximation of reality.

[11] Note that in equilibrium the strategic RA will always issue GR to a good project (see section 4).

The investors (and issuers) have some priors about the types of the RAs and they Bayesian update on their beliefs. Firstly, investors and issuers take into account the rating and update the reputation of the RA, before observing the outcome of the project. Given prior reputation q_t,

$$\text{If RA issues GR, } q_t^{GR} = \frac{\lambda q_t}{\lambda + (1-q_t)(1-\lambda)x} < q_t \tag{1}$$

$$\text{If not rated, } q_{t+1}^{N} = \frac{q_t}{1 - x(1-q_t)} > q_t \tag{2}$$

If the project is issued a good rating by the RA, the investors update their beliefs after observing the outcome of the project.

$$\text{If the project succeeds, } q_{t+1}^{S} = \frac{\lambda p_G q_t}{\lambda p_G q_t + \lambda p_G (1-q_t)} = q_t \tag{3}$$

$$\text{If the project fails, } q_{t+1}^{F} = \frac{\lambda(1-p_G)q_t}{\lambda(1-p_G)q_t + [\lambda(1-p_G) + (1-\lambda)x](1-q_t)} < q_t \tag{4}$$

We make the simplifying assumption that each issuer can only approach one RA for rating. Therefore, our model considers ratings shopping only to the extent that the issuer and the rating agency have negotiations before an official rating is issued. We do not explicitly study multiple ratings and herd behaviour of the RAs. While these are important issues that merit attention, they are not the focus of this paper. Here we look at the competition for market share among rating agencies and show that ratings inflation increases with competition.

Investors observe the rating decision and decide whether to invest. If they observe a GR from a RA with reputation q, their subjective belief that the project will succeed (using equation (1)) is given by

$$\begin{aligned} s(q,x) &= q^{GR} p_G + (1-q^{GR}) \frac{\lambda p_G}{\lambda + (1-\lambda)x} \\ &= \frac{\lambda q}{\lambda + (1-q)(1-\lambda)x} p_G + \left(1 - \frac{\lambda q}{\lambda + (1-q)(1-\lambda)x}\right) \frac{\lambda p_G}{\lambda + (1-\lambda)x} \\ &= \frac{\lambda p_G}{\lambda + (1-q)(1-\lambda)x} \end{aligned} \tag{5}$$

Given the required investment level X, investors are willing to finance the project *if and* *only if* $X \leq s(q,x)\Phi$, *i.e.* if the initial investment required for the project is no greater than its expected pay-off. Without loss of generality, assume $s(q_1,x_1) > s(q_2,x_2)$. We have 3 cases:

1. If X is such that a good rating from either RA is enough, *i.e* $X \leq s(q,x)\Phi$ for both q_1 and q_2, the firm can approach either RA.[12] We assume that in this case the firm will randomly choose one of the RAs, *i.e.* the project goes to both RAs with equal probability.[13]

2. If $s(q_2,x_2)\Phi < X < s(q_1,x_1)\Phi$, *i.e.* only the high reputation RA can issue ratings that can convince the investors to provide financing, hence the firm will go to RA1 and not RA2.

3. If $X > s(q_1,x_1)\Phi$, the project does not get financed.

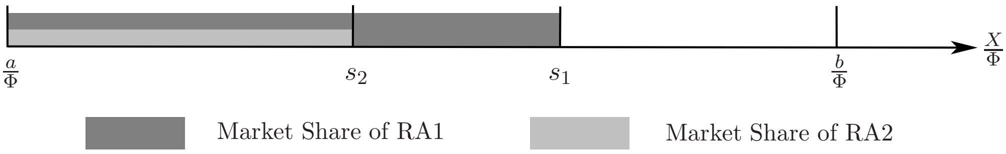


Figure 1: The Market for Ratings

Thus we get the following result as illustrated in Figure 1 -

$$\text{Probability that a project comes to RA1} = \frac{(s_1 - s_2) + \frac{1}{2}\left(s_2 - \frac{a}{\Phi}\right)}{\frac{b}{\Phi} - \frac{a}{\Phi}}$$
$$\text{Probability that a project comes to RA2} = \frac{\frac{1}{2}\left(s_2 - \frac{a}{\Phi}\right)}{\frac{b}{\Phi} - \frac{a}{\Phi}}$$

We set $(a,b) = (\lambda p_G \Phi, p_G \Phi)$, because any project with $X < \lambda p_G \Phi$ does not need a rating to be financed, and any project with $X > p_G \Phi$ is never worth financing *ex-ante*.

$$\text{The probability that a project comes to RA1} = \frac{s_1 - \frac{1}{2}\left(s_2 + \lambda p_G\right)}{p_G(1-\lambda)} \tag{6}$$
$$\text{The probability that a project comes to RA2} = \frac{\frac{1}{2}\left(s_2 - \lambda p_G\right)}{p_G(1-\lambda)} \tag{7}$$

[12]We assume that the issuers are only paid when projects succeed. This implies that the issuers will be indifferent between RAs (with different reputation) given that both can guarantee financing.

[13]Note that this is one of infinite many possible equilibria. Since the issuers are indifferent, we have an equilibrium for all probabilities ($\alpha \in (0,1)$) of approaching a specific RA . We focus on the case where $\alpha = \frac{1}{2}$. Our qualitative results do not depend on the choice of α.

Reputation plays a critical role in our model. The market share of the RAs depends on s, and thus on reputation q. Since the income from giving a GR is constant (denoted by I), the future profits of the RA will solely depend on its market share.[14] Moreover, the RA with a higher reputation enjoys additional benefits of being the market leader, because it owns entirely the proportion of the market that cannot be rated by its competitor but can be rated by itself, whereas its competitor can only share its market with the leader. This creates incentives for RAs to maintain or gain the market leader position and hence disciplines the RAs through competition.

We can now see that competition (modelled through market share) has two effects on lax behaviour: the market-sharing effect and the disciplining effect. The market-sharing effect refers to the fact that the RA finds lying and receiving income today more attractive as its expected future income is shared with another RA, and the disciplining effect refers to the fact that the RA finds lying less attractive in order to maintain/gain the advantages of being a market leader. We will show later that the market-sharing effect tends to dominate the disciplining effect and hence competition aggravates the lax behaviour of RAs in general.

4 Equilibrium

Definition 1. *The equilibrium in our model is a set of Markov Perfect strategies such that, at each period t, the strategic RA always*

(i) Gives a good rating to a good project.

(ii) Gives a good rating to a bad project with probability x_t, where $0 \leq x_t \leq 1$.

We look for a Markov Perfect Equilibrium in the sense that the equilibrium is "memoryless", *i.e.* the strategy of the strategic RA only depends on the current reputation of the two RAs. The equilibrium is also "symmetric", as the strategy function of both RAs (if they are both strategic) is the same. However, the RAs do not take actions simultaneously.

Let RA1 be a strategic RA and let $V_t(q_1, q_2)$ denote its discounted future profits,

[14] The assumption of fixed I allows us to cleanly model competition through market share but is not critical to our results. All our results go through if we assume that the fee is increasing in reputation. The plots are available upon request.

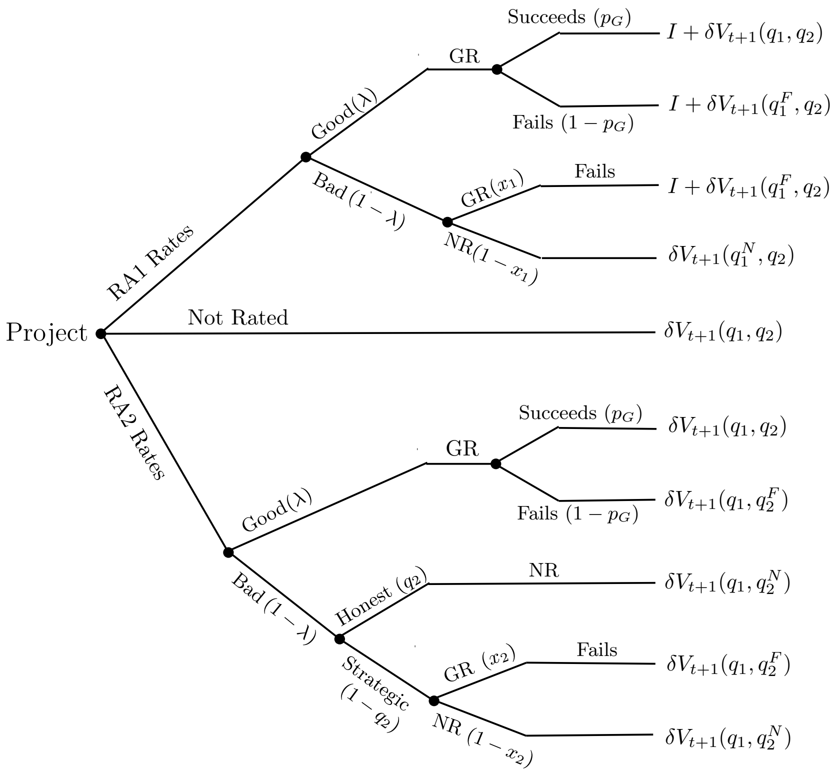


Figure 2: Decision-tree for Strategic RA1

given its reputation q_1 and its competitor's reputation q_2, and let δ be the discount rate. The RA's new reputation after it gives NR and the failure of a project following a GR are denoted by q_1^N and q_1^F respectively. A successful project with a GR leaves the RA's reputation unchanged. Note that q_1^F and q_1^N are functions of the strategy of the RA and its current reputation level. For notational simplicity, we suppress the time subscript of these reputation-updating functions.

Figure 2 shows the decision tree of RA1. Suppose it is approached for rating. If the project is good, RA1 gives it a GR and gets income I (see Proposition 2 below). On the other hand, if the project is bad, RA1 strategically decides whether to give a GR and get fees I or refuse rating. In case of NR, RA1's reputation rises as it gets a larger market share in the future. In case of a GR, RA1's reputation falls if the project fails and remains the same if it succeeds. This in turn determines the RA1's expected future income. A similar analysis applies if RA2 is approached for rating. In this case the fees go to RA2 and RA1 is only indirectly affected through a change in RA2's reputation. Note that since RA1 does not know the type of RA2, it has to take into account the possibility that RA2 is either honest or strategic.

$$
\begin{aligned}
V_t(q_1, q_2) = P(RA1rates)\Big\{ & P(Good)\big[I + p_G\delta V_{t+1}(q_1, q_2) + (1 - p_G)\delta V_{t+1}(q_1^F, q_2)\big] \\
& + P(Bad)\big[x_1(q_1, q_2)\big(I + \delta V_{t+1}(q_1^F, q_2)\big) + \big(1 - x_1(q_1, q_2)\big)\delta V_{t+1}(q_1^N, q_2)\big]\Big\} \\
& + P(RA2rates)\Big\{ P(Good)\big[p_G\delta V_{t+1}(q_1, q_2) + (1 - p_G)\delta V_{t+1}(q_1, q_2^F)\big] \\
& + P(Bad)\Big[(1 - q_2)x_2(q_1, q_2)\delta V(q_1, q_2^F) + \big[q_2 + \big(1 - q_2\big)\big(1 - x_2(q_1, q_2)\big)\big]\delta V(q_1, q_2^N)\Big]\Big\} \\
& + P(NotRated)\delta V_{t+1}(q_1, q_2) \quad (8)
\end{aligned}
$$

The objective function of RA1 is to maximise $V_t(q_1, q_2)$, the strategy being x_1. Note that RA1's strategy is only effectual when it rates a bad project. In all other cases, RA1's strategy is inconsequential.

Proposition 1. *There exists a unique x_1, where $0 \leq x_1 \leq 1$, given that $V_t(q_1, q_2)$ is an increasing function in q_1.*

Proof. See Appendix A.1 □

Intuitively, it is easy to see from equation (8) that $V_t(q_1, q_2)$ is linear in x_1. This ensures that RA1's maximisation problem has a unique solution.

Proposition 2. *A strategic RA does not have incentives to give NR to a good project.*

Proof. See Appendix A.2 □

Proposition 2 implies that a strategic RA always gives GR to a good project. This is because it gets a lower pay-off if it deviates from this strategy and gives a NR to a good project. The proposition follows directly from the pay-off structure of the RAs and the beliefs.

Proposition 3. *There exists a unique equilibrium as described in Definition 1.*

Proof. Follows from Propositions 1 and 2. □

Corollary 1. *Assume $p_G < 1$. Then the equilibrium strategy of the strategic RA is always positive, i.e. it inflates ratings with positive probability.*

Proof. See Appendix A.3 $\square$

Corollary 2. *Suppose the model ends in period T. Then the equilibrium strategy of the strategic RA is $x = 1$ at $t = T-1, T$.*

Proof. See Appendix A.4 $\square$

5 Model Solution

We solve the model numerically in an infinite horizon setting. The numerical solution is computed using backward induction, *i.e.* we first solve the model in the finite period case, and then increase the number of periods so that the equilibrium strategy converges to the infinite horizon solution. In the appendix, we present an analytical solution to the model in a 3-period setting. Our key results continue to hold, but the analytical solution requires additional simplyfying assumptions.

In an infinite period setting, V_t by itself is independent of t. Hence we suppress the time subscript for notational simplicity. However, the reputations evolve over time as investors (and issuers) update their beliefs. Let RA1 be the rating agency that behaves strategically. Then, RA1's value function takes the following form:

$$
\begin{aligned}
V(q_1, q_2) = \frac{\frac{1}{2}\big(s_1 - \lambda p_G\big)}{(1-\lambda)p_G} \Bigg\{ & \lambda \Big[I + p_G \delta V(q_1, q_2) + (1 - p_G)\delta V(q_1^F, q_2) \Big] + \\
& (1-\lambda)\Big[x_1(q_1, q_2)\big(I + \delta V(q_1^F, q_2)\big) + \big(1 - x_1(q_1, q_2)\big)\delta V(q_1^N, q_2) \Big] \Bigg\} \\
& + \frac{s_2 - \frac{1}{2}\big(s_1 + \lambda p_G\big)}{(1-\lambda)p_G} \Bigg\{ \lambda \Big[p_G \delta V(q_1, q_2) + (1 - p_G)\delta V(q_1, q_2^F) \Big] + \\
(1-\lambda)\Big[(1 - q_2) x_2(q_1, q_2) \delta V(q_1, q_2^F) & + \big[q_2 + \big(1 - q_2\big)\big(1 - x_2(q_1, q_2)\big) \big] \delta V(q_1, q_2^N) \Big] \Bigg\} \\
& + \frac{p_G - s_2}{(1-\lambda)p_G} \delta V(q_1, q_2) \qquad (9)
\end{aligned}
$$

where for $s(q_1, x_1) < s(q_2, x_2)$, $\frac{\frac{1}{2}\left(s_1 - \lambda p_G\right)}{(1-\lambda)p_G}$ is the probability that the issuer approaches RA1 for rating, $\frac{s_2 - \frac{1}{2}\left(s_1 + \lambda p_G\right)}{(1-\lambda)p_G}$ is the probability that the issuer approaches RA2 and $\frac{p_G - s_2}{(1-\lambda)p_G}$ is the probability that the project is not rated by either RA. The expression for $s(q_1, x_1) \geq s(q_2, x_2)$ is analogous.

We assume that the model ends at period T and solve the model backwards. We know that the strategic RA will always lie at period T and $T-1$ according to Corollary 2. For all $t < T-1$, the strategy of the RA depends on its own and its competitors' reputation. We solve for the equilibrium strategy of the RA described in Section 4. We look at the pay-offs from lying and being honest and determine the strategy. As long as $I + V_t(q_1^F, q_2) > V_t(q_1^N, q_2)$ for $x_t = 1$, RA1 will always choose to lie. Conversely, if $I + V_t(q_1^F, q_2) < V_t(q_1^N, q_2)$ for $x_t = 0$, RA1 will always tell the truth. In all other intermediate cases, there exists a unique x_t *s.t.* $I + V_t(q_1^F, q_2) = V_t(q_1^N, q_2)$ at which RA1 is indifferent between lying or not. Hence we deduce inductively the equilibrium strategies of RA1. As T goes to infinity, we approach the infinite horizon solution. Since $\delta < 1$, the Blackwell conditions are satisfied.

Using this procedure, we solve the model for various parameter values. At the first instance, we solve the model for a monopolistic RA. Next, we introduce competition in the form of RA2 and show that the additional competitive element is not sufficient to discipline the RAs. Furthermore, our results show that competition will in fact increase ratings inflation.

5.1 Monopolistic RA

First we consider the case where there is only one RA in the market. In order to make RA1 a monopolist, we set the reputation of RA2 to 0.

Figure 3 plots the strategy of the monopolistic RA for parameters $(\lambda,\ p_G,\ \delta) = (0.5,\ 0.7,\ 0.9)$.[15] We can clearly see the strategy of RA1 is 'u-shaped' in its reputation. Intuitively, the RA's strategy is determined by the trade-off between current fees and

[15] Note that we have chosen this set of parameters $(\lambda,\ p_G,\ \delta) = (0.5,\ 0.7,\ 0.9)$ for the purpose of illustration only, and verified that our results are robust to other parameter specifications, the plot of which are available upon request. In particular, robustness checks of the main results (Section 5.3) are presented in Appendix C.

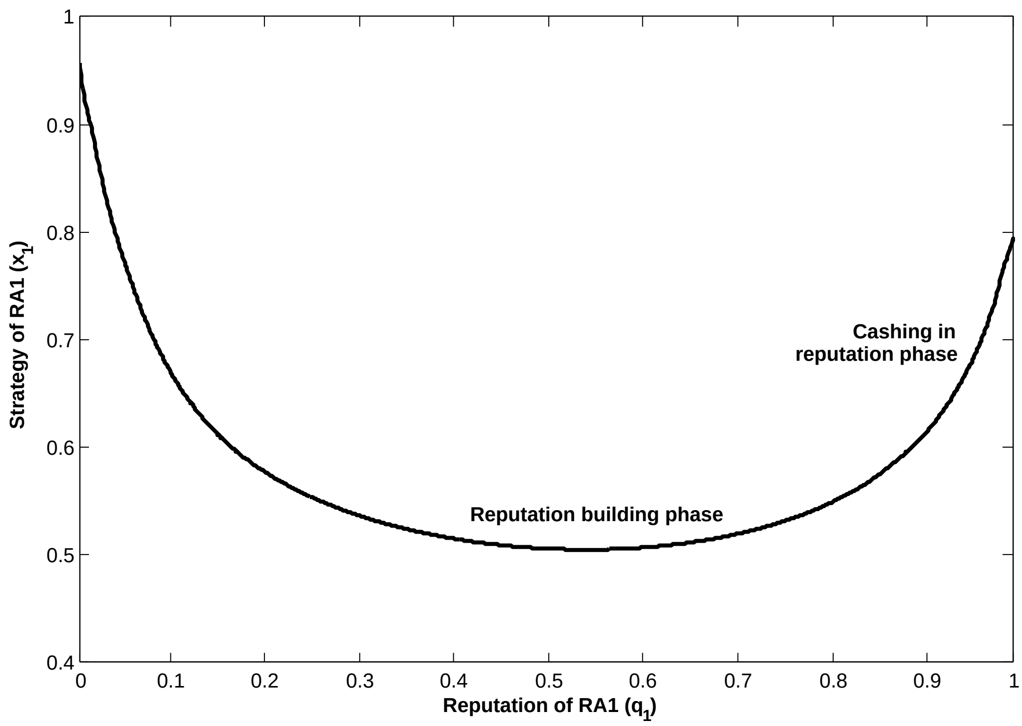
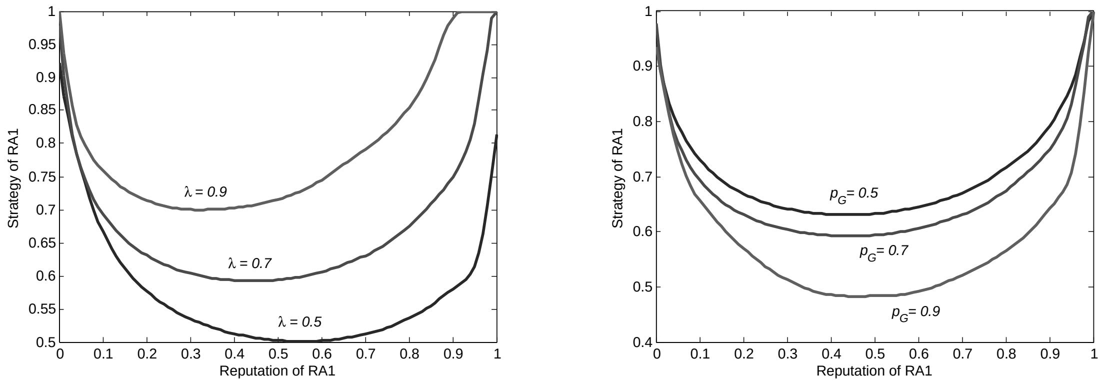


Figure 3: Strategy vs Reputation, Monopolistic RA (λ, p_G, δ, q_2) = (0.5, 0.7, 0.9, 0)

expected future income. When its reputation is very low, the RA's expected future income is very small compared to current fees, hence it has little incentive to behave honestly. When its reputation increases, the RA's future income becomes larger while current fees stay the same, the RA tends to lie less. However, when the RA's reputation is very high, the penalty for lying decreases, and the RA starts to lie more. The reason that the penalty for lying decreases with reputation is that investors attribute project failures to bad luck rather than lax behaviour when they believe that the RA is very likely to be of the honest type.

Moreover, we can see from Figure 4 that the strategy of RA1 is increasing in λ but decreasing in p_G.[16] The intuition is that, the reputational penalty of lying depends on how the investors update their beliefs. If projects are more likely to be good (higher λ) or if good projects are more likely to fail (lower p_G), then a failure is more likely to be attributed to bad luck rather than lying. Anticipating this smaller cost of lying on reputation, the RA would choose to lie more when λ increases or p_G decreases.

[16]We have also verified that this result holds in the case of competitive RAs, the plots of which are available upon request.



(a) Strategy of RA1 for different values of λ ($p_G = 0.7$)

(b) Strategy of RA1 for different values of p_G ($\lambda = 0.7$)

Figure 4: Strategy vs Reputation for different values of λ and p_G ($\delta = 0.9$)

5.2 Competitive RA

We now look at the impact of competition on the behaviour of rating agencies by introducing a second RA (RA2). Figure 5 plots the strategy of RA1 for parameter values (λ, p_G, δ) = (0.5, 0.7, 0.9). Figures 6 and 7 show cross-sections of this figure, for different values of q_2 and q_1 respectively.

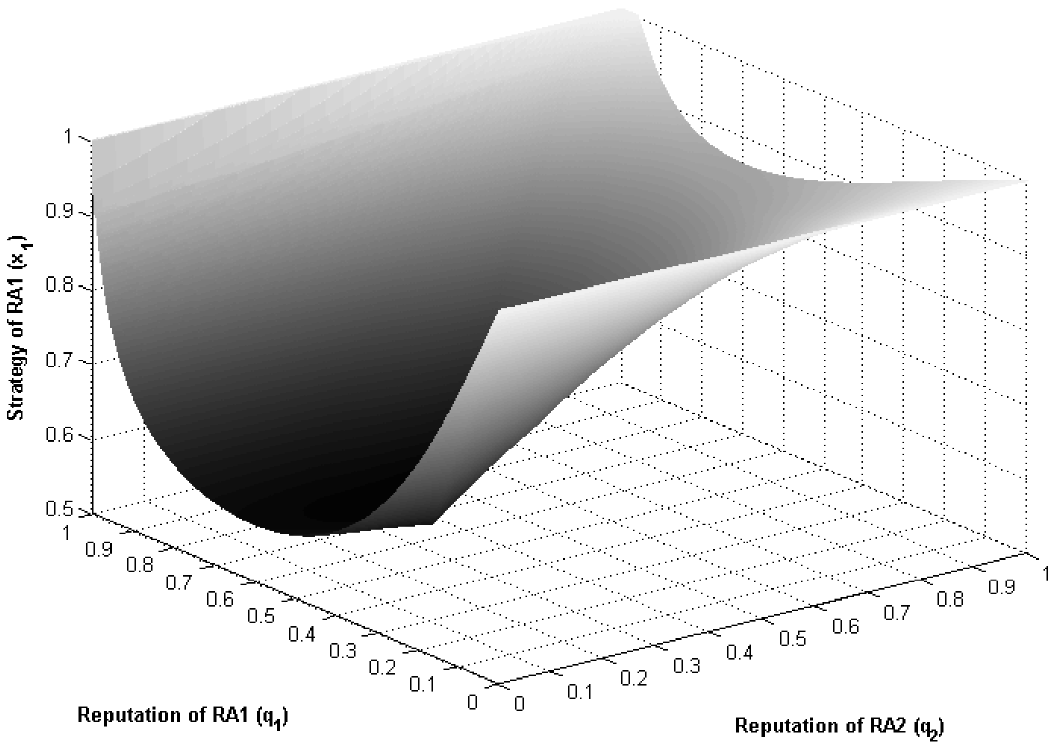
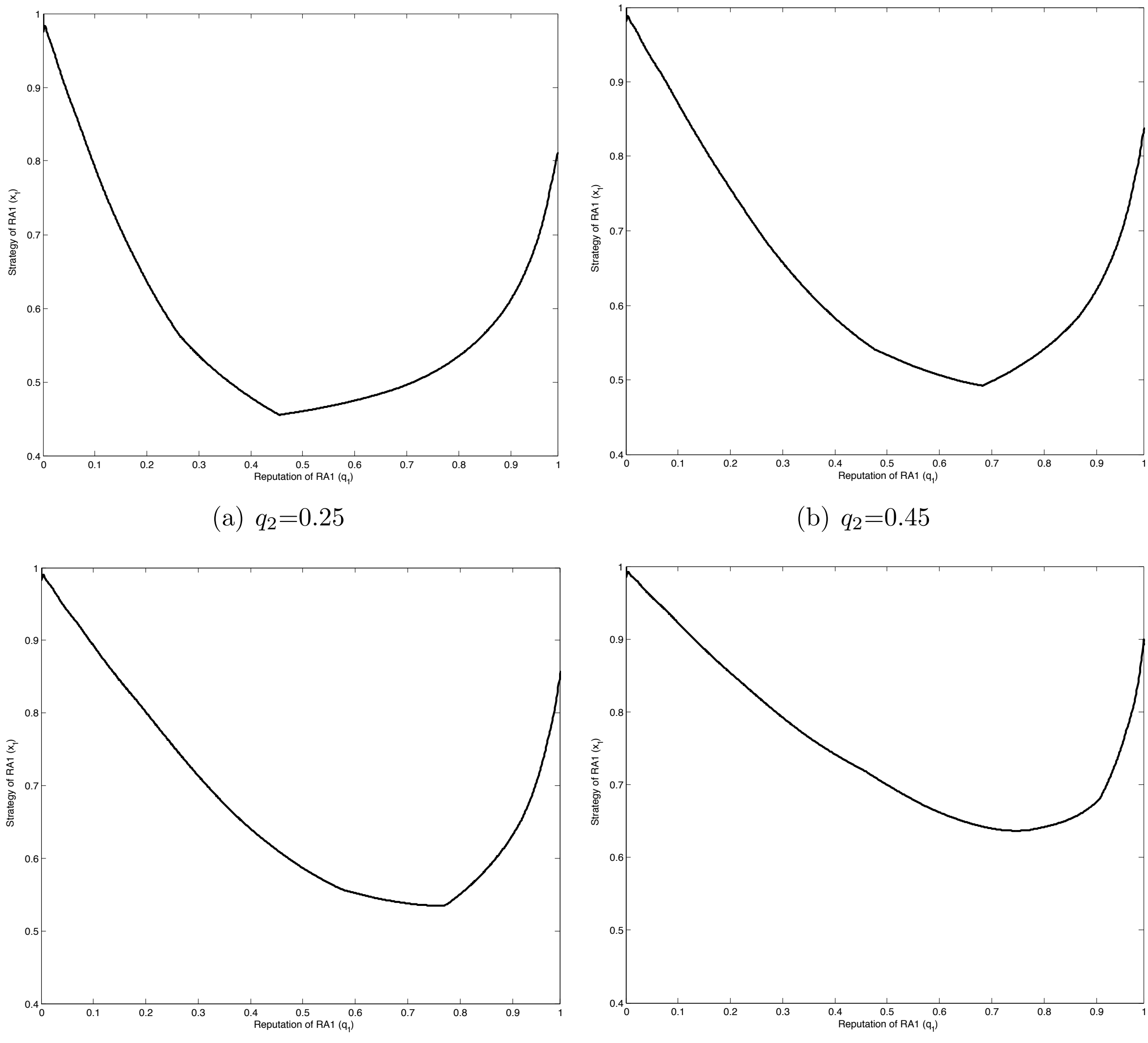


Figure 5: Strategy vs Reputation, (λ, p_G, δ) = (0.5, 0.7, 0.9)

Figure 6 shows the relationship between the reputation and strategy of RA1 for different values of the competing RA2's reputation. As we can see, the relationship between the reputation and strategy of RA1 remains 'u-shaped' as in the monopolistic case. Moreover, as the reputation of RA2 increases, the reputation at which RA1 has minimum x_1, *i.e.* is least likely to lie, also increases. This is not surprising as the disciplining effect is greatest when the reputation of the competing RA (RA2) is close to the reputation of RA1. This is because when the RAs' reputations are close, it is more likely that the market leadership will change, resulting in more disciplined behaviour. Conversely, if the two RAs have very different reputations, the disciplining effect is relatively weaker.

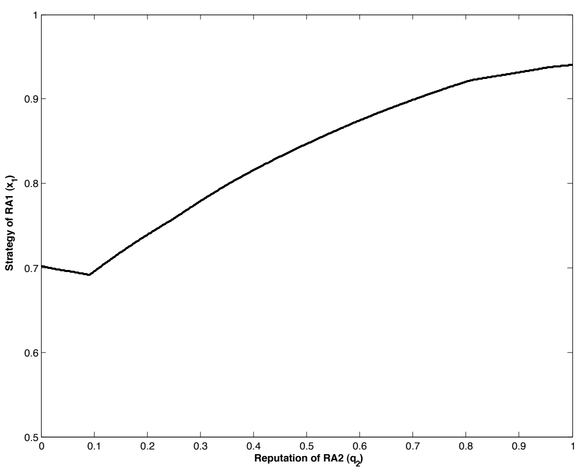


(a) q_2=0.25 (b) q_2=0.45

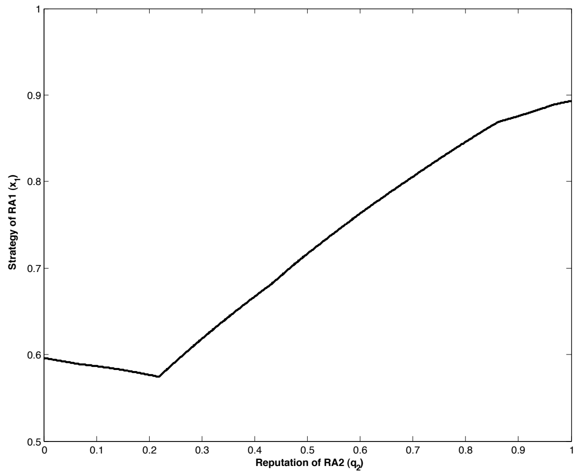
(c) q_2=0.55 (d) q_2=0.75

Figure 6: Strategy vs Reputation, $(\lambda,\, p_G,\, \delta) = (0.5,\, 0.7,\, 0.9)$, different values of q_2

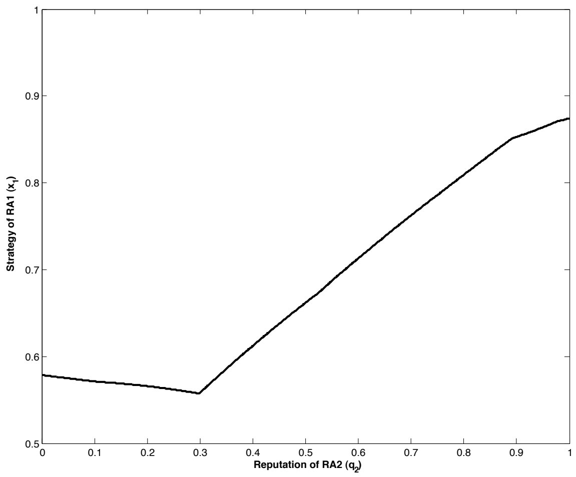
Moreover, as Figure 7 shows, the strategy of RA1 is initially decreasing with or flat in RA2's reputation, and then increasing. This effect of competition is a combination of the disciplining effect and the market-sharing effect. The disciplining effect is strongest



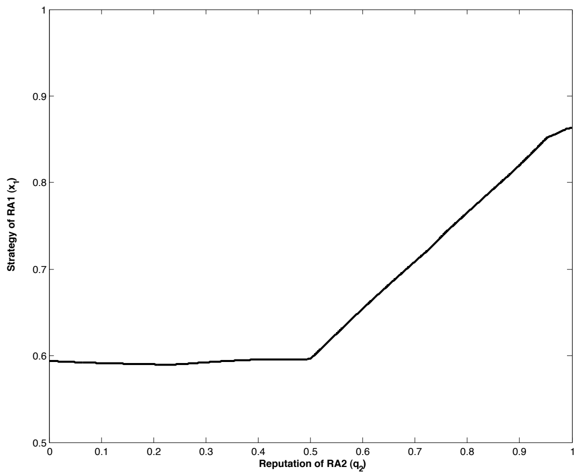
(a) q_1=0.25



(b) q_1=0.45



(c) q_1=0.55



(d) q_1=0.75

Figure 7: Strategy vs Reputation, different values of q_1, (λ, p_G, δ) = (0.5, 0.7, 0.9)

when the two RA's reputations are close, and weakest when the two RA's reputations are far apart, which implies that the probability of a change of market leader is very small. On the other hand, the market-sharing effect is always increasing in the competing RA's reputation. When the reputation of RA2 is low, the market-sharing effect is very small as RA2 can only take away a tiny fraction of market share. As RA2's reputation starts to increase, RA1 tends to lie less as the disciplining effect dominates the market-sharing effect. However, when RA2's reputation goes beyond a certain level, the market-sharing effect dominates as RA2's reputation becomes much higher than RA1's. Hence RA1 will lie more for high values of RA2's reputation, due to the dominance of the market-sharing effect.

Figures 8 and 9 show the expected profits of RA1 as a function of RA1 and RA2's reputation. We can clearly see that the expected profits of RA1 are increasing in its own reputation, and decreasing in its competitor's reputation, illustrating the market-sharing effect.

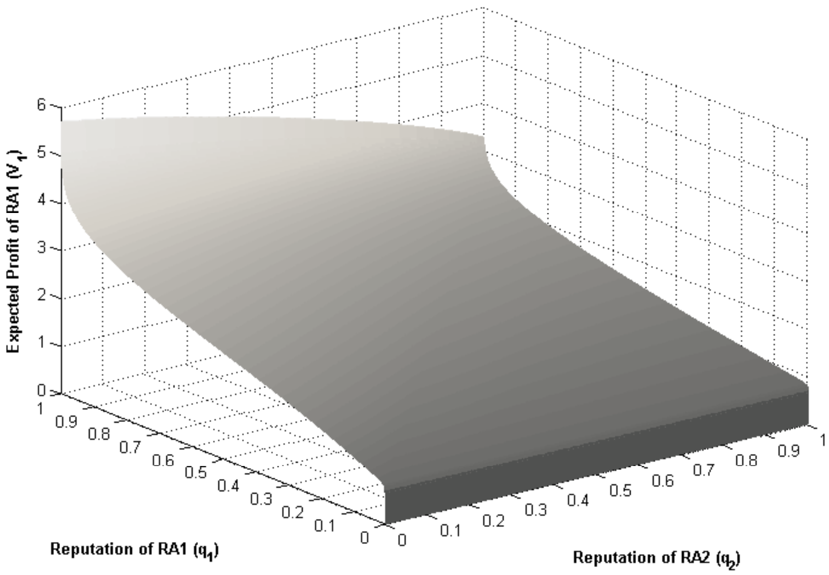


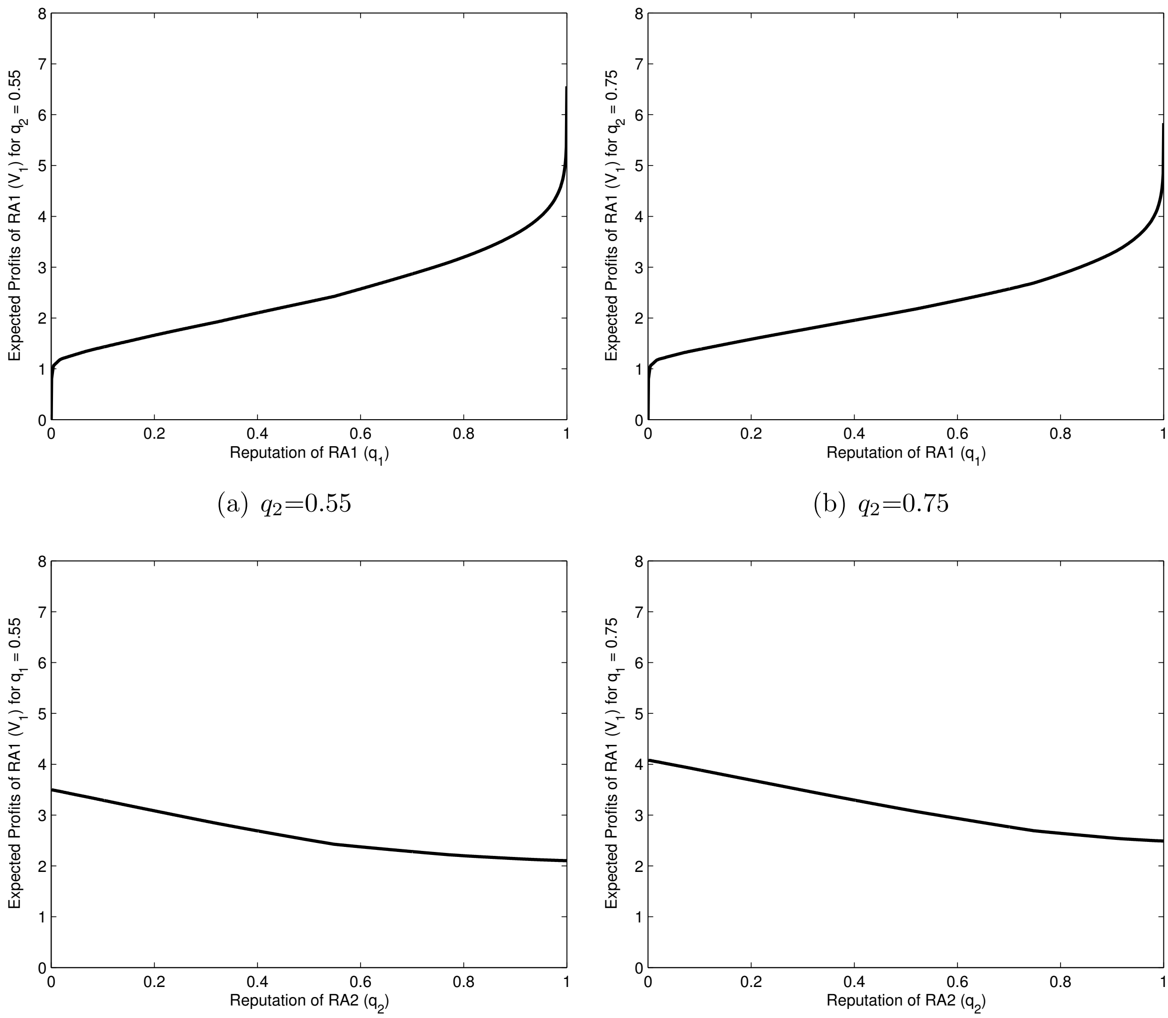
Figure 8: Expected Profits vs Reputation, $(\lambda,\, p_G,\, \delta) = (0.5,\, 0.7,\, 0.9)$

Finally, Figure 10 shows the convergence dynamics. It plots the change in RA1's strategy as the number of periods remaining increases. Reputation becomes less and less important as the number of periods remaining declines since there are fewer periods to reap the benefits of higher reputation. Thus ratings inflation increases. Note that as the number of periods remaining increases, the strategy converges, implying that we approach a long (infinite) horizon equilibrium.

In summary, our results show that introducing competition in the form of a second RA is not sufficient to discipline the RAs which always lie with positive probability in equilibrium. We now show that competition will actually increase the lax behaviour of RAs and reduce expected welfare.

5.3 Comparing Monopolistic and Competitive RA

It is often suggested that introducing more competition in the ratings industry can alleviate the problem of improper incentives and ratings inflation. However, our results show that competition is likely to worsen this situation and lead to more ratings inflation.



(a) q_2=0.55 (b) q_2=0.75

(c) q_1=0.55 (d) q_1=0.75

Figure 9: Expected Profits vs Reputation, different values of q_1 and q_2, $(\lambda,\ p_G,\ \delta) = (0.5,\ 0.7,\ 0.9)$

Figure 11 compares the strategic behaviour of RA1 under no competition, *i.e.* monopolistic RA ($q_2 = 0$), and under a competitive setting with different values of q_2. We observe that in most cases, RA1 is prone to greater ratings inflation relative to the monopolistic RA.

As described before, the implication of competition can be divided into the market-sharing effect and the disciplining effect. We can see that the market-sharing effect dominates the disciplining effect (*i.e.* competition aggravates lax behaviour) in most cases. The only case where competition may actually alleviate the lax behaviour of RA1 is when q_2 is very low (as shown in Figure 11(a)). This is because the market-sharing effect is weakest relative to the disciplining effect for low values of q_2. Intuitively, the disciplining effect only depends on the difference between q_1 and q_2, whereas the market-

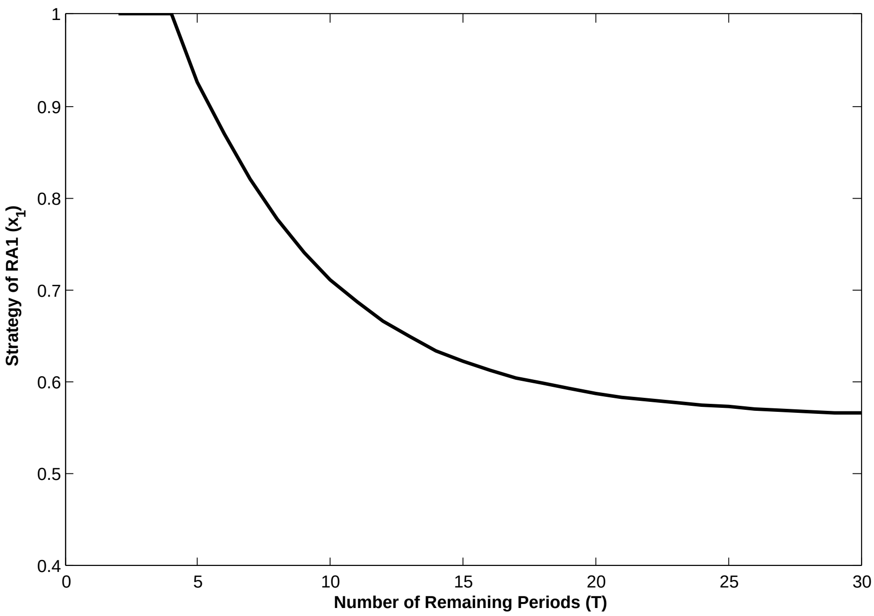


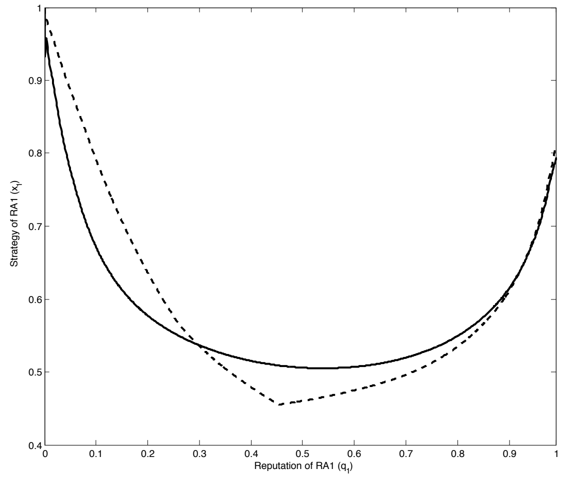
Figure 10: Convergence Dynamics of RA1

sharing effect increases with the absolute level of q_2. Hence the market-sharing effect tends to dominate the disciplining effect except for low values of q_2.

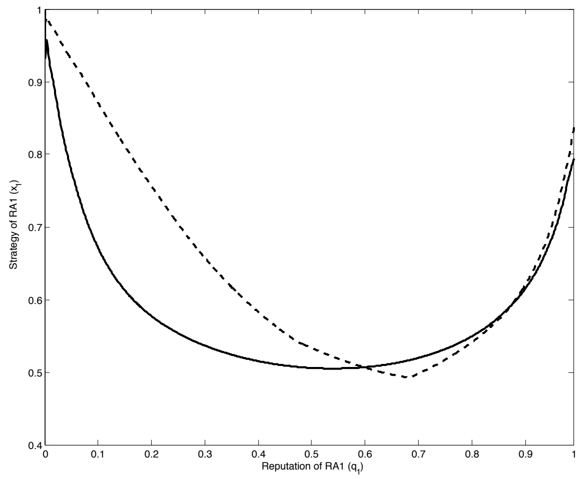
In order to assess the overall impact of competition, we compute the expected increase in lax behaviour of RA1 given its own reputation, assuming that the reputation of RA2 is uniformly distributed on $[0, 1]$. A positive value of this measure means the overall effect of enhanced competition on RA1 is to lie more (*i.e* inflate ratings more).

$$\text{Excess Lax Behaviour of RA1} = \int_{q_2 \in [0,1]} x_1(q_1, q_2)\, dq_2 \; - \; x_1(q_1, 0) \tag{10}$$

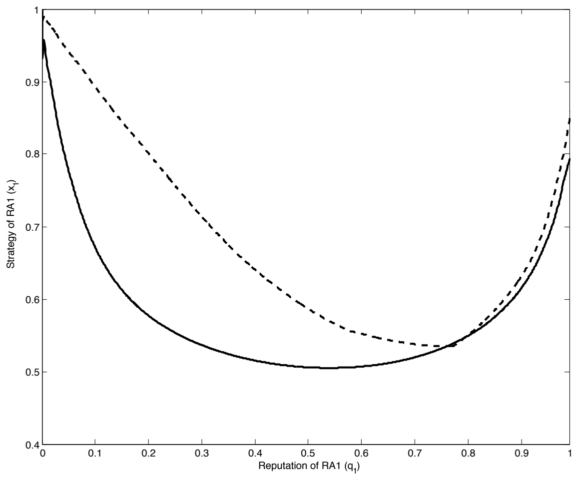
As shown in Figure 12, the expected increase in lax behaviour of RA1 is always positive, indicating that competition will, in general, aggravate ratings inflation. This is because a smaller market share will tend to reduce the reputational concerns of the RAs, and this market-sharing effect outweighs the disciplining effect brought by competition. Moreover, we can see that the expected increase in lax behaviour is increasing for low values of RA1's own reputation and decreasing for high values of RA1's reputation. The intuition is that, when the reputation of RA1 is low, the market share of RA1 is going to shrink significantly after introducing RA2 and the market-sharing effect of competition is strongest. However, when the reputation of RA1 is high, the impact of introducing RA2 on RA1's market share is small, hence the market-sharing effect becomes weaker



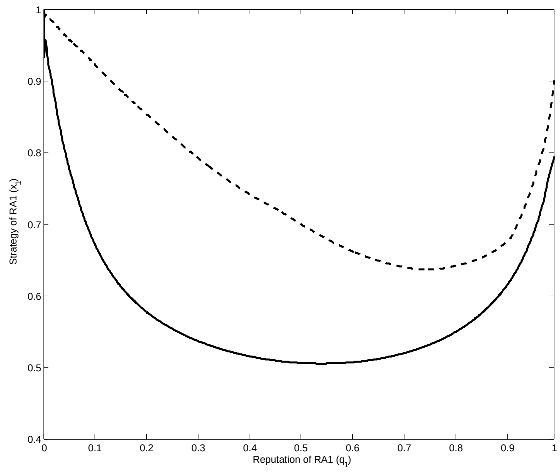
(a) q_2=0.25



(b) q_2=0.45



(c) q_2=0.55



(d) q_2=0.75

Figure 11: Comparing Monopolistic and Competitive RA, (λ, p_G, δ) = (0.5, 0.7, 0.9)
Solid line represents monopolistic RA while dashed line represents competitive RA with different values of q_2

and RA1 will lie relatively less. We verify that the excess lax behaviour, as defined above, is always positive for other values of λ and p_G in Appendix C.1.

In addition, we measure the expected total welfare in both monopolistic and duopolistic settings as defined below.

$$
\begin{aligned}
\text{Expected Total Welfare} &= E(\text{Project Payoff}) - E(\text{Financing Cost}) \\
&= P(\text{RA1 rates})\Big(\lambda p_G \Phi - E(X)\big(\lambda + (1-\lambda)(1-q_1)x_1\big)\Big) \\
&\quad + P(\text{RA2 rates})\Big(\lambda p_G \Phi - E(X)\big(\lambda + (1-\lambda)(1-q_2)x_2\big)\Big)
\end{aligned}
$$

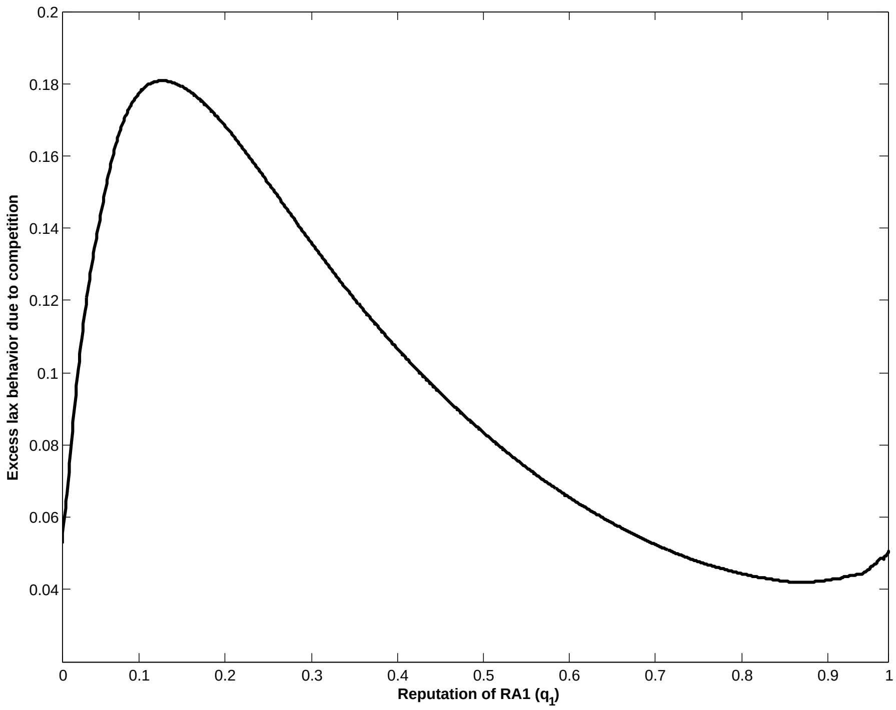


Figure 12: Excess Lax Behaviour of RA1 due to Competition, (λ, p_G, δ)=(0.5, 0.7, 0.9)

Figure 13 compares the total welfare[17] between the monopolistic case and the duopolistic case where both RAs have the same reputation. We can see that if a new RA is introduced with the same reputation as the incumbent RA, then the total welfare will always decrease, due to the fact that both RAs are more likely to inflate ratings.

Moreover, when we compare in Figure 14, the expected total welfare between the monopolistic case and the duopolistic case with fixed values of reputations of RA2, we can see that introducing competition will always lead to lower total welfare as long as the reputation of RA2 is lower than the reputation of RA1. However, total welfare may increase if the entrant RA has a higher reputation than the incumbent. Overall, this implies that competition is likely to adversely impact total welfare, unless we can introduce a new RA with much higher reputation than the incumbent. We check the robustness of this result for different values of λ and p_G in Appendix C.2.

[17]We are computing the welfare in one period only because it does not depend on time.

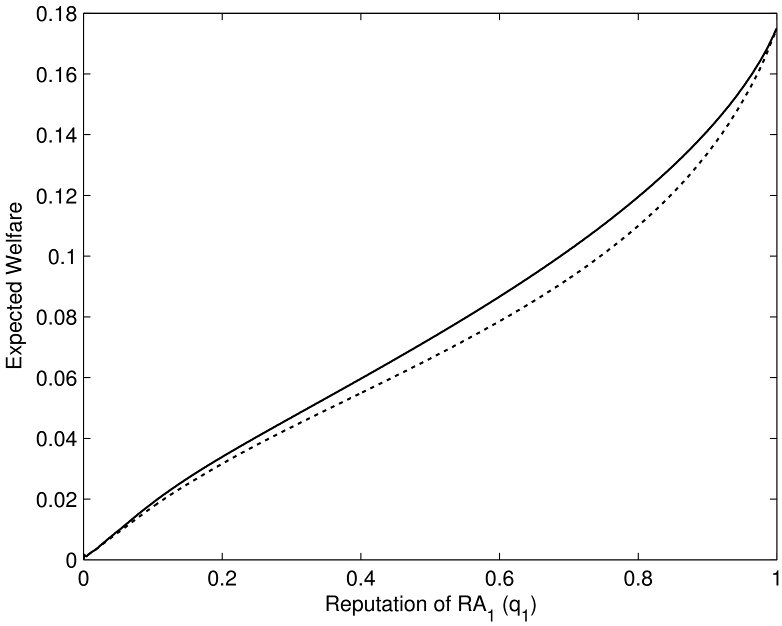
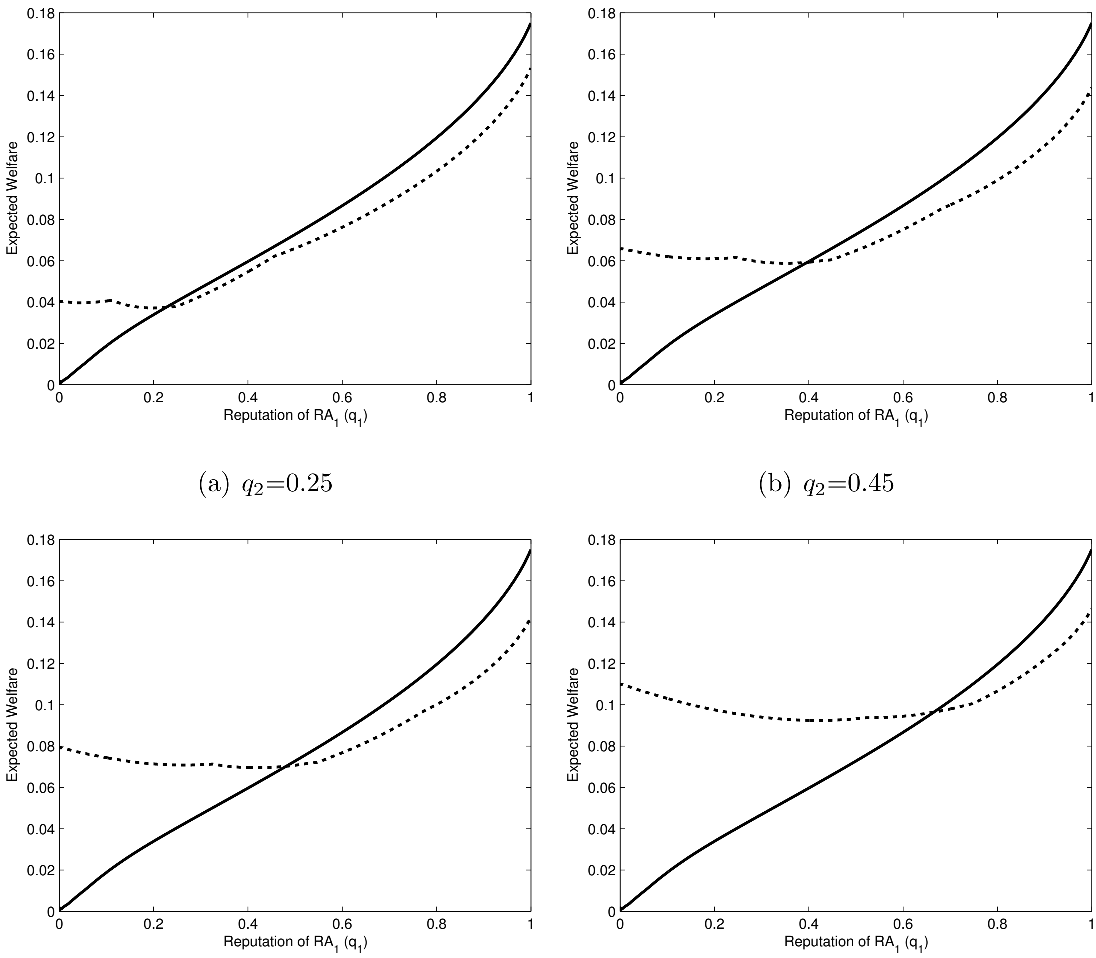


Figure 13: Expected Welfare - Competitor has same Reputation
Solid line represents monopoly while dashed line represents duopoly with $q_1 = q_2$



(a) q_2=0.25

(b) q_2=0.45

(c) q_2=0.55

(d) q_2=0.75

Figure 14: Expected Welfare - Competitor has different Reputation
Solid line represents monopoly while dashed line represents duopoly for different values of q_2
(λ, p_G, δ) = (0.5, 0.7, 0.9)

6 Conclusion

In this paper we show that competition can amplify ratings inflation and the lax behaviour of rating agencies, reducing total welfare. This result has important policy implications since it suggests that the most often cited solution to ratings inflation - enhanced competition in the ratings industry - is likely to render the situation worse. While we acknowledge that in order to focus on the implications of competition in the credit ratings industry, we have abstracted from other important issues such as herd behaviour, multiple ratings, and the quality of the models used by rating agencies, we believe that our results can serve as a baseline for evaluating the reform proposals currently being discussed.

One of the key thrusts of recent regulatory action in the credit ratings space has been to relax barriers to entry and enhance competition. In the US, the Securities and Exchange Commission has relaxed some barriers to entry and allowed several new CRAs in the US to obtain the Nationally Recognized Statistical Rating Organization (NRSRO) status. The European Union (EU) has gone further and has introduced new requirements as part of the proposed amendments to the EU Regulation on credit rating agencies, the so called 'CRA-III'. The new legislation seeks to place a cap on the market share of each ratings agency and requires issuers to rotate credit rating agencies periodically (see European Commission (2011) for details).

In the context of our model, the cap on the market share of rating agencies is likely to incentivise RAs to inflate ratings when their market share is close to the cap since they would no longer benefit from higher reputation. Furthermore, proposals to rotate RAs would mean that RAs would be assured of a market share, irrespective of their reputation. This would break the link between reputation and future income, thereby increasing ratings inflation. More broadly, proposals aimed at artificially enhancing competition are likely to exacerbate the market sharing effect, while doing little to increase the discipling effect.

One of the key findings in our model is that unless the new entrant RA has a higher reputation than the incumbent, increased competition is likely to adversely impact total welfare. However, it is unlikely that a new entrant would have sufficiently high reputation (and hence market share) to challenge the incumbents. It is more plausible to believe that the new entrants would start off as marginal players. Moreover, it is likely that under the current issuer pay model, they will continue to remain marginal players as their low reputation (and associated market share) would incentivise them to inflate ratings more than the established RAs. Interestingly, anecdotal evidence suggests that ratings issued by Dominion Bond Rating Service (DBRS), a relatively new player in the European market, are significantly more lenient than those issued by the more established players.

In conjunction with related work on multiple ratings and herd behaviour in the credit ratings industry, our results suggest that a fundamental reorganisation of the industry may be required to align the incentives. The conflict of interest highlighted in our paper is fundamental to the issuer-pay model and any meaningful attempt to resolve the conflict would require a fundamental shift in the way rating agencies are compensated. Empirical work by Xia and Strobl (2012) suggests that investor paid RAs can be a solution as they are unlikely to be affected by the conflict of interest highlighted in this paper and can have a discipling effect on the incumbent RAs. However, while an investor pay RAs can be a solution, free riding on the part of investors could result in insufficient revenues for such RAs, making it difficult for them to compete with the incumbents. Deb and Murphy (2009) argue that although free riding is a problem, the increasing use of ratings by institutions, coupled with the rise in the speed of information diffusion in the markets over the last few decades could, with proper regulatory encouragement, ensure that there are investors willing to subscribe to ratings issued by investor pay RAs.

References

Ashcraft, A., P. Goldsmith-Pinkham, and J. Vickery, 2010, "MBS ratings and the mortgage credit boom," *Federal Reserve Bank of New York Staff Reports*.

Bar-Isaac, H., 2003, "Imperfect competition and reputational commitment," *Economics Letters*, 89(2), 167–173.

Bar-Isaac, H., and J. Shapiro, 2011, "Credit Ratings Accuracy and Analyst Incentives," *American Economic Review P& P*, 101(3), 120–24.

——— , 2012, "Ratings Quality over the Business Cycle," *Working Paper*.

Barr, A., 2008, "S&P proposes changes to ratings process," `http://www.marketwatch.com/news/story/sp-plans-changes-ratings-cuomo/story.aspx?guid={3A7B5846-1E5E-48D0-ACCB-96A22E992ADF}`.

Beck, T., 2008, "Reforming the Credit Rating Agency Industry," `http://crisistalk.worldbank.org/2008/12/reforming-the-c.html`.

Becker, B., and T. Milbourn, 2011, "How did increased competition affect credit ratings?," *Journal of Financial Economics*, 101(3), 493–514.

Bolton, P., X. Freixas, and J. Shapiro, 2012, "The credit ratings game," *The Journal of Finance*, 67(1), 85–112.

Bouvard, M., and R. Levy, 2009, "Humouring both parties: a model of two-sided reputation," working paper.

Bulletin, E. M., May 2009, "Credit rating agencies: developments and policy issues," pp. 107 – 117.

Damiano, E., L. Hao, and W. Suen, 2008, "Credible Ratings," *Theoretical Economics*, (3), 325–365.

Deb, P., and G. Murphy, 2009, "Rating Agencies - An Alternative Model," Bank of England.

European Commission, 2011, "Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL amending Regulation (EC) No 1060/2009 on credit rating agencies," .

Fitch Ratings, 2007, "Inside the Ratings: What Credit Ratings Mean," .

Gellert, J. H., 2009, *United States House of Representatives, Committee on Financial Services.*

Griffin, J. M., and D. Y. Tang, 2011, "Did Credit Rating Agencies Make Unbiased Assumptions on CDOs?," *American Economic Review*, 101(3), 125–30.

Horner, J., 2002, "Reputation and Competition," *The American Economic Review*, 92(3).

Jenkinson, N., 2008, "Ratings in structured finance: what went wrong and what can be done to address shortcomings?," *BIS-CGFS Papers*, 32.

Klein, B., and K. B. Leffler, 1981, "The Role of Market Forces in Assuring Contractual Performance," *The Journal of Political Economy*, 89(4), 615–641.

Lowenstein, R., 2008, "Triple-A Failure," `http://www.nytimes.com/2008/04/27/magazine/27Credit-t.html`.

Mariano, B., 2012, "Market power and reputational concerns in the ratings industry," *Journal of Banking & Finance.*

Mathis, J., J. McAndrews, and J.-C. Rochet, 2009, "Rating the raters: Are reputation concerns powerful enough to discipline rating agencies?," *Journal of Monetary Economics*, 56(5), 657–674.

Sangiorgi, F., J. Sokobin, and S. Chester, 2009, "Credit-Rating Shopping, Selection and the Equilibrium Structure of Ratings," *Working Paper.*

Sangiorgi, F., and C. Spatt, 2011, "Opacity, Credit Rating Shopping and Bias," *Working Paper.*

SEC, 2008, "Summary Report of Issues Identified in the Commission Staff? Examinations of Select Credit Rating Agencies By the Staff of the Securities and Exchange Commission, July 8, 2008," .

Skreta, V., and L. Veldkamp, 2009, “Ratings Shopping and Asset Complexity,” *Journal of Monetary Economics*, 56(5), 678–695.

US Department of the Treasury - Financial Regulatory Reform: A New Foundation, 2009, `http://www.financialstability.gov/roadtostability/regulatoryreform.html`.

Watson, R., 2008a, “ESF/SIFMA response to CESR consultation paper on the role of credit rating agencies in structured finance,” .

——— , 2008b, “SIFMA/ESF response to IOSCO technical committee consultation report on the role of credit rating agencies in structured finance markets,” .

White, L. J., 2010, “Markets: The Credit Rating Agencies,” *Journal of Economic Perspectives*, (2), 211–26.

Xia, H., 2012, “Can Competition Improve the Information Quality of Credit Ratings?,” *Working Paper*.

Xia, H., and G. Strobl, 2012, “The Issuer-Pays Rating Model and Ratings Inflation: Evidence from Corporate Credit Ratings,” *Working Paper*.

Zelmer, M., 2007, “Reforming the Credit-Rating Process,” *Bank of Canada Financial System Review*, December, pp 51–57.

A Proofs

A.1 *Proof of Proposition 1*

There exists a unique x_1, where $0 \leq x_1 \leq 1$, given that $V_t(q_1, q_2)$ is an increasing function in q_1.

Proof. When the strategic RA (RA1) gets a bad project, it will get pay-off $\Psi(lie) = I + \delta V_t(q_1^F, q_2)$ if it gives the project a GR, and $\Psi(honest) = \delta V_t(q_1^N, q_2)$ if it refuses rating. Note that $q_1^F = \frac{\lambda(1-p_G)q_t}{\lambda(1-p_G)+(1-\lambda)(1-q_1)x_1}$ and $q_1^N = \frac{q_t}{1-x_1(1-q_t)}$, *i.e.* q_1^F is decreasing in x_1 and q_1^N is increasing in x_1. Given that $V_t(q_1, q_2)$ is increasing in q_1, it is easy to see that $\Psi(lie)$ is decreasing in x_1 and that $\Psi(honest)$ is increasing in x_1. Thus if we define x_1 such that

- $x_1 = 1$ if $\Psi(lie) \geq \Psi(honest)$
- $x_1 = 0$ if $\Psi(lie) \leq \Psi(honest)$ for
- $x_1 = x_1^*$ such that $0 < x_1^* < 1$ if $\Psi(lie) = \Psi(honest)$

it follows that x_1 is well-defined and unique. □

A.2 *Proof of Proposition 2*

The strategic RA does not have incentives to give NR to a good project.

Proof. Suppose that the strategic RA (RA1) gets a good project and that its strategy is x_1. Let's examine whether RA1 wants to deviate:

- if $x_1 = 1$, we have $\Psi(lie) \geq \Psi(honest)$, or $I + \delta V_t(q_1^F, q_2) \geq \delta V_t(q_1^N, q_2)$. If the RA1 gives NR to the good project, it will get $\delta V_t(q_1^N, q_2)$, and $I + p_G \delta V_t(q_1, q_2) + (1 - p_G)\delta V_t(q_1^F, q_2)$ otherwise. Since $I + p_G \delta V_t(q_1, q_2) + (1 - p_G)\delta V_t(q_1^F, q_2) \geq I + \delta V_t(q_1^F, q_2) \geq \delta V_t(q_1^N, q_2)$, RA1 does not want to deviate.

- if $x_1 = 0$, $q_1^N = q_1^F = q_1$, hence reputation becomes irrelevant and the RA does not have an incentive to give NR to the good project.

- if $0 < x_1 < 1$, we have $\Psi(lie) = \Psi(honest)$, so $I + p_G \delta V_t(q_1, q_2) + (1 - p_G)\delta V_t(q_1^F, q_2) \geq I + \delta V_t(q_1^F, q_2) = \delta V_t(q_1^N, q_2)$, and hence RA1 does not want to deviate.

Therefore RA1 does not have incentives to give NR to a good project. □

A.3 *Proof of Corollary 1*

Assume $p_G < 1$. Then the equilibrium strategy of the strategic RA is always positive.

Proof. Suppose that the equilibrium strategy is $x_1 = 0$. Then $q_1^N = q_1^F = q_1$ and we must have $I + \delta V_t(q_1, q_2) \leq \delta V_t(q_1, q_2)$. This is impossible as long as $I > 0$. Hence $x_1 = 0$ cannot be an equilibrium strategy. □

A.4 *Proof of Corollary 2*

Suppose the model ends in period T. Then the equilibrium strategy of the strategic RA is $x_t = 1$ at $t = T - 1, T$.

Proof. At $t = T$, the strategic RA does not have any reputational concerns. This implies that the strategy of strategic RA will be to always give GR if the project is bad, *i.e.*

$x_T = 1$.

Similarly, at $t = T - 1$ the strategic RA will always lie. Suppose that a bad project comes to strategic RA, say RA1. The expected pay-off of RA1 is

$$I + \delta V_{T-1}(q_1^F, q_2) = I + f(q_1^F, 1, q_2, 1)\delta I \tag{11}$$

if it lies, *i.e.* gives a good rating, and

$$\delta V_{T-1}(q_1^N, q_2) = f(q_1^N, 1, q_2, 1)\delta I \tag{12}$$

if it does not lie, *i.e.* gives no rating, where $f(q_1, x_1, q_2, x_2)$ is the probability that the project comes to RA1 in the next period. Using equations (5), (6) and (7) we have

- $f(q_1, x_1, q_2, x_2) = \frac{\frac{1}{2}\left(s(q_1,x_1)-\lambda p_G\right)}{p_G(1-\lambda)}$ if $s(q_1, x_1) \leq s(q_2, x_2)$
- $f(q_1, x_1, q_2, x_2) = \frac{s(q_1,x_1)-\frac{1}{2}\left(s(q_2,x_2)+\lambda p_G\right)}{p_G(1-\lambda)}$ otherwise

where $s(q, x) = \frac{\lambda p_G}{\lambda+(1-q)(1-\lambda)x}$.

Although in this case RA1 does have reputational concerns, these are not sufficient to prevent RA1 from being lax and not giving GR to bad projects. Since by being honest RA1 is giving up I today, in exchange for having a higher chance of getting I in the next period, it is not optimal for RA1 to be honest, given that RA1 is impatient (*i.e.* $\delta < 1$). Hence the optimal strategy of RA1 is to always lie, i.e. $x_{T-1} = 1$.

□

B Finite Horizon Solution

We assume the model lasts for three periods, $t = 1, 2, 3$, and the RAs maximise their expected total income over the three periods. As before, we compute the equilibrium strategy of the RAs using backward induction. We already know that the strategic RA will always lie in the last two periods, as shown in Corollary 2.

We solve for the equilibrium strategy at $t = 1$. Again, let's look at the decision of RA1. Since RA1 will always lie at $t = 2, 3$, the expected pay-off of RA1 at $t = 1$ is

$$\begin{aligned}
\Psi(lie) &= I + \delta V_2(q_1^F, q_2) = I + \delta f(q_1^F, 1, q_2, 1)I \\
&+ \delta^2 \Big\{ f(q_1^F, 1, q_2, 1)[\lambda p_G f(q_1^F, 1, q_2, 1) + ((1 - p_G)\lambda + (1 - \lambda))f(q_1^{FF}, 1, q_2, 1)] \\
&+ f(q_2, 1, q_1^F, 1)[\lambda p_G f(q_1^F, 1, q_2, 1) + \big(\lambda(1 - p_G) + (1 - \lambda)(1 - q_2)\big) f(q_1^F, 1, q_2^F, 1) \\
&+ (1 - \lambda)q_2 f(q_1^F, 1, q_2^N, 1)] \Big\} I \qquad (13)
\end{aligned}$$

if it lies, and

$$\begin{aligned}
\Psi(honest) &= \delta V_2(q_1^N, q_2) = \delta f(q_1^N, 1, q_2, 1)I \\
&+ \delta^2 \Big\{ f(q_1^N, 1, q_2, 1)[\lambda p_G f(q_1^N, 1, q_2, 1) + ((1 - p_G)\lambda + (1 - \lambda))f(q_1^{NF}, 1, q_2, 1)] \\
&+ f(q_2, 1, q_1^N, 1)[\lambda p_G f(q_1^N, 1, q_2, 1) + \big(\lambda(1 - p_G) + (1 - \lambda)(1 - q_2)\big) f(q_1^N, 1, q_2^F, 1) \\
&+ (1 - \lambda)q_2 f(q_1^N, 1, q_2^N, 1)] \Big\} I \qquad (14)
\end{aligned}$$

if it is honest, where $f(q_1, x_1, q_2, x_2)$ is the probability that the project comes to RA1 next period, given its reputation q_1, its strategy x_1, its competitor's reputation q_2 and its competitor's strategy x_2.

As described in Section 4, we look for an equilibrium of the game by examining the trade-off facing RA1, *i.e.* the difference between expressions (13) and (14). If the pay-off from lying is greater then $x_1 = 1$ and we have a pure-strategy equilibrium in which RA1 always lies; if the pay-off from not lying is greater then $x_1 = 0$ and we have a pure-strategy equilibrium in which RA1 never lies; otherwise we have a mixed-strategy equilibrium in which RA1 is indifferent between lying and not lying, given some prior beliefs about its strategy, *i.e.* $0 < x_1 < 1$.

To derive an analytical solution to this game, we make a simplifying assumption that $p_G = 1$ and $\delta = 1$. This assumption implies that the reputation of the strategic RA goes to zero if it gives a GR to a bad project since now every good project succeeds and every bad project fails. This simplifies expressions (13) and (14) and allows us to derive the equilibrium strategy of RA1. This assumption is relaxed in Section 5.

The expression of market share of RA1 depends on whether RA1 has a higher probability of success than its competitor. Given that the strategy of the strategic RA in the last two periods is to always lie, the RA with a higher reputation will have a higher market share in any single period. Hence we compute the strategy of RA1 in different ranges of the reputation of RA2.

Proposition 4. *The equilibrium strategy at $t = 1$ assuming $p_G = 1$ and $\delta = 1$ is*

$$x_1 = \begin{cases} 0 & \text{if } A \leq \frac{\lambda q_1}{2\left(\lambda q_1 + (1-q_1)\right)} \\ 1 - \frac{(1-2A)\lambda q_1}{2A(1-q_1)} & \text{if } \frac{\lambda q_1}{2\left(\lambda q_1 + (1-q_1)\right)} < A < \frac{1}{2} \\ 1 & \text{if } A \geq \frac{1}{2} \end{cases}$$

where A is the solution to the equation

$$\Psi(lie) - \Psi(honest) = I - \delta(2A - \min\{A,B\})I - \delta^2\Big(\lambda(2A - \min\{A,B\})^2 +$$
$$(2B - \min\{A,B\})\big[\lambda(2A - \min\{A,B\}) + 2(1-\lambda)(1-q_2)A + (1-\lambda)q_2 A\big]\Big)I = 0$$

and $B = \frac{\frac{1}{2}\left(s(q_2,1) - \lambda p_G\right)}{p_G(1-\lambda)}$.

Furthermore, in equilibrium, x_1 is decreasing in q_1. Moreover, x_1 is increasing in q_2 using first order Taylor approximation.

Proof. Since $p_G = 1$, the reputation of RA1 (*i.e.* the strategic RA) will go to zero if it gives a GR to a bad project since now every good project succeeds and every bad project fails. So the expected pay-off from giving a GR to a bad project is I. This simplifies expressions (13) and (14) and allows us to derive RA1's equilibrium strategy.

The expected pay-off from being honest is

$$\Psi(honest) = \delta f(q_1^N,1,q_2,1)I + \delta^2\Big(f(q_1^N,1,q_2,1)\lambda f(q_1^N,1,q_2,1)$$
$$+ f(q_2,1,q_1^N,1)[\lambda f(q_1^N,1,q_2,1) + (1-\lambda)(1-q_2)f(q_1^N,1,q_2^F,1) + (1-\lambda)q_2 f(q_1^N,1,q_2^N,1)]\Big)I$$

Using equations (6) and (7) and noting that RA1 will always lie in periods $t = 2,3$, this can be rewritten as

$$\Psi(honest) = \delta(2A - \min\{A,B\})I + \delta^2\Big(\lambda(2A - \min\{A,B\})^2$$
$$+ (2B - \min\{A,B\})[\lambda(2A - \min\{A,B\}) + 2(1-\lambda)(1-q_2)A + (1-\lambda)q_2 A]\Big)I$$

where $A = \frac{\frac{1}{2}\left(s(q_1^N,1)-\lambda p_G\right)}{p_G(1-\lambda)}$ and $B = \frac{\frac{1}{2}\left(s(q_2,1)-\lambda p_G\right)}{p_G(1-\lambda)}$

The expected pay-off from lying is I, since the RA's reputation goes to zero

$$\Psi(lie) = I$$

We look for a equilibrium of the game by examining RA1's trade-off between lying and not lying. If the pay-off from lying is greater when $x_1 = 1$, we have a pure-strategy equilibrium in which RA1 always lies; if the pay-off from not lying is greater when $x_1 = 0$, we have a pure-strategy equilibrium in which RA1 never lies; otherwise we have a mixed-strategy equilibrium in which RA1 is indifferent between lying or not given some prior beliefs about its strategy, *i.e.* $0 < x_1 < 1$.

We now solve the equation $\Psi(lie) - \Psi(honest) = 0$. We do this in 2 stages. In the first stage, we solve the equation in terms of A and then using the expression for A, we solve for the equilibrium value of x_1.

For $A < B$ we have

$$\Psi(lie) - \Psi(honest) = \delta^2(1-\lambda)(2-q_2)A^2 - \big(\delta + 2B\delta^2\lambda + 2B\delta^2(1-\lambda)(2-q_2)\big)A + 1$$

Assuming $\delta = 1$, the solution is

$$A = B + \frac{1 + 2B\lambda - \sqrt{(1+2B\lambda)^2 + (1-\lambda)(2-q_2)\varrho}}{2(1-\lambda)(2-q_2)}$$

which is valid[18] as long as $\varrho = B^2\big(2-(1-\lambda)q_2\big) + B - 1 > 0$.

[18] *i.e.* A is real and less than B.

Note that B can be simplified to

$$B = \frac{\lambda q_2}{2\big(1 - q_2(1-\lambda)\big)}$$

We can see that B is bounded above by $\frac{1}{2}$. Therefore $\varrho \leq 0$ and we can rule out the case above.

Now for $A \geq B$ we have

$$\Psi(lie) - \Psi(honest) = -4\delta^2\lambda A^2 - \big(2\delta - 2B\delta^2\lambda + B\delta^2(1-\lambda)(2-q_2)\big)A + \delta B + 1$$

Assuming $\delta = 1$, the solution is

$$A = B + \frac{\sqrt{\big(2 + 6\lambda B + B(1-\lambda)(2-q_2)\big)^2 - 16\lambda\varrho} - \big(2 + 6\lambda B + B(1-\lambda)(2-q_2)\big)}{8\lambda}$$

which is valid[19] given $\varrho = B^2\big(2 - (1-\lambda)q_2\big) + B - 1 \leq 0$.

Note that A can also be expressed as

$$A = \frac{\sqrt{\big(2 - 2\lambda B + B(1-\lambda)(2-q_2)\big)^2 + 16\lambda(B+1)} - \big(2 - 2\lambda B + B(1-\lambda)(2-q_2)\big)}{8\lambda}$$

Applying first-order Taylor approximation[20], we have

$$A \simeq \frac{B+1}{2 - 2\lambda B + B(1-\lambda)(2-q_2)} = \frac{B+1}{2 - 2\lambda B + \frac{\lambda q_2}{2} + B(1-2\lambda)}$$

[19] *i.e.* A is real and greater than B.
[20] That is, $\sqrt{N^2 + d} \simeq N + \frac{d}{2N}$.

Substituting for B, the first order derivative of A with respect to q_2 is

$$\frac{\lambda\Big((1-\lambda)(3\lambda-2)q_2^2+4(1-\lambda)q_2+8\lambda\Big)}{\Big(4-q_2(4\lambda^2-6\lambda+4)-\lambda(1-\lambda)q_2^2\Big)^2}$$

It can be shown that the minimum of the above is proportional to $-\frac{4(1-\lambda)}{3\lambda-2}+8\lambda$ and is attained when $q_2=-\frac{2}{3\lambda-2}$. When $\lambda<\frac{2}{3}$, the minimum is always positive, hence A is increasing in q_2. When $\lambda>\frac{2}{3}$, the derivative reaches zero when $q_2=\frac{-4(1-\lambda)\pm\sqrt{\big(16(1-\lambda)^2-32\lambda(1-\lambda)(3\lambda-2)}}{2(1-\lambda)(3\lambda-2)}$, which is negative. Hence the minimum is positive for $q_2>0$. Therefore A is always increasing in q_2.

Now we have shown that there always exists a solution which depends on the parameter ϱ. Since A always has a solution, we can use it to find the equilibrium strategy x_1 in terms of A, *i.e.* we will look for the value of x_1 such that $\frac{\frac{1}{2}\big(s(q_1^N,1)-\lambda p_G\big)}{p_G(1-\lambda)}=A$.

Note that assuming $p_G=1$ implies $\lambda p_G=\lambda$. Using this and equation (5), the above expression can be rewritten as $\frac{\lambda q_1^N}{\lambda q_1^N+1-q_1^N}=2A$, where $q_1^N=\frac{q_1}{1-(1-q_1)x_1}$.

Solving, we obtain

$$x_1=1-\frac{(1-2A)\lambda q_1}{2A(1-q_1)}$$

for $0<x_1<1$. This holds when $\frac{\lambda q_1}{2\big(\lambda q_1+(1-q_1)\big)}<A<\frac{1}{2}$. Clearly, we have x_1 increasing in A and decreasing in q_1. □

Proposition 4 implies that the strategy of RA1 depends on its own and its competitor's reputation. When A is large, RA1 always gives a GR to a bad project. Conversely, when A is small RA1 behaves honestly and gives NR to bad projects. In the intermediate range, RA1 has a mixed strategy, with $0<x_1<1$. Note that the lower threshold for A is increasing with RA1's reputation.

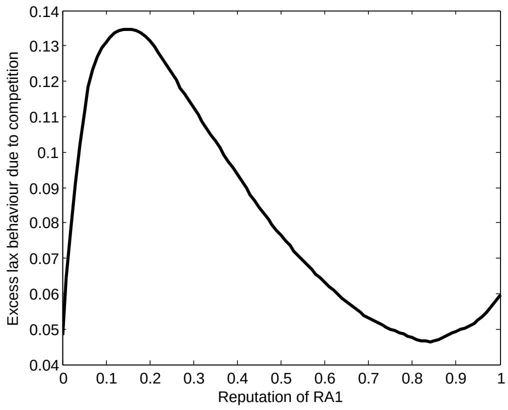
The results imply that RA1 tends to lie less as its reputation increases. The intuition behind this result is straightforward. Since we assumed $p_G = 1$, the reputation of RA1 goes to zero immediately after a project fails. This means that the cost of lying increases with RA1's reputation while the benefit of lying stays constant. Hence it is not surprising that RA1 prefers to lie less as its reputation increases.[21]

Moreover, RA1's strategy tends to increase with RA2's reputation. As explained before, competition has two opposite effects on the behaviour of RA1: the disciplining effect and the market-sharing effect. When the reputation of its opponent increases, RA1 will find it less attractive to increase its own reputation given a smaller expected future market share, and hence will behave more laxly. On the other hand, RA1 may have incentives to behave honestly when RA2's reputation increases in order to maintain its market leader position. Our analysis shows that the market-sharing effect tends to dominate the disciplining effect, using first order Taylor approximation. One potential explanation could be that, in our model, the market share of a rating agency is determined not only by its reputation relative to that of its competitor, but also by the absolute level of its reputation. That is, even a monopolistic RA cannot behave totally laxly, because otherwise its reputation would become too low to credibly rate most projects. Therefore, the incentives of a RA to maintain good reputation, even in absence of competition, render the disciplining effect of competition weaker. We believe this is reasonable because in reality, given rational investors, a monopolistic RA would not have unbounded market powers.

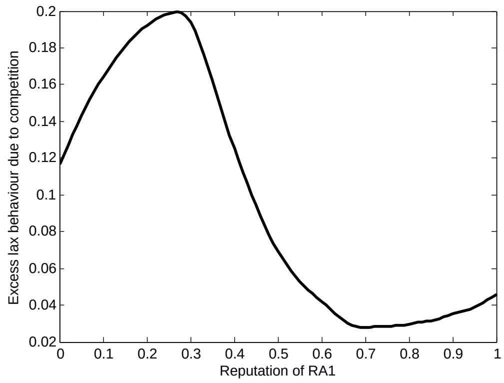
[21] Our results in section 5 show that this is no longer true if $p_G < 1$. The penalty on reputation will be smaller as the reputation of RA increases, *i.e.* the cost of ratings inflation can decrease with reputation, resulting in a 'u-shaped' relationship between strategy and reputation.

C Robustness Check

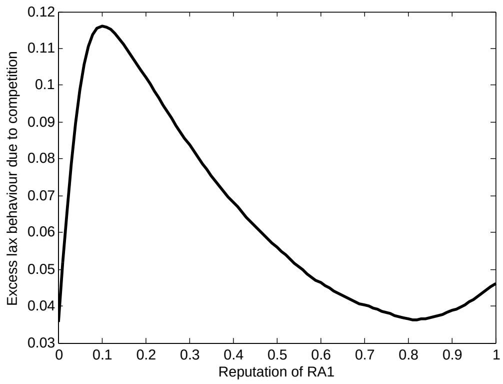
C.1 Excess Lax Behaviour



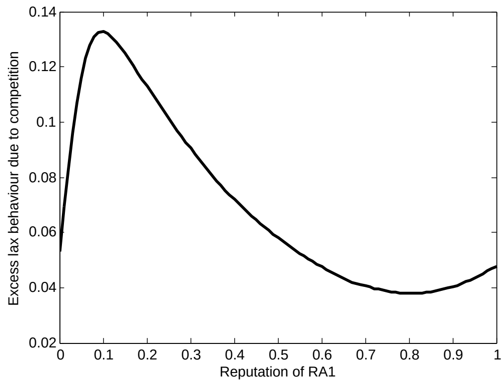
(a) λ=0.5, p_G=0.5



(b) λ=0.5, p_G=0.9

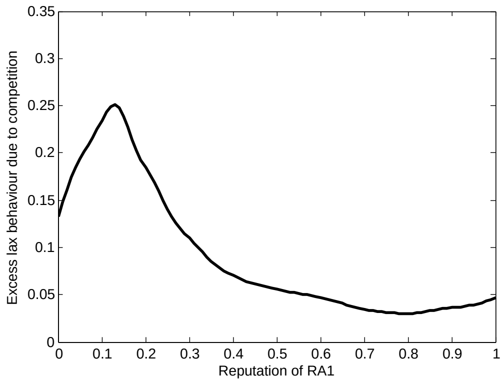


(c) λ=0.7, p_G=0.5

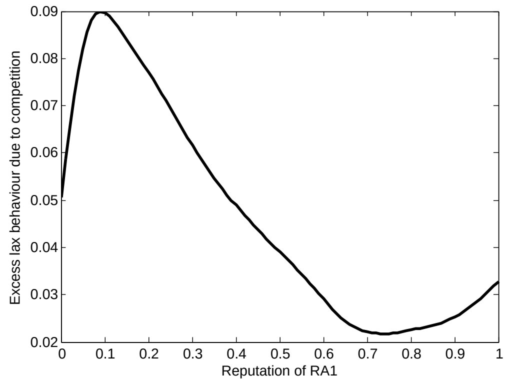


(d) λ=0.7, p_G=0.7

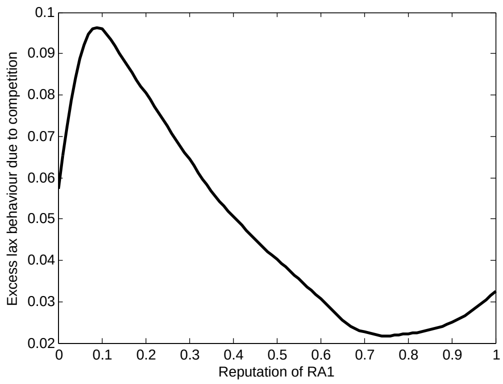
Figure 15: Excess Lax Behaviour for different values of λ and p_G ($\delta = 0.9$)



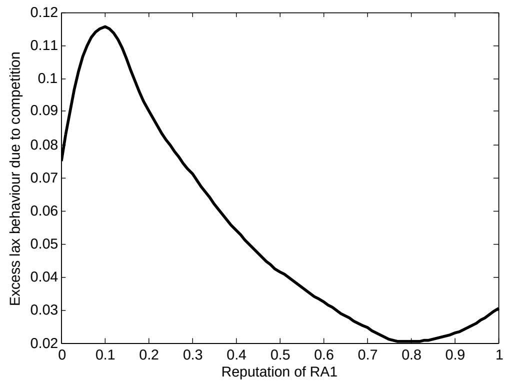
(a) λ=0.7, p_G=0.9



(b) λ=0.9, p_G=0.5



(c) λ=0.9, p_G=0.7



(d) λ=0.9, p_G=0.9

Figure 16: Excess Lax Behaviour for different values of λ and p_G (continued)

C.2 Expected Total Welfare

The reputation of RA1 (q_1) above which the expected total welfare is always greater in the monopoly case than in the duopoly case, for different values of q_2 ($\delta = 0.9$)

Parameter values	$\mathbf{q_2 = 0.25}$	$\mathbf{q_2 = 0.45}$	$\mathbf{q_2 = 0.55}$	$\mathbf{q_2 = 0.75}$
$\lambda = \mathbf{0.5}$, $\mathbf{p_G = 0.5}$	$q_1 = 0.23$	$q_1 = 0.45$	$q_1 = 0.52$	$q_1 = 0.69$
$\lambda = \mathbf{0.5}$, $\mathbf{p_G = 0.7}$	$q_1 = 0.23$	$q_1 = 0.45$	$q_1 = 0.52$	$q_1 = 0.69$
$\lambda = \mathbf{0.5}$, $\mathbf{p_G = 0.9}$	$q_1 = 0.23$	$q_1 = 0.45$	$q_1 = 0.52$	$q_1 = 0.69$
$\lambda = \mathbf{0.7}$, $\mathbf{p_G = 0.5}$	$q_1 = 0.15$	$q_1 = 0.45$	$q_1 = 0.51$	$q_1 = 0.67$
$\lambda = \mathbf{0.7}$, $\mathbf{p_G = 0.7}$	$q_1 = 0.15$	$q_1 = 0.45$	$q_1 = 0.51$	$q_1 = 0.67$
$\lambda = \mathbf{0.7}$, $\mathbf{p_G = 0.9}$	$q_1 = 0.15$	$q_1 = 0.45$	$q_1 = 0.51$	$q_1 = 0.67$
$\lambda = \mathbf{0.9}$, $\mathbf{p_G = 0.5}$	$q_1 = 0.13$	$q_1 = 0.45$	$q_1 = 0.51$	$q_1 = 0.66$
$\lambda = \mathbf{0.9}$, $\mathbf{p_G = 0.7}$	$q_1 = 0.13$	$q_1 = 0.45$	$q_1 = 0.51$	$q_1 = 0.66$
$\lambda = \mathbf{0.9}$, $\mathbf{p_G = 0.9}$	$q_1 = 0.13$	$q_1 = 0.45$	$q_1 = 0.51$	$q_1 = 0.66$

Market Frictions, Interbank Linkages and Excessive Interconnections

Pragyan Deb

p.deb@lse.ac.uk

Financial Markets Group

London School of Economics and Political Science

Abstract

This paper studies banks' decision to form financial interconnections using a model of financial contagion that explicitly takes into account the crisis state of the world. This allows us to model the network formation decision as optimising behaviour of competitive banks, where they balance the benefits of forming interbank linkages against the cost of contagion. We use this framework to study various market frictions that can result in excessive interconnectedness that was seen during the crisis. In this paper, we focus on two channels that arises from regulatory intervention - deposit insurance and the too big to fail problem.

Keywords: Contagion, network formation, financial crises, deposit insurance, too-big-to-fail.

JEL Classifications: C70, G21, D85, G01, G28.

1 Introduction

The financial crisis that hit the world economy in 2007 has made clear the interconnected nature of the global financial system. What started as difficulties in the US sub-prime mortgage market rapidly escalated and spilled over to markets all over the world. Banks, uncertain of the interconnections and fearful of contagion became less willing to lend money to each other. Interbank lending rates started to rise and soon the market for short-term lending dried-up. The credit crunch ultimately triggered a bank run at the British mortgage lender Northern Rock - something not seen in the UK for over 140 years and in Western Europe for the last 15 years.

The mechanics of financial contagion are well studied and the literature broadly takes two approaches - examining direct balance sheet linkages and indirect linkages operating through fire sales and expectations (see Allen, Babus, and Carletti (2009) for a survey of the literature on financial crises). The benefits of financial interlinkages and networks are also well known and stem primarily from better liquidity sharing and diversification. However, the literature on networks and contagion has mostly focussed on comparing the social costs and benefits of different network structures as opposed to the banks decision to form the network.

This paper attempts to fill this gap and develops a simple model to study banks' decision to form financial interconnections. This framework allows us to study various market frictions that can influence this decision and result in a wedge between socially optimal network formation and excessive interconnectedness that was seen during the crisis. While such suboptimal interconnections can arise from the several sources such as the underestimation of the crisis probability and disaster myopia, the focus of this paper is on two sources of market frictions that arises from regulatory intervention - deposit insurance and the existence of systemically important financial institutions (SIFI). While the former is an explicit government guarantee, the latter is a more subtle implicit guarantee that is embedded in the modern financial system.

Most papers that study contagion in the banking system model the crisis state as an exogenous perturbation of the model. In these models, the crisis state is therefore modeled as a zero probability event. In order to study banks' decision whether or not to form an interconnection, we develop a simplified model of financial contagion with direct linkages through an interbank market that *explicitly* takes into account the crisis state of the world. This allows us to endogenise the network formation decision with banks' recognising the risk of contagion when forming an interconnection in the form of an interbank market to share liquidity. We model this decision as an optimising behaviour of banks, where they balance the benefit of sharing liquidity against the cost of contagion.

We use this framework to explore some of the market frictions that can lead to suboptimal interconnections. One obvious market friction that can lead to such suboptimal network formation is underestimation of the risk of crisis which can stem from a variety of sources such as disaster myopia. While important, this is not the focus of our paper and we only briefly touch upon this potential source of suboptimal interconnections. We instead focus on deposit insurance and the existence of SIFIs. We show that explicit deposit insurance and more implicit too big to fail type perceptions of government guarantees to SIFIs creates a wedge between social and private optimality. In the presence of these implicit and explicit guarantees, competitive banks find it optimal to participate in the interbank market even when the risk of contagion is high and it is socially suboptimal to do so. This results in excessive interconnections in equilibrium.

We model contagion through direct balance-sheet linkages and show how these connections can result in contagious bank runs. Following Bryant (1980) and Diamond and Dybvig (1983), we model bank runs in a setting where depositors have uncertain consumption needs and the long term investment is costly to liquidate. In these papers, if depositors believe that other depositors will withdraw then all agents find it rational to redeem their claims and a bank run occurs. Another equilibrium exists where everybody believes no panic will occur and agents withdraw their funds according to their consumption needs. The theory is silent on which of the two equilibria will be selected and depositors belief is coordinated by 'sun-spots'.

Allen and Gale (2000) study how the banking system responds to contagion when banks are connected under different network structures. Banks perfectly insure against liquidity shocks by exchanging interbank deposits. The connections created by swapping deposits, however, expose the system to contagion. The authors show that incomplete networks are more prone to contagion than complete structures. Better connected networks are more resilient since the proportion of the losses in one bank's portfolio is transferred to more banks through interbank agreements. To show this, they take the case of an incomplete network where the failure of a bank may trigger the failure of the entire banking system. They prove that, for the same set of parameters, if banks are connected in a complete structure, then the system is more resilient with regard to contagious effects.

Dasgupta (2004) also explores how linkages between banks, represented by cross holding of deposits, can be a source of contagious breakdowns. However, using the global games techniques developed by Morris and Shin (2000, 2003), Dasgupta isolates an unique equilibrium which depends on depositors private signal about the banks' fundamentals. In the same spirit, Brusco and Castiglionesi (2007) show that there is a positive probability of bankruptcy and propagation of a crises across regions when banks keep interbank deposits and this may lead to excessive risk. Rochet and Vives (2004) and Goldstein and Pauzner (2005) also use global games techniques to study banking crises. Chen, Goldstein, and Jiang (2010) establishes the empirical applicability of the global games approach using a detailed dataset on mutual funds.

Our paper is also related to the literature on financial networks. Allen and Babus (2009) provide a comprehensive survey of this literature, most of which studies network effect rather than network formation. These papers focus on how different network structures respond to the breakdown of a single bank in order to identify network structures that are more resilient. Gale and Kariv (2007) study the process of exchange in financial networks and show that when networks are incomplete, substantial costs of intermediation can arise and lead to uncertainty of trade as well as market breakdowns. Eisenberg and Noe (2001) investigate default by firms that are part of a single clearing mechanism and show the existence of a clearing payment vector that defines the level of connections between firms. The authors develop an algorithm that allows them to evaluate the effects of small shocks on the system. This algorithm produces a natural measure of systemic risk based on how many waves of defaults are required to induce a given firm in the system to fail. Similarly, Minguez-Afonso and Shin (2007) use lattice-theoretic methods to study liquidity and systemic risk in high-value payment systems.

Gai and Kapadia (2010) develop a model of contagion using techniques from the literature on complex networks to assess the fragility of the financial system based on banks' capital buffers, the degree of connectivity and the liquidity of the market for failed banking assets. They find that while greater connectivity reduces the likelihood of widespread default, the shocks may have a significantly larger impact on the system when they do occur. Moreover, the resilience of the network to large shocks depends on shocks hitting particular fragile points associated with structural vulnerabilities.

There are some models that study endogenous network formation. Leitner (2005) constructs a model where the success of an agent's investment in a project depends on the investments of other agents in the network. Since endowments are randomly distributed across agents, an agent may not have enough cash to make the necessary investment. In this case, agents may be willing to bail out other agents to prevent the collapse of the whole network. Leitner examines the design of optimal financial networks that minimize the trade-off between risk sharing and the potential for collapse. In a related paper, Kahn and Santos (2010) investigate whether banks choose the optimal degree of mutual insurance against liquidity shocks. They show that when there is a shortage of exogenously supplied liquidity, which can be supplemented by bank liquidity creation, the banks generally fail to find the correct degree of interdependence. In aggregate, they become too risky.

Babus (2009) models the decision of the bank to ex-ante commit to ensure each other against the risk of contagion using a network formation game approach and shows that when banks endogenously form networks to respond to contagion risk, financial stability is supported. The model predicts a connectivity threshold above which contagion does not occur, and banks form links to reach this threshold. However, an implicit cost associated with being involved in a link prevents banks from forming more connections than required by the connectivity threshold. Banks manage to form networks where contagion rarely occurs. Castiglionesi and Navarro (2007) are also interested in decentralizing the network of banks that is optimal from a social planner perspective. In a setting where banks invest on behalf of depositors and there are positive network externalities on the investment returns, fragility arises when banks that are not sufficiently capitalized gamble with depositors' money. When the probability of bankruptcy is low, the decentralized solution approximates the first best.

The rest of the paper is organised as follows. Section 2 sets up our model and outlines the costs and benefits of participating in the interbank market. Section 3 solves the model for the banks' optimal participation strategy. In the absence of externalities and market failures, competitive banks optimal strategy is socially optimal. Section 4 introduces market frictions and explores how these frictions can create a wedge between the socially optimal and privately optimal participation strategy, resulting in excessive interconnectedness in equilibrium. Section 5 concludes. The proofs are presented in the appendix.

2 Model Setup

We consider an economy which extends over three dates, $t = 0, 1, 2$, with two non-overlapping 'regions' A and B. We assume that each region is populated by a continuum of depositors and a single representative competitive bank. Banks are competitive in the sense that they makes zero profit in equilibrium and offer depositors the highest possible return. There is a single good that is used for both consumption and investment and serves as the numeraire.

Depositors We consider a continuum of *ex-ante* identical depositors who each live for three periods. The depositors have liquidity preferences as in Diamond and Dybvig (1983) - a depositor is either impatient and wishes to consume at date 1 or patient and wishes to consume at date 2. The number of patient and impatient depositors depend on the state of the world.

Depositors are uncertain about their consumption needs at date 0, but learn their preferences privately at date 1. Let ψ denote the fraction of impatient depositors, who wish to consume at period 1. A depositors' utility function is -

$$u(c_1, c_2) = \begin{cases} c_1 & \text{with probability} \quad \psi \\ c_2 & \text{with probability} \quad 1 - \psi \end{cases}$$

where c_t is consumption at date t. Depositors are endowed with a unit of the consumption good at date 0 only, which they can invest in a storage technology or deposit at the bank. If they choose to deposit at the bank, depositors can choose to withdraw at date 1 or date 2. Given depositor preferences, impatient depositors always withdraw at date 1, while patient depositors decide strategically.

States of the World The distribution of patient and impatient depositors depend on the state of the world. In '*normal*' times, we assume that each region can have high or low demand for liquidity, such that $\psi \in \{\omega + x, \omega - x\}$. These liquidity states are negatively correlated across regions. We further assume that both states are equally likely, implying that aggregate liquidity (ω) is constant in normal times.

Additionally, there exists a '*crisis*' state when the liquidity demand is very high. In order to simplify our calculations, we assume that in the crisis state all consumers are impatient, *i.e.* $\psi = 1$.[1] These crisis states are low probability events and are independently and identically distributed across the two regions, with probability γ, where $\gamma \approx 0$. Figure 1 summaries the possible states of the world.

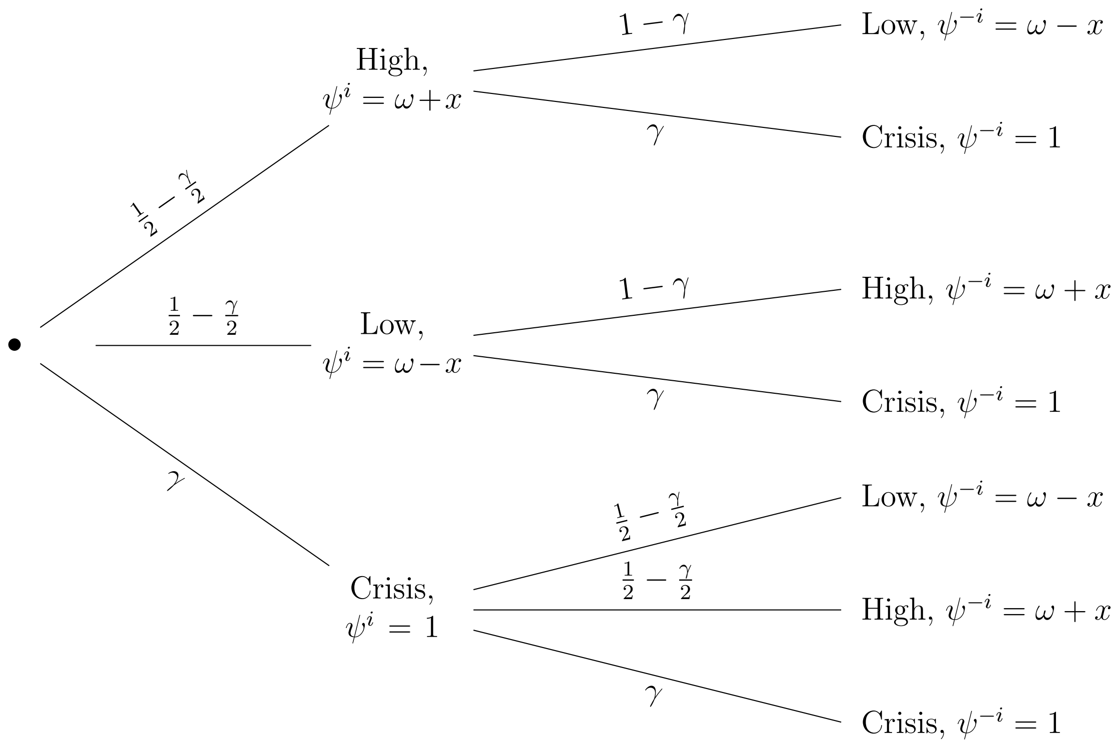


Figure 1: Liquidity States, $i \in \{A, B\}$

Assets Following Allen and Gale (2000), we assume that there are two types of

[1]Note that for our results to hold, we only require that the liquidity demand be reasonably high during the crisis state, *i.e.* $\psi >> \omega + x$. This is because when liquidity demand is sufficiently high, the patient depositors also find it optimal to withdraw, effectively making $\psi = 1$.

assets. The *liquid asset* represents storage technology. One unit of consumption good invested in the liquid asset at date t produces one unit of consumption good at period $t+1$. The other asset is the *long asset* which takes two periods to mature. Investment in the long asset can take place only at $t=0$, and the asset provides a return $R>1$ in period $t=2$. If the asset is liquidated prematurely at period $t=1$, it produces $0<r<1$ units of the consumption good. Figure 2 summaries the cash flows from the two assets.

t	0	1	2
Liquid Asset	-1	1	
Liquid Asset		-1	1
Long Asset	-1	r *(liquidated)*	R *(not liquidated)*

Figure 2: Asset Cash Flows

Deposit Contracts At date 0, banks collect deposits from depositors and offers demand deposit contracts. The demand deposit contacts' allow depositors to withdraw either c_1 units of consumption at date 1 or c_2 units of consumption at date 2. Following Dasgupta (2004), we focus on deposit contacts that offer conversion of deposits into cash at par at $t=1$, *i.e.* $c_1=1$. Note that the deposit contract is entered into at date 0 and cannot be contingent on the depositor type since the depositors learn their preferences (patient or impatient) privately.

The role of banks is to invest on the behalf of depositors and insure them against liquidity shocks. By pooling the assets of a large number of consumers, the bank can offer insurance to consumers against their uncertain liquidity demands, giving depositors some of the benefits of the long asset without subjecting them to liquidation costs in case they turn out to be impatient depositors. Our model generates multiple equilibria in line with Diamond and Dybvig (1983) and Allen and Gale (2000), *i.e.* even when the bank is solvent, patient depositors may wish to withdraw their deposits at $t=1$, if they believe that enough other patient depositors will do the same, leading to a bank run - the so called 'sunspot' equilibria. Using the global games techniques outlined in Morris and Shin (2000, 2003), Dasgupta (2004) and Goldstein and Pauzner (2005), we can generate unique equilibria in our model by introducing an arbitrarily small amount of uncertainty. However, in order to keep our model simple, we assume a deterministic return R at period $t=2$ and focus on the no-run efficient equilibrium where, as long as the bank remains solvent, patient depositors wait till period $t = 2$ to withdraw their funds. The existence of deposit insurance ensures that no-run is the dominant strategy.

Banks Competitive banks choose their balance sheets at date t=0 such that they offer depositors highest possible returns. They hold a fraction of their portfolio, y in the liquid asset and invest the remainder $(1 - y)$ in the long asset. They also decide whether or not to participate in the interbank market to insure against regional liquidity shocks. They do this by exchanging interbank deposits at date $t = 0$. Since ex-ante, bank deposits are identical, the banks budget constraint is not affected by the exchange of interbank deposits.

At date 1, the bank uses the liquid asset to service withdrawals by depositors. If the bank holds excess liquidity, *i.e.* $y \geq \psi$, it can invest the surplus in the liquid asset and use it to repay patient depositors in period 2. However, in case liquid assets are not sufficient, *i.e.* $y < \psi$, the bank has to rely on the interbank market or liquidate the long asset.

Given the asset payoffs, there exists an obvious pecking order for banks. Banks uses the liquid asset in the first instance. Next, they meet any shortfall by tapping into the interbank market. This is only possible when the banks collectively decide to participate in the interbank market at $t = 0$ and exchange interbank deposits b. Liquidation is costly and is only used as a last resort. Banks optimally choose their portfolio taking into account the expected withdrawals at date 1.

Interbank Market Since liquidity states are negatively correlated across regions, aggregate liquidity is constant in '*normal*' times. Participation in the interbank market allows banks' to share liquidity and insure against regional liquidity shocks (Allen and Gale, 2000; Dasgupta, 2004). In particular, banks hold the average level of liquidity demand, ω of the short asset and insure against regional liquidity shocks by holding interbank deposits x. This allows banks to avoid costly liquidation of the long asset. However, participation in the interbank market also exposes banks to possible contagion.

In our setting, this contagion risk is modelled by the '*crisis*' state, which occurs with a small probability γ. We assume that the interbank market shuts down in the crisis state, resulting in contagion. While we acknowledge that there are different ways to model the impact of the crisis state on the interbank market, we follow the literature on financial networks and contagion and assume that interbank deposits disappear during the crisis state. We believe that this assumption is consistent with the evidence from the recent financial crisis, where the interbank market shutdown in the aftermath of the collapse of Lehman Brothers in September 2008.[2] Furthermore, in case of insolvency, interbank deposits are likely to be junior to retail deposits and even if the interbank deposits pay-off a fraction of their value after liquidation, this is unlikely to be available at $t = 1$.

We can illustrate the contagion channel through an example. Suppose region A is in the normal state while there is a crisis in region B. In case the banks decide not to participate in the interbank market, region A banks would ex-ante hold enough of the liquid asset to ensure that they have enough liquidity to be viable in both high an low liquidity states. The bank run would therefore be limited to the region B bank.

On the other hand, participation in the interbank market allows banks to share liquidity and optimise their holding of the liquid asset. Since liquidity shocks are negatively correlated, they can use their interbank deposits to address any liquidity shortfalls in the high liquidity demand state. However, this dependence on the interbank market exposes them to contagion risk - the crisis in region B can result in a run on the region A bank. This is because the region A bank, reliant on the interbank market, would now hold insufficient liquidity to pay impatient depositors in the high liquidity state. Therefore, when the interbank market shuts down due to the crisis in region B, the region A bank is forced to liquidate the long asset, resulting in a run on the bank. Note that in our model, contagion takes place only when high liquidity demand coincides with a crisis in the other region. This is because our model abstracts from other channels of contagion such as asset fire sales. If these channels are taken into account, the probability of contagion is likely to increase significantly.

3 Model Solution - Social Optimality

We can now solve for the banks decision to participate in the interbank market. We do this in 3 stages. First, we look at the case where the bank does not participate in the interbank market. Second, we introduce the interbank market. Finally, we compare

[2]See Bank of International Settlements (2009).

the costs and benefits of participation in the interbank market and solve for the bank's optimal participation strategy.

No Interbank Market

The bank offers deposit contract (c_1, c_2) at date 0 that allows depositors to withdraw either $c_1 = 1$ at date 1 or c_2 at date 2. That bank also chooses its balance sheet and holds a fraction of its portfolio, y in the liquid asset and invests the remainder $(1 - y)$ in the long asset such that it maximises depositor utility. Note that during the crisis state, all consumers are impatient and withdraw their deposits at $t = 1$. Therefore, unless the bank portfolio consists of only liquid asset - which makes it redundant, the bank goes bust. Therefore, when choosing the balance sheet, the bank does not to take into account the crisis state and instead optimally chooses its balance sheet taking into account only the high and low liquidity demand states during normal times.

At date 1, the proportion of impatient depositors, ψ withdraw their deposits. If the bank holds enough liquid assets, *i.e.* $y \geq \psi$, the bank pays out the impatient depositors $c_1 = 1$ and invests the surplus in the short asset. At date 2, it pays patient depositors c_2, which equals the surplus cash from date 1, $(y - \psi)$ and the returns from the long asset, $(1-y)R$. On the other hand, if banks hold insufficient liquid assets, *i.e.* $y < \psi$, the bank is forced to liquidate a part of its long asset in order to pays out the impatient depositors $c_1 = 1$ at date 1. It then uses the remaining long asset to pay the patient depositors c_2 at date 2. In case the bank is unable to pay patient depositors $c_2 \geq 1$ at date 2, the patient depositors optimally withdraw their deposits at date 1, resulting in a run on the bank. Additionally, if the bank is unable to pay $c_1 = 1$ to all depositors withdrawing at date 1, it is declared insolvent and goes bust.

Banks face uncertain liquidity demand at date 1 as the distribution of patient and impatient depositors depend on the state of the world. Since liquidating the long asset prematurely at date 1 is costly for the bank, $0 < r < 1$, the bank holds liquidity in the form of investment in the liquid asset. However, this liquidity has an opportunity cost - foregone investment in the long asset. Therefore, banks face a trade off and choose their optimal portfolio taking into account the relative costs of liquidity and liquidation.

Proposition 5. *When not participating in the interbank market, the bank optimally holds* $y = \omega + x$ *of the liquid asset.*

Proof. See Appendix A.1 □

Proposition 5 implies that the bank finds it optimal to hold enough liquidity to pay impatient depositors in the high liquidity demand state, although this has an opportunity cost in terms of surplus liquid asset and foregone investment in the low liquidity demand state. When choosing their balance sheets, bank trade-off the costs of liquidation against the opportunity cost of investing in the long asset. For the rest of this paper, we assume that the relative to the cost of liquidation outweighs the benefit of holding more of the long asset. More formally,

Assumption 1. *We assume that for a given return on the long asset, liquidation is sufficiently costly such that,*

$$r < \frac{R}{2R-1}$$

Using Proposition 5, we can now write down the depositor pay-off's in different states of the world.

Proposition 6. *The depositor's payoff in different states of the world when the bank does not participate in the interbank market is as follows -*

Liquidity State	*Probability*	c_1	c_2
High Liquidity $(\omega + x)$	$\frac{1}{2} - \frac{\gamma}{2}$	1	R
Low Liquidity $(\omega - x)$	$\frac{1}{2} - \frac{\gamma}{2}$	1	$\frac{2x+(1-\omega-x)R}{1-\omega+x} < R$
Crisis	γ	$\omega + x + (1-\omega-x)r < 1$ *Bank Run, everyone withdraws at date 1*	

There is no run on the bank in normal times.

We consider each case in turn. Suppose the demand for liquidity is high, *i.e.* $\psi = \omega + x$. Since the bank hold $y = \omega + x$ of the liquid asset, the withdrawal demand from impatient depositors exactly matches up with the availability of the liquid asset and $c_1 = 1$. The investment in the long asset, $(1 - y)$ also matches up with the demand from patient depositors, $(1-\omega-x)$. Thus, $c_2 = R$. Since $c_2 > c_1 = 1$, patient depositors find it optimal to withdraw their deposits at date 2 and there is no run on the bank.

On the other hand, if the demand for liquidity is low, *i.e.* $\psi = \omega - x$, the bank has excess liquidity at date 1. The demand from impatient depositors at date 1 equals $\omega - x$, while the bank holds $y = \omega + x$ of the liquid asset. The bank invests this surplus cash in the liquid asset at date $t = 1$. At date 2, the bank has a cash flow of $2x$ from the liquid asset and $(1-\omega-x)R$ from the long asset, which it uses to pay $(1-\omega+x)$ patient depositors $c_2 = \frac{2x+(1-\omega-x)R}{1-\omega+x}$.

Corollary 3. *In the low liquidity demand state with no interbank market,* $1 < c_2 < R$

Proof. See Appendix A.2 □

$c_2 > 1$ ensures that there is no run on the bank and patient depositors wait till date 2 to withdraw their deposits. However, since the bank had excess liquidity at date 1, $c_2 < R$. This represents the cost of holding the liquid asset. Banks are forced to hold this excess liquidity in order to avoid liquidating the long asset in the high liquidity demand case.

Finally, in the crisis state, all depositors withdraw their deposit at date 1 and the bank is forced to liquidate the long asset. At date 1, the bank gets $(\omega+x)$ from the liquid asset and $(1-\omega-x)r$ from liquidating the long asset. Since $r < 1$, $c_1 = \omega+x+(1-\omega-x)r < 1$ and the bank goes bust.

We can now calculate the expected pay-off of the depositors, taking into account the ex-post distribution of patient and impatient depositors.

Proposition 7. *Expected pay-off of the depositors with no interbank market is -*

$$\begin{aligned}\Pi^{NIB} =& \left(\frac{1}{2} - \frac{\gamma}{2}\right)[(\omega+x) + (1-\omega-x)R] + \gamma\,[\omega + x + (1-\omega-x)r] \\ &+ \left(\frac{1}{2} - \frac{\gamma}{2}\right)\left[(\omega-x) + (1-\omega+x)\cdot\frac{2x+(1-\omega-x)R}{1-\omega+x}\right] \\ =& R(1-\omega-x) + \omega + x - (1-\omega-x)(R-r)\gamma\end{aligned}$$

Interbank Market

We now introduce the interbank market. Participation in interbank market expands the bank's sources of liquidity. During normal times, since liquidity shocks are negatively correlated, the interbank market allows banks in different regions to share liquidity and economise on the holding of the liquid asset. Since aggregate liquidity (ω) is fixed during normal times, a bank in a region with high liquidity demand can, through the interbank market, tap the surplus liquidity in the other region with low liquidity demand. This allows banks to offer a higher c_2 to patient depositors in normal times. However, this comes at a cost since it makes the bank dependent on the interbank market. The bank optimally no longer holds enough liquid assets to meet the demands of impatient depositors at date 1. When the interbank market disappears during a crisis, banks are exposed to contagion risk. When faced with high liquidity demand, banks are forced to liquidate the long asset.

Proposition 8. *When participating in the interbank market, the bank optimally holds $y = \omega$ of the short asset and exchanges interbank deposits of size x.*

Proof. See Appendix A.3 □

Proposition 8 holds as long as liquidity shocks are materially significant, *i.e.* crisis in the other region can lead to contagion and crisis through the disappearance of the interbank market. Formally,

Assumption 2. *We assume that liquidity shock*

$$x > \frac{(1-\omega)(R-1)}{\frac{R}{r}-1}$$

The above condition is weak and as satisfied for all plausible parameters values. This result is similar to the condition derived in Allen and Gale (2000), but in our setting the bank explicitly take into account the probability of the crisis state and potential contagion. As before, we look at the depositor's payoff on a case by case basis.

Proposition 9. *The depositor's payoff in different states of the world when the bank participates in the interbank market is as follows -*

<table>
<tr><th>Liquidity State</th><th>Other Region</th><th>Probability</th><th>c_1</th><th>c_2</th></tr>
<tr><td rowspan="2">High Liquidity
$(\omega + x)$</td><td>No Crisis</td><td>$\left(\frac{1}{2} - \frac{\gamma}{2}\right)(1-\gamma)$</td><td>1</td><td>$R$</td></tr>
<tr><td>Crisis</td><td>$\left(\frac{1}{2} - \frac{\gamma}{2}\right)\gamma$</td><td colspan="2">$\omega + (1-\omega)r < 1$
Contagion: Bank Run, everyone withdraws at date 1</td></tr>
<tr><td rowspan="2">Low Liquidity
$(\omega - x)$</td><td>No Crisis</td><td>$\left(\frac{1}{2} - \frac{\gamma}{2}\right)(1-\gamma)$</td><td>1</td><td>$R$</td></tr>
<tr><td>Crisis</td><td>$\left(\frac{1}{2} - \frac{\gamma}{2}\right)\gamma$</td><td>1</td><td>$\frac{x+(1-\omega)R}{1-\omega+x} < R$</td></tr>
<tr><td>Crisis</td><td colspan="2">γ</td><td colspan="2">$\omega + (1-\omega)r < 1$
Bank Run, everyone withdraws at date 1</td></tr>
</table>

We consider each case in turn, starting with the normal times, when there is no crisis in either region. Suppose region A banks are in the high liquidity demand state, *i.e.* $\psi = \omega + x$. At date 1, region A bank faces a withdrawal demand of $\omega + x$ from impatient depositors. The bank holds $y = \omega$ of the short asset and can withdraw its interbank deposit of x. Since the liquidity shock in region B is negatively correlated, the bank in this region faces a withdrawal demand of $\omega - x$ and can use the excess cash (arising from $y = \omega$ holding of the short asset) to pay the region A bank. Thus, the cash flows match up and $c_1 = 1$.

At date 2, the region A bank receives $(1-\omega)R$ from the long asset and has to make an interbank payment of xR to the region B bank. Since it faces a withdrawal demand of $(1 - \omega - x)$ from patient depositors, it offers $c_2 = R$. Region B bank receives $(1 - \omega)$ from the long asset and xR from the region A bank, which matches up with the demand of $(1 - \omega + x)$ from patient depositors and once again $c_2 = R$. Since $c_2 > c_1$, patient depositors find it optimal to wait and withdraw their deposits at date 2 and there is no run on either of the banks.

An analogous analysis holds for the case where there is no crisis and the region A bank faces low liquidity demand, *i.e.* $\psi = \omega - x$. Once again, there is no aggregate liquidity shock and the cash flows match up with $c_1 = 1$ and $c_2 = R$.

We now turn to the cases where there is a crisis in region B. The region A bank gets exposed to this crisis through the interbank market, or more precisely, through the disappearance of the interbank market on which it is now dependant. Suppose the bank is region A faces a high liquidity demand, *i.e.* $\psi = \omega + x$. Since it only holds $y = \omega$ of the short asset and the interbank market disappears due to the crisis, the bank in region A is forced to liquidate a fraction $l = \frac{x}{r}$ of the long asset, even though there is no crisis in region A. As long as $l \leq (1 - \omega)$, the bank is able to pay date 1 depositors c_1. However, if Assumption 2 holds, after liquidating l at date 1, the region A bank does not have enough long assets to offer patient depositors $c_2 \geq 1$ and patient depositors find it optimal to withdraw their deposits at date 1 and there is a run on the bank.

Corollary 4. *If Assumption 2 holds, contagion results in a run on the bank during the high liquidity demand state.*

Proof. See Appendix A.4 □

Contagion results in a run on the bank and it is forced to liquidate all its assets. The payoff for depositors is $c_1 = \omega + (1 - \omega)r < 1$ and the bank goes bust.

Conversely, consider the case where there is a crisis in region B, but the region A bank faces a low liquidity demand state, *i.e.* $\psi = \omega - x$. In this case $y = \omega$ of the short term asset is sufficient to meet the demands of the impatient depositors and $c_1 = 1$. However, the interbank market disappears due to crisis in region B and the region A bank is forced to invest the surplus cash (x) in the short asset. Therefore, due to the disappearance of the interbank market, it can no longer earn return R on this surplus cash.

At date 2, the region A bank receives $(1 - \omega)R$ from the long asset and has x cash, with with it has to meet the withdrawal demand of $(1 - \omega + x)$ from patient depositors. Therefore, $c_2 = \frac{x+(1-\omega)R}{1-\omega+x}$. Clearly, $1 < c_2 < R$. Since we assume that there are no fire sale externalities, there is no run on the region A bank due to the crisis in region B. Nevertheless, the region A bank still suffers due to contagion since it can no longer pay patient depositors $c_2 = R$.

Finally, we consider the crisis state in region A. The state in region B is now irrelevant. All depositors in region A withdraw their deposit at date 1 and the bank is forced to liquidate the long asset. At date 1, the bank gets (ω from the short asset and $(1 - \omega)r$ from liquidating the long asset. Since $r < 1$, $c_1 = \omega + (1 - \omega)r < 1$ and the bank goes bust.

As before, we calculate the expected pay-off of the depositors, when the bank decides to participate in the interbank market.

Proposition 10. *Expected pay-off of the depositors when the bank participates in the interbank market is -*

$$
\begin{aligned}
\Pi^{IB} =& \left(\frac{1}{2}-\frac{\gamma}{2}\right)(1-\gamma)\left[(\omega+x)+(1-\omega-x)R\right]+\left(\frac{1}{2}-\frac{\gamma}{2}\right)\gamma\left[\omega+(1-\omega)\,r\right]\\
&+\left(\frac{1}{2}-\frac{\gamma}{2}\right)(1-\gamma)\left[(\omega-x)+(1-\omega+x)R\right]+\gamma\left[\omega+(1-\omega)r\right]\\
&+\left(\frac{1}{2}-\frac{\gamma}{2}\right)\gamma\left[(\omega-x)+(1-\omega+x)\cdot\frac{x+(1-\omega)R}{1-\omega+x}\right]\\
=&\omega+(1-\omega)R-\frac{3}{2}(1-\omega)(R-r)\gamma+\frac{1}{2}(1-\omega)(R-r)\gamma^2
\end{aligned}
$$

Participation strategy

We are now in a position to derive the bank's participation strategy. Our analysis suggests that the bank faces a tradeoff - participation in the interbank market allows the bank to offer higher deposit contract (higher c_2) to depositors during normal times, but exposes it to contagion during the crisis state. Using Propositions 6 and 9 we can compare depositor payoff from not participating and participating in the interbank market in different states of the world.

The main benefit from participating in the interbank market arises during normal times, when the bank is able to offer higher deposit contracts during the low liquidity demand state. When banks do not participate in the interbank market, they are forced to hold excess liquidity in order to avoid costly liquidation of the long asset. However, in the low liquidity demand state this excess liquidity in the short asset means that the bank has to offer a lower level of consumption to patient depositors at date 2. Thus $c_2 = \frac{2x+(1-\omega-x)R}{1-\omega+x} < R$. On the other hand, participation in the interbank market allows the bank to offer depositors $c_2 = R$ by using the interbank market to earn a return R on its surplus cash.

The bank also gains marginally in the state where there is a crisis in the other region, but low liquidity demand in the bank's own region. This is because, in the low liquidity demand state, the bank does not need to access the interbank liquidity market at date $t = 1$. Although the bank looses out via-a-vis the normal state in which the bank could access the interbank market at date $t = 2$ and offer depositors $c_2 = R$, it still gains vis-a-vis non participation in the interbank market because participating in the interbank market allows banks to hold lower levels of liquidity (ω as opposed to $\omega + x$), which allows the bank to offer a marginally higher return $\frac{x+(1-\omega)R}{1-\omega+x}$ $\big(> \frac{2x+(1-\omega-x)R}{1-\omega+x}$, the return the bank can offer when it decides not to participate in the interbank market$\big)$.

Conversely, the main cost of participation in the interbank market manifests itself through the contagion channel, when the crisis in the other region results in a run on the bank in the high liquidity demand state. When the bank does not participate in the interbank market, it holds enough liquidity to pay impatient depositors in the high liquidity state without liquidating the long asset. This in turn allows the bank to pay patient depositors $c_2 = R$. However, when participating in the interbank market, the bank is dependent on the interbank market to pay impatient date 1 depositors during the high liquidity demand state. When the interbank market disappears, the bank is forced to liquidate its long asset. If assumption 2 holds, this results in a run on the bank, all depositors receive $c_1 = \omega + (1-\omega)r < 1$ and the bank goes bust.

During the crisis state, there is a run on the bank and it goes bust irrespective of whether or not it decided to participate in the interbank market. However, when participating in the interbank market, depositors face a marginally larger loss, since the bank now holds lower levels of liquidity (ω as opposed to $\omega + x$), and therefore in case of a run, can offer depositors $c_1 = \omega + (1-\omega)r < \omega + x + (1-\omega-x)r$, the return the bank can offer when it does not participate in the interbank market.

The bank optimally chooses its participation strategy by comparing the cost and benefit of participating in the interbank market, *i.e* by comparing Π^{NIB} and Π^{IB} (Proposition 7 and 10). This, in turn, depend on the probability of the crisis state, γ.

Proposition 11. $\Pi^{IB} - \Pi^{NIB}$ *is decreasing in crisis probability* γ*, i.e. the benefits of participating in the interbank market are decreasing in* γ*.*

Proof. See Appendix A.5 □

Since the main benefits of participating in the interbank market arise in normal times and the loss arises during the crisis state, the expected gain from participation in the interbank market falls as the probability of crisis increases.

Proposition 12. *There exists a* γ^* *such that* $\Pi^{IB} > \Pi^{NIB}$ *for* $\gamma < \gamma^*$ *and* $\Pi^{IB} < \Pi^{NIB}$ *for* $\gamma > \gamma^*$*.*

Proof. See Appendix A.6 □

As long as the probability of crisis is less than the critical level of γ, γ^*, *i.e.* $\gamma \leq \gamma^*$, the optimal strategy for the bank is to participate in the interbank market. However, as γ rises, the potential costs from contagion increases. When the probability of crisis is relatively significant, $\gamma > \gamma^*$ the bank finds it optimal not to participate in the interbank market.

Proposition 12 gives us the socially optimal threshold level for the crisis probability γ^* beyond which the potential costs of contagion outweigh the liquidity sharing benefits of network formation. In the absence of market frictions, private and social costs are aligned and the privately optimal strategy is also socially optimal.

4 Market Frictions - Private Optimality

In this section, we discuss some market frictions that can result in excessive interconnectedness. These market frictions create a wedge between the socially optimal and privately optimal participation strategy, resulting in suboptimal network formation that was seen during the crisis. The frictions can be broadly divided into 2 categories - (i) market frictions that result in underestimation of the probability of crisis; (ii) frictions where the probability of crisis is correctly measured but banks still optimally choose to form suboptimal interconnections.

While the underestimation of the risk of crisis is important and can stem from a variety of sources such as disaster myopia, this is not the focus of our paper and we only briefly touch upon this potential source of excessive interconnectedness. We instead focus on deposit insurance and the too-big-to-fail problem. We show that explicit deposit insurance and more implicit too big to fail type perceptions of government guarantees to SIFIs creates a wedge between social and private optimality. In the presence of these implicit and explicit guarantees, competitive banks find it optimal to participate in the interbank market even when the risk of contagion is high and it is socially suboptimal to do so ($\gamma > \gamma^*$). This results in excessive interconnectedness in equilibrium.

Deposit Insurance

Most jurisdictions require banks to participate in some form of government guaranteed deposit insurance with the intention of protecting bank deposit holders from losses that could occur in the event of a bank failure. While primarily a means of protecting individual (especially small) bank depositors, deposit insurance also serves to protect the financial system from inefficient bank runs and panics - the so called 'sunspot equilibrium' highlighted in Diamond and Dybvig (1983) and Allen and Gale (2000) where even solvent banks may fail if enough depositors believe that other depositors will withdraw their deposits.

However, deposit insurance has the potential to undermine market discipline as depositors bear little or none of the risk associated with bank failures and therefore seek higher returns, even if it entails contagion risk. This in turn compels competitive banks to offer the highest possible returns - which in our setting is possible by optimising on the holdings of the short asset and participating in the interbank market. This creates a wedge between socially optimal γ^* and depositors desired γ, leading to excessive suboptimal interconnections in equilibrium.

Since deposit guarantee is contingent on bank failure, in our model, it is only paid out in the crisis state. Furthermore, since the deposit contracts in our model offers conversion of deposits into cash at par at $t = 1$, we model deposit insurance as a guarantee that depositors get $c_1 = 1$ at $t = 1$ even if the bank is insolvent and fails. This effectively means that all depositors withdraw $c_1 = 1$ at $t = 1$, implying that the patient depositors loose out on $c_2 = R$ at $t = 2$.

Put differently, we assumes that the deposit insurance scheme only protects depositor capital and does not guarantee interest. We can make the alternative assumption that the deposit insurance also guarantees interest, which in this setting would imply that the deposit insurance scheme guarantees R to patient depositors at $t = 2$. This alternative assumption strengthens our results.

We can now revisit depositor's payoff from participating and not participating in the interbank market respectively. First consider Proposition 6 with deposit insurance.

Proposition 13. *The depositor's payoff in different states of the world when the bank does not participate in the interbank market and is protected by deposit insurance is as follows -*

Liquidity State	*Probability*	c_1	c_2
High Liquidity $(\omega + x)$	$\frac{1}{2} - \frac{\gamma}{2}$	1	R
Low Liquidity $(\omega - x)$	$\frac{1}{2} - \frac{\gamma}{2}$	1	$\frac{2x+(1-\omega-x)R}{1-\omega+x} < R$
Crisis	γ	1 *Bank Run, deposit insurance pays out*	

Comparing Proposition 6 and 13, with deposit insurance, depositors receive $c_1 = 1 > \omega + x + (1 - \omega - x)r$ during the crisis state, while their payoff remains the same during the normal state. Note that here we assume that the deposit guarantee is externally funded. This is in line with most deposit insurance schemes, where deposit insurance is government guaranteed, with minimal ex-ante fees. This assumption is also in line with the UK's Financial Services Compensation Scheme (FSCS) which envisages an ex-post 'pay-as-you-go' levy on the banking system to recover the costs of deposit insurance, supplemented by borrowing from the government in case large payouts become necessary. For example, following the default of five relatively large banks in 2008 which necessitated payouts to nearly 4 million depositors, the government provided loan facilities of approximately £20 billion to the FSCS. Under the terms of the loan, only the interest on the loan is payable with any recoveries offset against the principal loan amount.[3]

Using Proposition 13, we can calculate the expected pay-off of depositors.

Proposition 14. *Expected pay-off of the depositors when the bank does not participate in the interbank market and is protected by deposit insurance is -*

$$\begin{aligned}\Pi_{DI}^{NIB} = &\left(\frac{1}{2} - \frac{\gamma}{2}\right)[(\omega + x) + (1 - \omega - x)R] + \gamma \\ &+ \left(\frac{1}{2} - \frac{\gamma}{2}\right)\left[(\omega - x) + (1 - \omega + x) \cdot \frac{2x + (1 - \omega - x)R}{1 - \omega + x}\right] \\ =& R(1 - \omega - x) + \omega + x - (1 - \omega - x)(R - 1)\gamma\end{aligned}$$

[3]According to the HM Treasury, "The Treasury's loans to the FSCS attract interest, consistent with state aid rules set by the European Commission. State aid rules ensure that interventions by national governments do not distort competition or trade within the EU, and so help to promote a level-playing field across Europe for financial services." Therefore, it is likely that without the EU restrictions, the loans would not have attracted any interest. There terms of the loan are due to be reviewed in 2012, with the aim of agreeing on a 'credible and affordable loan repayment schedule for the £20 billion of legacy costs'.

In addition, the expected cost of deposit insurance is -

$$
\begin{aligned}
C_{DI}^{NIB} &= (1-\omega-x)(1-r)\gamma \\
&= \Pi_{DI}^{NIB} - \Pi^{NIB}
\end{aligned}
$$

Now, considering Proposition 9 with deposit insurance.

Proposition 15. *The depositor's payoff in different states of the world when the bank participates in the interbank market and is protected by deposit insurance is as follows -*

<table>
<tr><th>Liquidity State</th><th>Other Region</th><th>Probability</th><th>c_1</th><th>c_2</th></tr>
<tr><td rowspan="2">High Liquidity
$(\omega+x)$</td><td>No Crisis</td><td>$\left(\frac{1}{2}-\frac{\gamma}{2}\right)(1-\gamma)$</td><td>1</td><td>$R$</td></tr>
<tr><td>Crisis</td><td>$\left(\frac{1}{2}-\frac{\gamma}{2}\right)\gamma$</td><td colspan="2">1
Contagion: Bank Run, deposit insurance pays out</td></tr>
<tr><td rowspan="2">Low Liquidity
$(\omega-x)$</td><td>No Crisis</td><td>$\left(\frac{1}{2}-\frac{\gamma}{2}\right)(1-\gamma)$</td><td>1</td><td>$R$</td></tr>
<tr><td>Crisis</td><td>$\left(\frac{1}{2}-\frac{\gamma}{2}\right)\gamma$</td><td>1</td><td>$\frac{x+(1-\omega)R}{1-\omega+x} < R$</td></tr>
<tr><td>Crisis</td><td colspan="2">γ</td><td colspan="2">1
Bank Run, deposit insurance pays out</td></tr>
</table>

Once again, the existence of deposit insurance reduces depositors loss during the crisis state since they now receive $c_1 = 1 > \omega + (1-\omega)r$. Note that the gain from deposit insurance is marginally higher when banks participate in the interbank market since in the absence of deposit insurance, the banks liquidation value during the crisis state is lower when participating in the interbank market. This is driven by the fact that banks optimally invest less in the liquid asset (ω) when participating in the interbank market vis-a-vis when not participating in the interbank market $(\omega + x)$, reducing their liquidation values.

As before, participation in the interbank market exposes the bank to contagion and a crisis in the other region results in a run on the bank in the high liquidity demand state. However, with deposit insurance, depositors still get $c_1 = 1 > \omega + (1-\omega)r$, depositors payoff in the absence of deposit insurance. Using Proposition 15, we can calculate the expected pay-off of depositors.

Proposition 16. *Expected pay-off of the depositors when the bank participates in the* *interbank market and is protected by deposit insurance is -*

$$\begin{aligned}\Pi^{IB}_{DI} = &\left(\frac{1}{2}-\frac{\gamma}{2}\right)(1-\gamma)\left[(\omega+x)+(1-\omega-x)R\right]+\left(\frac{1}{2}-\frac{\gamma}{2}\right)\gamma \\ &+\left(\frac{1}{2}-\frac{\gamma}{2}\right)(1-\gamma)\left[(\omega-x)+(1-\omega+x)R\right]+\gamma \\ &+\left(\frac{1}{2}-\frac{\gamma}{2}\right)\gamma\left[(\omega-x)+(1-\omega+x)\cdot\frac{x+(1-\omega)R}{1-\omega+x}\right] \\ =&\omega+(1-\omega)R-\frac{3}{2}(1-\omega)(R-1)\gamma+\frac{1}{2}(1-\omega)(R-1)\gamma^2\end{aligned}$$

In addition, the expected cost of deposit insurance is -

$$\begin{aligned}C^{IB}_{DI} &= \left(\frac{1}{2}-\frac{\gamma}{2}\right)\gamma(1-\omega)(1-r)+\gamma(1-\omega)(1-r) \\ &= \frac{3}{2}(1-\omega)(1-r)\gamma-\frac{1}{2}(1-\omega)(1-r)\gamma^2 \\ &= \Pi^{IB}_{DI}-\Pi^{IB}\end{aligned}$$

Overall from Propositions 14 and 16, it is clear that for every level of γ, $\Pi^{IB}-\Pi^{NIB}$ is strictly greater with deposit insurance. Therefore, the threshold level of γ up to which depositors find it optimal for banks to participate in the interbank market is greater in the presence of deposit insurance.

Proposition 17. *The critical level of γ is higher with deposit insurance, i.e. $\gamma^{DI} > \gamma^*$.*

Proof. See Appendix A.7 □

Figure 3 shows that in the presence of deposit insurance, depositors' expected gain from participating in the interbank market is positive for all $\gamma < \gamma^{DI}$ (dashed line). Thus for all such γ depositors prefer to deposit their endowment with the bank that participates in the interbank market and offers the deposit contract $(c_1 = 1, c_2 = R)$ in the 'normal' state of the world. Although this exposes the depositors to the crisis state in the other region, deposit insurance guarantees that the depositors get $c_1 = 1$ in the crisis state. Since banks are competitive in our setting, in order to attract deposits, they are forced to participate in the interbank market and offer depositors deposit contracts $(c_1 = 1, c_2 = R)$ in the 'normal' state of the world.

In contrast, in the absence of deposit insurance, depositors gain from participation in

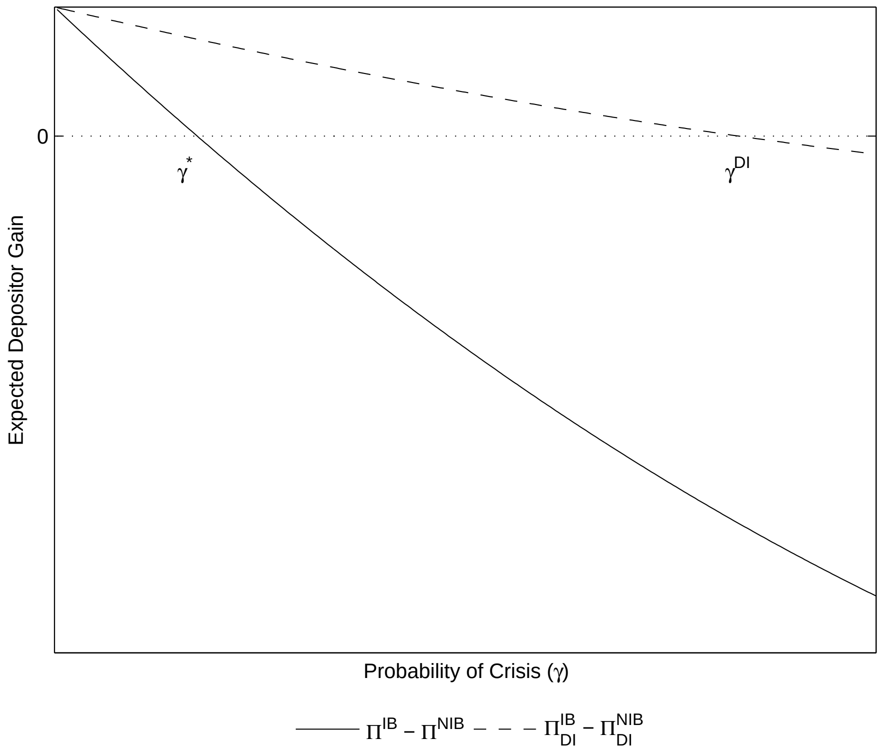


Figure 3: Expected Gain from Participating in the Interbank Market with and without deposit insurance.

the interbank market only for $\gamma < \gamma^*$ (solid line). This is because depositors now only receive the liquidation value during the crisis state, *i.e.* $c_1 < 1$ in the crisis state. Thus, for $\gamma^* > \gamma > \gamma^{DI}$, banks optimally participate in the interbank market only in the presence of deposit insurance, *i.e.* deposit insurance incentives banks to form interconnections which banks would otherwise not have formed.

Intuitively, in our model, deposit insurance acts as a state contingent cash transfer from the government to the depositors through the banking system. During the crisis state, depositors loss is limited to $c_1 = 1$ as opposed to the liquidation value of the bank where $c_1 < 1$. This incentivises depositors to take on greater risk as they do not internalise the cost of providing this insurance creating a wedge between social and private optimality.

Proposition 18. *For $\gamma^* < \gamma < \gamma^{DI}$, banks participation in the interbank market is socially sub-optimal.*

Proof. See Appendix A.8 □

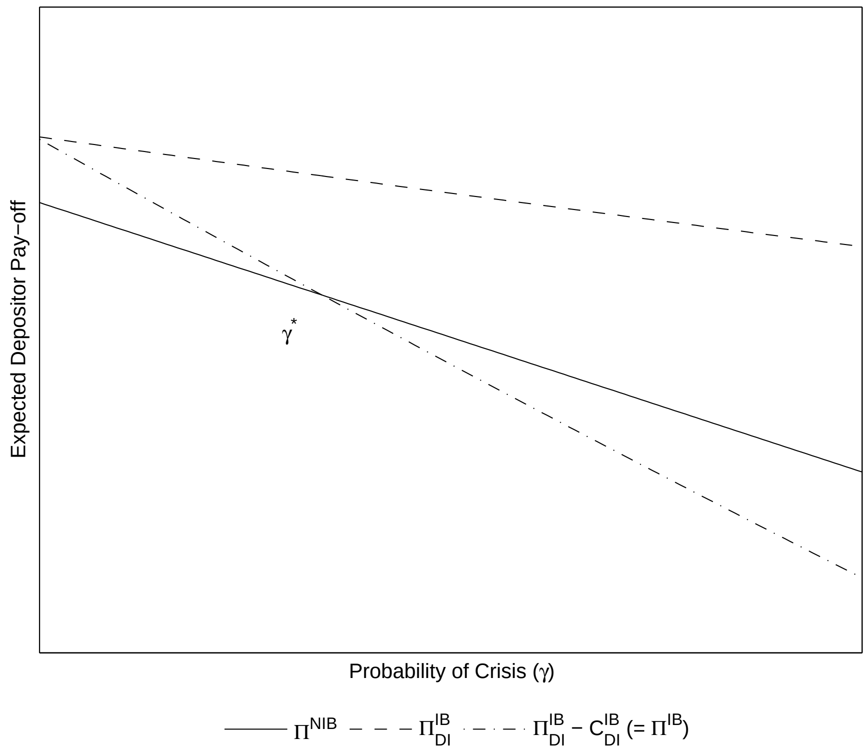


Figure 4: Expected Depositor Pay-off under (i) no interbank market; (ii) interbank market with deposit insurance; and (iii) interbank market with deposit insurance taking into account the expected cost of insurance.

In order to compare the welfare effects of deposit insurance, we need to explicitly take into account the cost to the government of providing this insurance. More formally, for $\gamma^* > \gamma > \gamma^{DI}$, we need to compare the expected depositor payoff when the bank participates in the interbank market taking into account the expected cost to the government of providing deposit insurance, *i.e.* $\Pi^{IB}_{DI} - C^{IB}_{DI}$ with the expected depositor payoff in the absence of deposit insurance when the bank does not participate in the interbank market, *i.e.* Π^{NIB}.

Proposition 18 formally shows that $\Pi^{NIB} > \Pi^{IB}_{DI} - C^{IB}_{DI}$, but more intuitively it is easy to see from Figure 4 that participation in the interbank market is socially suboptimal for any $\gamma > \gamma^*$. The solid line represents the expected depositor payoff in the absence of deposit insurance when the bank does not participate in the interbank market (Π^{NIB}) while the dashed line shows the expected depositor payoff when the bank participates in the interbank market and the depositor is protected by deposit insurance (Π^{IB}_{DI}). Finally, the dot-dashed line represents the expected depositor payoff when the bank participates in the interbank market taking into account the expected cost to the government of providing deposit insurance ($\Pi_{DI}^{IB} - C_{DI}^{IB}$).

The key thing to note here is that the once the cost of deposit insurance is taken into account, the expected depositor payoff from participating in the interbank market is equal to the depositor payoff in the absence of deposit insurance, *i.e.* $\Pi_{DI}^{IB} - C_{DI}^{IB} = \Pi_{IB}$. This follows directly from Proposition 16, but intuitively, since deposit insurance in our model is essentially a state contingent transfer from the government to the depositors, if its costs are properly taken into account, it is no coincidence that the depositors gain is exactly offset by the governments cost of providing this insurance.

Now it is easy to see that for all $\gamma^* < \gamma < \gamma^{DI}$, the benefits of participating in the interbank market after accounting for the cost of deposit insurance is less than the benefits of not participating in the interbank market (note that from Propositions 14, $\Pi^{NIB} = \Pi_{DI}^{NIB} - C_{DI}^{NIB}$). Therefore, socially it is not optimal for banks to participate in the interbank for all $\gamma > \gamma^*$. In other words, deposit insurance creates a wedge between the socially optimal γ^* and the privately optimal γ^{DI}.

Systemically Important Financial Institutions

While deposit insurance is an example of an explicit market friction that can result in a wedge between private and social optimality of network formation, the existence of Systemically Important Financial Institutions (SIFI) is an example of an implicit market friction that can result in suboptimal network formation. While there is no consensus on the precise definition of a SIFI, in general terms they can be defined as institutions whose disorderly failure, because of their size, complexity and systemic interconnectedness, would cause significant disruption to the wider financial system and economic activity.[4]

Such institutions benefit from an implicit government guarantee that they will not be allowed to fail and in the event of a crisis, they will be bailed out. Interestingly, one of the 'criteria' for being a SIFI is their interconnectedness. Therefore, in the context of our model, if we assume that the bank decides to participate in the interbank market and forms a network, it will be classified as a SIFI. In this setting, it is clear that the optimal strategy for the depositor and the bank would be to participate in the interbank

[4]See Financial Stability Board (2010, 2011)

market and form a network, irrespective of the crisis probability.

This result stems from the fact that the SIFI status ensures that during the crisis state, the bank benefits from a government bailout and can continue to function. The key difference from deposit insurance is that when the crisis state materialises, deposit insurance guarantees that in the event of insolvency, depositors get at least $c_1 = 1$ at $t = 1$ when the bank fails. However, in the case of a SIFI, the bank is not allowed to fail and receives government support to continue operating as normal and honour all its commitments, including the deposit contract ($c_1 = 1, c_2 = R$). Furthermore, since there is no insolvency, there is no possibility of a crisis state and therefore no contagion channel.

This implies that our model is restricted to the normal state of the world, with zero probability of the crisis state. Therefore in the case of a SIFI, the bank and the depositors will always find it privately optimal to form a network irrespective of probability of a crisis.

Underestimation of γ

Market frictions such as deposit insurance and the existence of SIFIs can result in excessive interconnections even in cases where the probability of crisis is correctly measured. The other source of excessive and sub-optimal network formation is the underestimation of crisis probability, γ. Such underestimation can stem from a variety of sources.

There is an extensive literature on disaster myopia - misperception of the probability of rare events that arises from a set of heuristics commonly used to form judgements under uncertainty. Tversky and Kahneman (1974) highlight the so called 'availability heuristic', which states that agents base probabilistic assessments on the ease with which an event can be brought to mind: how recently it has occurred, how severe are its effects and how personal is the experience.

Car crashes are the classic example. These often arise from disaster myopia, as drivers systematically under-estimate the probability of a pile-up and drive too fast. The longer the period since the last crash, the greater the risk-taking. After a lengthy stretch of clear motorway, risk appetite may become too healthy, and risk-taking too great, relative to the true probability of disaster. In other words, drivers are disaster-myopic. Kunreuther, Ginsberg, Miller, Sagi, Slovic, Borkan, and Katz (1978) conducted a widely cited study of disaster protection. As described in Kunreuther (1996), a 1974 survey of more than 1,400 homeowners in hurricane-prone areas of the US found that only 22% of respondents had voluntarily adopted any protective measures. Haldane (2011) discusses the role of such biases in the context of the financial crisis.

Given that before the recent financial crisis, bank runs were not seen in the UK for over 140 years and in Western Europe for the last 15 years, it is possible that such heuristic biases could have result in the underestimation of the probability of the crisis. In the context of our model, this underestimation of γ would have resulted in excessive interconnectedness between banks.

5 Conclusion

In order to study the role of market frictions that can result in the excessive interconnectedness that was seen during the crisis, we develop a simplified model of financial contagion. Departing from the literature which generally models contagion as an exogenous perturbation of the system, our model explicitly take into account the crisis state of the world. This allows us to endogenise the network formation decision with banks' recognising the risk of contagion when forming a network. We model the decision to form a network as an optimising behaviour of banks, where they balance the benefit of forming a network against the cost of contagion.

We use this framework to study how various market frictions can create a wedge between the socially optimal and the privately optimal threshold for network formation, resulting in excessive interconnectedness of the financial system. While such suboptimal network formation can arise from the several sources that can lead to the underestimation of the crisis probability, our results suggest that even in cases where the probability of crisis is perfectly observed, regulatory intervention in the form of deposit insurance and the existence of systemically important financial institutions can result in socially suboptimal network formation and excessive interconnectedness in equilibrium.

References

Allen, F., and A. Babus, 2009, "Networks in Finance," *The network challenge: strategy, profit, and risk in an interlinked world*, p. 367.

Allen, F., A. Babus, and E. Carletti, 2009, "Financial crises: theory and evidence," *Annu. Rev. Financ. Econ.*, 1(1), 97–116.

Allen, F., and D. Gale, 2000, "Financial contagion," *Journal of political economy*, pp. 1–33.

Babus, A., 2009, "The Formation of Financial Networks," *Discussion Paper 06-093, Tinbergen Institute.*

Bank of International Settlements, 2009, "79th Annual Report," *Basel, Switzerland.*

Brusco, S., and F. Castiglionesi, 2007, "Liquidity coinsurance, moral hazard, and financial contagion," *The Journal of Finance*, 62(5), 2275–2302.

Bryant, J., 1980, "A model of reserves, bank runs, and deposit insurance," *Journal of Banking & Finance*, 4(4), 335–344.

Castiglionesi, F., and N. Navarro, 2007, *Optimal fragile financial networks*. Tilburg University.

Chen, Q., I. Goldstein, and W. Jiang, 2010, "Payoff complementarities and financial fragility: Evidence from mutual fund outflows," *Journal of Financial Economics*, 97(2), 239–262.

Dasgupta, A., 2004, "Financial contagion through capital connections: a model of the origin and spread of bank panics," *Journal of the European Economic Association*, 2(6), 1049–1084.

Diamond, D., and P. Dybvig, 1983, "Bank runs, deposit insurance, and liquidity," *The Journal of Political Economy*, pp. 401–419.

Eisenberg, L., and T. Noe, 2001, "Systemic risk in financial systems," *Management Science*, pp. 236–249.

Financial Stability Board, 2010, "Reducing the moral hazard posed by systemically important financial institutions," *FSB Recommendations and Time Lines*.

——— , 2011, "Progress in the Implementation of the G20 Recommendations for Strengthening Financial Stability," *Report of the Financial Stability Board to G20 Finance Ministers and Central Bank Governors*.

Gai, P., and S. Kapadia, 2010, "Contagion in financial networks," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 466(2120), 2401.

Gale, D., and S. Kariv, 2007, "Financial networks," *The American economic review*, 97(2), 99–103.

Goldstein, I., and A. Pauzner, 2005, "Demand–Deposit Contracts and the Probability of Bank Runs," *the Journal of Finance*, 60(3), 1293–1327.

Haldane, A. G., 2009, "Rethinking the Financial Network," *speech delivered at the Financial Student Association, Amsterdam*.

——— , 2011, "Risk Off," *Bank of England*.

Kahn, C., and J. Santos, 2010, "Liquidity, Payment and Endogenous Financial Fragility," in *EFA 2005 Moscow Meetings Paper*.

Kunreuther, H., 1996, "Mitigating disaster losses through insurance," *Journal of risk and Uncertainty*, 12(2), 171–187.

Kunreuther, H., R. Ginsberg, L. Miller, P. Sagi, P. Slovic, B. Borkan, and N. Katz, 1978, *Disaster insurance protection: Public policy lessons*. Wiley New York.

Leitner, Y., 2005, "Financial networks: Contagion, commitment, and private sector bailouts," *The Journal of Finance*, 60(6), 2925–2953.

Minguez-Afonso, G., and H. Shin, 2007, "Systemic Risk and Liquidity in Payment Systems," *working paper London School of Economics*.

Morris, S., and H. Shin, 2000, "Rethinking Multiple Equilibria in Macroeconomic Modelling," *Cowles Foundation Discussion Papers*.

——— , 2003, “Global games: theory and applications,” *Advances in Economics and Econometrics: Theory and Applications, Volume One*, pp. 56–114.

Rochet, J., 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.

Tversky, A., and D. Kahneman, 1974, “Judgment under uncertainty: Heuristics and biases,” *American Association for the Advancement of Science*, 185(4157), 1124.

A Proofs

A.1 Proposition 5

When not participating in the interbank market, the bank optimally holds $y = \omega + x$ of the liquid asset.

Clearly, given the opportunity cost of the liquid asset, banks do not optimally hold liquid asset of size greater than $\omega + x$.[5] But suppose holds $y = \omega + x - \epsilon$ of the liquid asset, where $\epsilon \to 0$. We show that given assumption 1, such a deviation is not optimal and the bank optimally holds $y = \omega + x$ of the liquid asset.

First consider the high liquidity demand case, *i.e.* $\psi = \omega + x$. Since the bank only holds $y = \omega + x - \epsilon$ of the liquid asset, in order to pay $\omega + x$ of impatient depositors $c_1 = 1$ at date $t = 1$, it has to liquidate $l = \frac{\epsilon}{r}$ of the liquid asset. Assuming that the bank can do so without causing a run on the bank, it can no longer pay patient depositors $c_2 = R$ at date $t = 2$.

$$\begin{aligned} c_2 &= \frac{\left(1 - \omega - x + \epsilon - \frac{\epsilon}{r}\right) R}{1 - \omega - x} \\ &= R - \underbrace{\frac{\epsilon R}{1 - \omega - x}\left(\frac{1}{r} - 1\right)}_{\text{Loss due to liquidation}} \end{aligned}$$

Next, we consider the low liquidity demand case, *i.e.* $\psi = \omega - x$. After paying out impatient depositors $c_1 = 1$ at date $t = 1$, the bank has surplus cash of $2x - \epsilon$. However, this surplus cash is now less compared to the case where the bank holds $y = \omega + x$ of the liquid asset. Therefore,

$$\begin{aligned} c_2 &= \frac{2x - \epsilon + (1 - \omega - x + \epsilon) R}{1 - \omega + x} \\ &= \underbrace{\frac{2x + (1 - \omega - x) R}{1 - \omega + x}}_{c_2 \text{ when } y = \omega + x} + \underbrace{\frac{\epsilon (R - 1)}{1 - \omega + x}}_{\text{Gain due to smaller y}} \end{aligned}$$

[5] Note that the bank can increase its liquidation value during the crisis state by holding more of the liquid asset. However, since by assumption all consumers are impatient and withdraw their deposits at $t = 1$, the bank will always go bust unless it only holds liquid assets - in which case it becomes redundant.

Comparing the loss and the gain, taking into account the ex-post distribution of patient and impatient depositors, we can show that the loss due to liquidation outweighs the gain due to lower opportunity cost of investing in the liquid asset.

$$
\begin{aligned}
&(1-\omega-x)\cdot\frac{\epsilon R}{1-\omega-x}\Big(\frac{1}{r}-1\Big) > (1-\omega+x)\cdot\frac{\epsilon(R-1)}{1-\omega+x} \\
&\implies R\Big(\frac{1}{r}-1\Big) > (R-1) \\
&\implies r < \frac{R}{2R-1}
\end{aligned}
$$

which is true by Assumption 1.

A.2 Corollary 3

In the low liquidity demand state with no interbank market, $1 < c_2 < R$*, where* $c_2 = \frac{2x+(1-\omega-x)R}{1-\omega+x}$

$$
\begin{aligned}
&c_2 > 1 \\
&\implies \frac{2x+(1-\omega-x)R}{1-\omega+x} > 1 \\
&\implies 2x+(1-\omega-x)R > 1-\omega+x \\
&\implies (1-\omega-x)(R-1) > 0
\end{aligned}
$$

Since $(1-\omega-x) > 0$ and $R > 1 \implies c_2 > 1$.

$$
\begin{aligned}
&c_2 < R \\
&\implies \frac{2x+(1-\omega-x)R}{1-\omega+x} < R \\
&\implies 2x+(1-\omega-x)R < (1-\omega+x)R \\
&\implies x < R
\end{aligned}
$$

Since $x < 1$ and $R > 1 \implies c_2 < R$.

A.3 Proposition 8

When participating in the interbank market, the bank optimally holds $y = \omega$ of the liquid asset and exchanges interbank deposits of size x.

We show this in two stages. First, we show that the bank hold just enough liquidity (in the form of investment in the liquid asset and interbank deposits) to pay impatient depositors $c_1 = 1$ at date $t = 1$ during '*normal*' times. Next, we show that banks find it optimal to hold this liquidity in the form of ω of the liquid asset and interbank deposits of size x.

Given the opportunity cost of liquid assets, it is obvious that the bank would not hold liquidity greater than $\omega + x$, the demand from depositors in the high liquidity demand state. But suppose they hold less. They can do this either by holding smaller interbank deposits or by investing less in the liquid asset. We consider both cases in turn.

Case 1(a): Suppose the bank invests $y = \omega$ in the liquid asset and holds interbank deposits of size $x - \epsilon$, where $\epsilon \to 0$. We can show that such a deviation is not optimal and banks are better off holding interbank deposits of size x. Since interbank deposits disappear during the crisis state, their size does not matter in such states and it is sufficient to compare the payoffs during '*normal*' times.

Suppose the bank faces high liquidity demand, *i.e.* $\psi = \omega + x$ and has to meet the withdrawal demand of $\omega + x$ from impatient depositors at $t = 1$. The bank holds ω of the liquid asset and can withdraw the interbank deposit of $x - \epsilon$. This leaves the bank with a deficit of ϵ which it must meet by liquidating a fraction $l = \frac{\epsilon}{r}$ of the long asset. However, due to early liquidation of the long asset, the bank can no longer pay date 2 depositors $c_2 = R$. At date 2, the bank gets $(1 - \omega - \frac{\epsilon}{r})R$ from its remaining long assets and has to pay $(x - \epsilon)R$ for the interbank deposits and meet withdrawal demand of $(1 - w - x)$. This implies that

$$
\begin{aligned}
c_2 &= \frac{\big(1 - \omega - \frac{\epsilon}{r} - (x - \epsilon)\big)R}{(1 - \omega - x)} \\
&= R - \underbrace{\frac{\epsilon R}{1 - \omega - x}\left(\frac{1}{r} - 1\right)}_{\text{Loss due to liquidation}}
\end{aligned}
$$

Clearly, this is decreasing in ϵ and the bank can get $c_2 = R$ if it sets $\epsilon = 0$.

Next, when the bank faces low liquidity demand, *i.e.* $\psi = \omega - x$, it has a cash surplus of ϵ at date 1, after paying out $x - \epsilon$ for interbank deposits and $(\omega - x)$ to the impatient depositors. At date 2, it uses this surplus cash, $(1 - \omega)R$ return from the long asset, and $(x - \epsilon)R$ from interbank deposits to pay $(1 - \omega + x)$ of patient depositors. This implies that

$$
\begin{aligned}
c_2 &= \frac{\epsilon + (1-\omega)R + (x-\epsilon)R}{(1-\omega+x)} \\
&= R \quad \underbrace{-\frac{\epsilon(R-1)}{1-\omega+x}}_{\text{Loss due to cash surplus}}
\end{aligned}
$$

Once again, this is decreasing in ϵ and the bank finds it optimal to set $\epsilon = 0$ and increase $c_2 = R$. Therefore, given $y = \omega$ holdings of the liquid asset, the bank optimally holds interbank deposit of size x.

Case 1(b): Suppose instead the bank invests $y = \omega - \epsilon$ in the liquid asset and holds interbank deposits of size x. During '*normal*' times, the analysis remains the same as above. However, since the bank now holds only $y = \omega - \epsilon$ of the liquid asset, its liquidation value decreases. During the crisis state, as well as in the high liquidity demand state when there is a run on the bank, the banks liquidation value,

$$
\begin{aligned}
c_1 &= (\omega - \epsilon) + (1 - \omega + \epsilon)r \\
&= \omega + (1-\omega)r \quad \underbrace{-\epsilon(1-r)}_{\text{Loss due to smaller y}}
\end{aligned}
$$

Therefore, the bank can increase its liquidation value by setting $\epsilon = 0$. Finally, consider the case where the bank faces low liquidity demand, *i.e.* $\psi = \omega - x$ and there is a crisis in the other region. Interbank deposits disappear and the bank has surplus cash of $x - \epsilon$ at date 1 after paying out $c_1 = 1$ to $(\omega - x)$ impatient depositors. At date 2, it uses this surplus cash and $(1 - \omega + \epsilon)R$ return from the long asset to pay $(1 - \omega + x)$ of patient depositors. This implies,

$$
\begin{aligned}
c_2 &= \frac{(x-\epsilon)+(1-\omega+\epsilon)R}{(1-\omega+x)} \\
&= \frac{x+(1-\omega)R}{(1-\omega+x)} \quad \underbrace{+\frac{\epsilon(R-1)}{1-\omega+x}}_{\text{Gain due to smaller y}}
\end{aligned}
$$

Although the bank can increase c_2 in this case, it is easy to see that for all $\gamma < \frac{1}{2}$, in expectation this gain is less than the loss during the no crisis state. Therefore, banks cannot do better by holding less liquidity, either by holding smaller interbank deposits or by investing less in the liquid asset.

We now show that banks find it optimal to hold this liquidity in the form of ω of the liquid asset and interbank deposits of size x.

Case 2(a): Suppose the bank holds $\omega-\epsilon$ of the liquid assets and interbank deposits of size $x+\epsilon$. When the bank faces high liquidity demand, the cash flows match up during normal times and $c_1 = 1$ and $c_2 = R$. However, during the low liquidity demand state, the bank has a liquidity shortfall of 2ϵ, even during normal times - $(\omega-\epsilon)$ cash inflow from the investment in the liquid asset vs the outflow of $(x+\epsilon)$ on account of interbank deposits and $(\omega - x)$ to pay impatient depositors. The bank is forced to liquidate $l = \frac{2\epsilon}{r}$ of the long asset. This implies that

$$
\begin{aligned}
c_2 &= \frac{\left(1-\omega+\epsilon-\frac{2\epsilon}{r}+(x+\epsilon)\right)R}{(1-\omega+x)} \\
&= R \underbrace{-\frac{2\epsilon R}{1-\omega+x}\left(\frac{1}{r}-1\right)}_{\text{Loss due to liquidation}}
\end{aligned}
$$

Clearly, this is decreasing in ϵ and the bank can get $c_2 = R$ if it sets $\epsilon = 0$. During the crisis state in the other region, the interbank deposits disappear, but there is no bank run in the low liquidity demand case and the bank has a smaller cash surplus of $(x-\epsilon)$

at date 1. Therefore,

$$\begin{aligned} c_2 &= \frac{(x-\epsilon)+(1-\omega+\epsilon)R}{(1-\omega+x)} \\ &= \frac{x+(1-\omega)R}{(1-\omega+x)} \quad \underbrace{+\frac{\epsilon(R-1)}{1-\omega+x}}_{\text{Gain due to smaller y}} \end{aligned}$$

As in Case 1(b), although the bank can increase c_2 in this case, for all $\gamma < \frac{1}{2}$, in expectation this gain is less than the loss during the no crisis state. Finally, in the event of a bank run, the banks liquidation value falls.

$$\begin{aligned} c_1 &= (\omega-\epsilon)+(1-\omega+\epsilon)r \\ &= \omega+(1-\omega)r \quad \underbrace{-\epsilon(1-r)}_{\text{Loss due to smaller y}} \end{aligned}$$

Therefore, the bank is better off with $\epsilon = 0$.

Case 2(b): Suppose the bank holds $\omega+\epsilon$ of the liquid assets and interbank deposits of size $x-\epsilon$. As before, during normal times, the cash flows match up in the high liquidity demand state. But during the low liquidity demand state, the bank has surplus liquidity of 2ϵ - $(\omega+\epsilon)$ cash inflow from the investment in the liquid asset vs the outflow of $(x-\epsilon)$ on account of interbank deposits and $(\omega-x)$ to pay impatient depositors. This implies that

$$\begin{aligned} c_2 &= \frac{2\epsilon+(1-\omega-\epsilon)R+(x-\epsilon)R}{(1-\omega+x)} \\ &= R \quad \underbrace{-\frac{2\epsilon(R-1)}{(1-\omega+x)}}_{\text{Loss due to cash surplus}} \end{aligned}$$

This is decreasing in ϵ and the bank can get $c_2 = R$ if it sets $\epsilon = 0$.

During the crisis state in the other region, the interbank deposits disappear, but there is no bank run in the low liquidity demand case and the bank has a surplus cash of $(x+\epsilon)$

at date 1. Therefore,

$$\begin{aligned} c_2 &= \frac{(x+\epsilon)+(1-\omega-\epsilon)R}{(1-\omega+x)} \\ &= \frac{x+(1-\omega)R}{(1-\omega+x)} \quad \underbrace{-\frac{\epsilon(R-1)}{1-\omega+x}}_{\text{Loss due to cash surplus}} \end{aligned}$$

Once again we see that the optimal strategy for the bank is to set $\epsilon = 0$, *i.e.* hold $y = \omega$ of the liquid asset and exchange interbank deposits of size x

However, in the event of a bank run, the liquidation value of the bank increases when it holds $y = \omega + \epsilon$ of the liquid asset.

$$\begin{aligned} c_1 &= (\omega+\epsilon)+(1-\omega-\epsilon)r \\ &= \omega+(1-\omega)r \quad \underbrace{+\epsilon(1-r)}_{\text{Gain due to larger y}} \end{aligned}$$

Based on this, we can compute the expected gain during liquidation and the expected loss when there is no bank run.

Expected gain during bank runs: $\left(\frac{1}{2}-\frac{\gamma}{2}\right)\gamma\epsilon(1-r)+\gamma\epsilon(1-r)$

Expected loss when no bank run: $\left(\frac{1}{2}-\frac{\gamma}{2}\right)(1-\gamma)2\epsilon(R-1)+\left(\frac{1}{2}-\frac{\gamma}{2}\right)\gamma\epsilon(R-1)$

As long as $\frac{1-r}{R-1} < \frac{\gamma^2-\gamma+2}{\gamma(1-\gamma)}$, the expected loss is greater than the expected gain and banks find it optimal to hold ω of the liquid asset and interbank deposits of size x. The above condition is relatively weak and holds for most parameter values. Intuitively, it says that the probability of the crisis should not be too high relative to the liquidation cost and the return of the long term assets. For practical purposes, it is only violated when the liquidation cost is close to 0 and/or R is very close to 1 and at the same time, the probability of the crisis is very high.

A.4 Corollary 4

If Assumption 2 holds, contagion results in a run on the bank during the high liquidity demand state.

In the case of a crisis in region B, the region A bank is forced to liquidate a fraction

$l = \frac{x}{r}$ of the long asset to meet the withdrawal demand of $\psi = \omega + x$ during the high liquidity state and pay impatient depositors $c_1 = 1$. Note that this implicitly assumes that $l \leq (1 - w)$. In case $l > (1 - w)$, the bank cannot pay $c_1 = 1$ to date one depositors even after it liquidates all its long assets. It pays $c_1 < 1$ and therefore goes bust.

Assuming this does not hold, the bank has $(1 - \omega - l)$ of the long asset left, with which it has to pay $(1 - \omega - x)$ patient depositors. Thus,

$$(1 - \omega - l)R = (1 - \omega - x)c_2$$
$$\implies c_2 = \frac{(1 - \omega - \frac{x}{r})R}{1 - \omega - x}$$

For there to be a run on the bank, we need to show,

$$\begin{aligned}
&c_2 < 1 \\
&\implies (1 - \omega - \frac{x}{r})R < 1 - \omega - x \\
&\implies (1 - \omega)R - x\frac{R}{r} < 1 - \omega - x \\
&\implies (1 - \omega)(R - 1) < x\left(\frac{R}{r} - 1\right) \\
&\implies x > \frac{(1 - \omega)(R - 1)}{\frac{R}{r} - 1}
\end{aligned}$$

Given Assumption 2, this is always true.

A.5 Proposition 11

$\Pi^{IB} - \Pi^{NIB}$ *is decreasing in crisis probability* γ*, i.e. the benefits of participating in the interbank market are decreasing in* γ*.*

Using Propositions 7 and 10, we have

$$\begin{aligned}
\Pi^{IB} - \Pi^{NIB} =& \omega + (1 - \omega)R - \frac{3}{2}(1 - \omega)(R - r)\gamma + \frac{1}{2}(1 - \omega)(R - r)\gamma^2 \\
& - R(1 - \omega - x) + \omega + x - (1 - \omega - x)(R - r)\gamma \\
=& x(R - 1) - \frac{1}{2}(1 - \omega + 2x)(R - r)\gamma + \frac{1}{2}(1 - \omega)(R - r)\gamma^2
\end{aligned}$$

Taking the partial derivative with respect to γ,

$$\frac{\partial}{\partial\gamma}(\Pi^{IB} - \Pi^{NIB}) = -\frac{1}{2}(1-\omega+2x)(R-r) + (1-\omega)(R-r)\gamma$$
$$\text{Now, } \frac{\partial}{\partial\gamma}(\Pi^{IB} - \Pi^{NIB}) < 0 \text{ if } \gamma < \frac{1-\omega+2x}{2(1-\omega)}$$

This is always true for all $\gamma < \frac{1}{2}$, implying that $\Pi^{IB} - \Pi^{NIB}$ is decreasing in the crisis probability γ.

A.6 Proposition 12

There exists a γ^* such that $\Pi^{IB} > \Pi^{NIB}$ for $\gamma < \gamma^*$ and $\Pi^{IB} < \Pi^{NIB}$ for $\gamma > \gamma^*$.

We solve for γ^* such that $\Pi^{IB} - \Pi^{NIB} = 0$.

$$\Pi^{IB} - \Pi^{NIB} = x(R-1) - \frac{1}{2}(1-\omega+2x)(R-r)\gamma + \frac{1}{2}(1-\omega)(R-r)\gamma^2 = 0$$

Solving for γ^*, we have

$$\gamma^* = \frac{\frac{1}{2}(1-\omega+2x)(R-r) - \sqrt{\frac{1}{4}(1-\omega+2x)^2(R-r)^2 - 2(1-\omega)(R-r)x(R-1)}}{(1-\omega)(R-r)}$$

Using Proposition 11, it is clear that $\Pi^{IB} > \Pi^{NIB}$ for $\gamma < \gamma^*$ and $\Pi^{IB} < \Pi^{NIB}$ for $\gamma > \gamma^*$.

A.7 Proposition 17

The critical level of γ is higher with deposit insurance, i.e. $\gamma^{DI} > \gamma^*$.

From the proof of Proposition 12, we have the gain from participating in the interbank market as a function of γ in the absence of deposit insurance.

$$\Pi^{IB} - \Pi^{NIB} = x(R-1) - \frac{1}{2}(1-\omega+2x)(R-r)\gamma + \frac{1}{2}(1-\omega)(R-r)\gamma^2$$

Similarly, we can calculate the gain from participating in the interbank market as a function of γ in the presence of deposit insurance.

$$\Pi_{DI}^{IB} - \Pi_{DI}^{NIB} = x(R-1) - \frac{1}{2}(1-\omega+2x)(R-1)\gamma + \frac{1}{2}(1-\omega)(R-1)\gamma^2$$

Comparing these two equations, we can show that,

$$\begin{aligned}
&(\Pi_{DI}^{IB} - \Pi_{DI}^{NIB}) > (\Pi^{IB} - \Pi^{NIB}) \\
\implies &(\Pi_{DI}^{IB} - \Pi_{DI}^{NIB}) - (\Pi^{IB} - \Pi^{NIB}) > 0 \\
\implies &\frac{1}{2}(1-\omega+2x)(1-r)\gamma - \frac{1}{2}(1-\omega)(1-r)\gamma^2 > 0 \\
\implies &(1-\omega+2x) > (1-\omega)\gamma
\end{aligned}$$

Since $0 < \gamma < 1$ and $x > 0$, this condition holds.

This implies that for every level of γ, the gains from participating in the interbank market are greater in the presence of deposit insurance. This in turn implies that $\gamma^{DI} > \gamma^*$.

Using the expression for $\Pi_{DI}^{IB} - \Pi_{DI}^{NIB}$, we can explicitly solve for the value of γ^{DI}. Solving, we have

$$\gamma^{DI} = \frac{2x}{1-\omega}$$

Note that we ignore the other root $\gamma^{DI} = 1$ as we are interested in small γ.

A.8 Proposition 18

For $\gamma^* < \gamma < \gamma^{DI}$, banks participation in the interbank market is socially sub-optimal.

For $\gamma^* < \gamma < \gamma^{DI}$, in the absence of deposit insurance the bank optimally does not participate in the interbank market and the expected depositor payoff is Π^{NIB} (Proposition 7). On the other hand, with deposit insurance, it is privately optimal for the bank to participate in the interbank market and the expected depositor payoff is Π_{DI}^{IB} (Proposition 16). In order to assess the social optimality of participation in the interbank market, we need to explicitly take into account the costs of deposit insurance, C_{DI}^{IB}. Participation is socially sub-optimal when $\Pi^{NIB} > \Pi_{DI}^{IB} - C_{DI}^{IB}$.

From Proposition 16, we have $C_{DI}^{IB} = \Pi_{DI}^{IB} - \Pi^{IB}$. Substituting, it is sufficient to show that $\Pi^{NIB} > \Pi^{IB}$. This always holds for all $\gamma > \gamma^*$ (Proposition 12). Therefore, for $\gamma^* < \gamma < \gamma^{DI}$, banks participation in the interbank market is socially sub-optimal.

Competition, Premature Trading and Excess Volatility

Pragyan Deb *Bonsoo Koo* *Zijun Liu*

p.deb@lse.ac.uk *bonsoo.koo@monash.edu* *z.liu@lse.ac.uk*

Financial Markets Group
London School of Economics and Political Science

Abstract

A substantial body of research suggests that it is difficult to account for all of the asset price volatility in terms of news. This paper attempts to explain the excess volatility puzzle as a consequence of competitive interaction between market participants. We develop a model of competitive interaction in the presence of noisy, but verifiable information. Our model shows that in the presence of competitive pressures, market participants find it optimal to act prematurely on unverified, noisy information. This premature reaction leads to lower total profits and excess market volatility in equilibrium. Our model also suggests that the spike in volatility at the closing time of the market can be modelled as a direct consequence of premature trading.

Keywords: Premature trading, Competition, Unverified Information, Excess Volatility

JEL Classifications: G12, G14

1 Introduction

One of the enduring puzzles in the finance literature is excess volatility. Shiller (1981) showed that equity markets are more volatile than what fundamentals, *i.e.* changes in the level of dividends and interest rates, suggest. Although Kleidon (1986) cast some doubt on Shiller's findings, there exists accumulating evidence that not all asset price movements reflect changes in fundamental values. LeRoy and Porter (1981), Roll (1984, 1988), Mankiw, Romer, and Shapiro (1985), West (1988), Cutler, Poterba, and Summers (1989) and Ebrahim and Mathur (2001) amongst several others have all documented a significant amount of volatility that cannot be explained by changes in fundamentals. After an extensive survey, Gilles and LeRoy (1991) conclude *"this finding of excess volatility is robust and is difficult to explain within the representative consumer, frictionless market model"*.

In this paper, we attempt to explain the excess volatility puzzle as a consequence of competitive interaction between market participants. We assume that the initial information received by market participants is very noisy, often taking the form of a rumour or speculation. This could be attributed to information uncertainty *à la* Zhang (2006) and Epstein and Schneider (2008).

Ideally, market participants should wait to verify the initial noisy information before trading on it. However, we show that in the presence of competitive pressures, market participants, fearful of losing out on the trading opportunity, find it optimal to act prematurely on unverified information. This result stems from the fact that each market participant is afraid that if he does not act quickly enough, he will miss out on the opportunity. We show that this premature reaction leads to excess market volatility and lower expected profits in equilibrium.

Furthermore, high frequency data has shown systematic patterns in the dynamics of intraday volatility. In particular, there is strong evidence suggesting that market volatility spikes as the market closing time approaches. We show that this spike in volatility at the closing time of the market can be modelled a direct outcome of premature trading. Intuitively, as the market closing time draws near, market participants tend to increase premature trading out of fear of not being able to use the information at hand.

We assume that if they wait till the next trading day, the information is revealed to the market and ceases to be a profitable trading opportunity. Thus as the market nears its closing time, the time left to execute the trade decreases. As a result, market participants tend to trade even more on unverified noisy initial information, resulting in the spike in market volatility during the end of the trading day.

We model competitive interaction between market participants in a three-period rational expectations equilibrium (REE) model, where informed traders trade against a price setting competitive market maker, in presence of a noise trader. Informed traders initially receive a noisy signal which they can verify over time. This evolving nature of information plays a key role in our setting. Informed traders can either trade early on unverified, noisy signal or they can choose to wait and trade on verified information.

The informed traders' payoff from either action is contingent on the actions of other informed traders. The uncertainty surrounding the actions of other traders leads to coordination failure and lower profits in equilibrium. The intuition behind the coordination failure in our model is similar to the one employed in Abreu and Brunnermeier (2003). While in Abreu and Brunnermeier (2003), the uncertainty surrounding the action of informed traders (rational arbitragers in their setting) results in the persistence of asset price bubbles, in our model it leads to premature trading and excess volatility.

The rest of the paper is organised as follows. Section 2 places our paper in the existing academic literature. In Section 3, we describe the basic setup of our model. We then solve the model in Section 4 for the baseline one informed trader case, followed by two and N informed traders respectively. Section 5 outlines the implication of our results for total expected profits of the informed traders and market volatility. We also look at the impact of market closing in this section. Section 6 concludes.

2 Literature Review

Our paper is related to the extensive market microstructure literature on price formation in the presence of asymmetric information. Asymmetric information is generally modelled in the form of informed traders – agents with private information unavailable to the wider market. These informed traders are distinct from insiders, usually defined as corporate officers with fiduciary obligations to the shareholders. Informed traders trade in order to profit from their private information. Noise traders are liquidity motivated, smoothing their inter-temporal consumption stream through portfolio adjustment. Bagehot (1971) provides a detailed discussion of the trading motivations and the resultant distinction between noise traders who possess no special informational advantages and are primarily liquidity motivated and informed traders with private information who try to maximise profits.

In these models, the informed traders' trade against the price setting competitive market maker, in the presence of noise traders. The market maker makes losses on the trades with the informed traders, but recoups these loses on trades with the noise traders, making zero profit on average. Sequential trade models such as Glosten and Milgrom (1985) and Easley and O'Hara (1987) show that a rational market maker will quote regret free ex-post bid and ask prices which reflect the expected value of the security given the observed order. Implicit in the bid-ask formulation is the assumption that the market maker takes into account the direction of trade and the set of public information includes all information at that time, including the knowledge of the trade itself. The bid-ask spread is therefore increasing in information asymmetry and the degree of asset value volatility. The market maker trades off the reduction in losses to the informed trader from the wider spread against the opportunity cost in terms of profits from trading with uninformed noise traders with reservation prices.

Our modelling strategy is similar to the one employed in the seminal Kyle (1985) paper. In the static version of the Kyle (1985) model, the market maker sets the price after observing the aggregate order flow – a batch clearing model. The market maker sets the price equal to his best estimate, given his belief about the insiders trading strategy. Kyle derives a perfect Bayesian Nash equilibrium strategy where the informed trader's profit is increasing in his informational advantage and market depth. Intuitively, greater market depth implies that an additional order from the informed trader would not lead to a large change in prices, allowing the informed trader to trade more aggressively on his private information. Kyle (1985) also extends this static model to a dynamic setting focussing on the profit maximising temporal decision of the informed trader. In the dynamic version of the model, if the insider takes a larger position on the early periods, his early profits increase but this comes at a cost of revealing his private information to the market. As a result, the prices in the later trading rounds worsen. The optimal strategy for the informed trader is to exploit the informational advantage over time by hiding behind the noise traders.

Several papers check the robustness of the Kyle (1985) set-up. Biais and Rochet (1997) demonstrate that multiple equilibria exist in a discrete setting. Rochet and Vila (1994) consider a setting where the informed trader observes the amount of noise trading and show that the expected profits stay the same. Rochet and Vila (1994) show that the equilibrium in their setting coincides with Kyle (1989) where informed traders submit demand schedules. Vives (1995) shows that the market clearing process, where a competitive risk neutral market maker submits demand schedules based on public information and the informed trader submits market orders is equivalent to the classic Kyle (1985) set-up. Bagnoli, Viswanathan, and Holden (2001) provide a comprehensive overview of the existence of linear equilibria in static Kyle (1985)-type models.

Holden and Subrahmanyam (1992) generalise Kyle's model to incorporate competition amongst informed traders. They show that such competition results in high trading volumes and rapid revelation of private information, which is compatible with strong-form market efficiency. Relative to Kyle's model, markets are more efficient, volumes are higher and the profits of insiders are much lower. They argue, in the same spirit as Spiegel and Subrahmanyam (1992) that insider trading may not be a concern in the presence of competition amongst agents. However, they do not allow for informational gains from waiting which is usually the case in financial markets. Moreover, in practice, we hardly observe strong-form market efficiency their conclusion envisages. By allowing for the noisy information available to the informed traders, our model is more realistic and produces different results, which suggests that enhanced competition may hinder price discovery and lead to excess market volatility. In this regard, our paper complements Holden and Subrahmanyam (1992).

Admati and Pfleiderer (1988) also develop a model of strategic play by informed and uninformed traders and allow for the uninformed traders to have discretion as to which time period they would trade in. They show that this can result in concentrated bouts of trading, similar to the spike in volatility at the start and close of trading. While this finding is consistent with our model, Admati and Pfleiderer approach excess volatility from concentrated-trading stemming from the strategic behaviour of noise traders while we focus on the behaviour of informed traders who strategically decide whether to trade early on unverified information or wait and verify the accuracy of the information.

Our paper also contributes to the growing strand of recent literature which studies the effect of unverified initial information or ambiguous interpretations of incoming information on asset returns and volatility. For example, Barron and Karpoff (2004) find that trading volume reactions to public announcements are most sensitive to announcement precision among low-transaction cost securities and in low-cost trading regimes. More recently,Chen and Zhao (2012) find that the informed trading effect is both independent of and stronger than that of information uncertainty in determining price momentum. Lu, Chen, and Liao (2010) also examine the effects of information uncertainty and information asymmetry on corporate bond yield spreads.

Our paper also contributes to the literature on the spike in volatility at the closing time of the market. Research based on high frequency data suggests that returns volatility varies systemically over the trading day, with Wood, McInish, and Ord (1985) and Harris (1986) documenting the existence of a distinct 'U-shaped' pattern in return volatility over the trading day. In particular, these papers document spikes in market volatility as the market closing time approaches. Several other studies such as McInish and Wood (1990a), Foster and Viswanathan (1990), Niemeyer and Sandas (1994), Aitken, Brown, Buckland, Izan, and Walter (1996) also document such spikes. In the presence of a lunch break in the Japanese stock exchange, Andersen, Bollerslev, and Cai (2000) find that the Nikkei 225 index volatility is significantly higher at the opening of the morning and close of the afternoon session, with an increase in volatility immediately before and after the lunch break, resulting in two distinct 'U-shapes', one in the morning and one in the afternoon.

A number of studies have sought to rationalise this 'U-shaped' pattern in intraday volatility by strategic interaction of asymmetrically informed agents. As discussed above, the Admati and Pfleiderer (1988) model relies on the behaviour of some uninformed traders who strategically execute their trades in order to minimise their trading costs. Brock and Kleidon (1992) on the other hand model this as a result of portfolio rebalancing at trading halts based on the idea that the optimal portfolio is a function of the ability to trade. They show that transactions demand at open and close is greater and less elastic than at other times in the trading day.

Slezak (1994) develops a multi-period model in which market closure delays the resolution of uncertainty, threby redistributing risk across time and agents. Since agents are risk averse in their model, this redistribution affects equilibrium price, altering risk premia and liquidity cost. This allows them to generate a variety of mean and variance effect, including those mirroring the observed spike in closing time volatility. Foster and Viswanathan (1990) analyze the implications of adverse selection in securities markets for the intertemporal behavior of trading volume, trading costs, and price volatility, when there is periodic variation in the information advantage of an informed trader. Based on a variant of the Kyle (1985) model, they show that because the price is an important source of information for uninformed liquidity traders, the informed trader has the greatest advantage when the market first opens; and, the longer the market is closed, the more significant is the advantage of the informed trader at the opening.

3 Model Setup

We consider a three period model, $t = 0, 1, 2$, with three types of agents - informed traders, a noise trader and a market maker. There is a risky asset that pays a liquidating dividend V at $t = 2$, where

$$V = \begin{cases} H & \text{with probability} \quad \frac{1}{2} \\ L & \text{with probability} \quad \frac{1}{2} \end{cases}$$

Let $\bar{V} = \frac{1}{2} \cdot H + \frac{1}{2} \cdot L$ denote the expected value of the asset at $t = 0$.

The asset can be traded at $t = 0$ and $t = 1$ via a competitive market maker. The asset generates a signal s about V, which is observed privately by the informed trader(s) at $t = 0$. The signal can be either *'positive'*$(= s_H)$ or *'negative'*$(= s_L)$. The signal, $s(\in \{s_H, s_L\})$ is correct with probability λ and wrong with probability $1 - \lambda$. In order to make the signal informative, we assume $\lambda > \frac{1}{2}$. Thus, if the informed trader receives a positive signal, then at $t = 2$, the asset is worth H with probability λ and L with probability $1 - \lambda$. Formally, λ is the conditional probability which can be expressed as

$$\lambda = \begin{cases} \Pr(V = H, t = 2 | s = s_H, t = 0) \\ \Pr(V = L, t = 2 | s = s_L, t = 0) \end{cases}$$

Likewise,

$$1 - \lambda = \begin{cases} \Pr(V = L, t = 2 | s = s_H, t = 0) \\ \Pr(V = H, t = 2 | s = s_L, t = 0) \end{cases}$$

The signal can be verified costlessly, but takes one period to do so. Thus, if the informed trader decides to verify the signal, he has to wait and trade at period $t = 1$. This information structure plays a key role in our model and in the presence of competitive pressures, leads to premature trading. Informed traders strategically decide whether to trade immediately based on the noisy signal at $t = 0$, or wait and verify the signal and trade at $t = 1$, at which time the informed trader has perfect information. We assume that the informed trader submits market order and can only trade once which can be motivated by high transaction costs.

Following the established literature, we assume that the noise trader submits demand which can be d or $-d$ with probability $\frac{1}{2}$. This implies that if the informed trader decides to trade, he can either submit demand d or $-d$. Any other demand would instantly reveal his identity to the market maker. We assume noise trader trades at both $t = 0$ and $t = 1$.

Following Kyle (1985), trading in our model takes place through a two-step auction mechanism. The informed and noise traders simultaneously submit market orders which are observed by the market maker. After observing the demand, the market maker sets the price and trades the quantity that clears the market. However, unlike the Kyle (1985) model where the market maker only observes aggregate order flow, we assume a model of call markets, where the market maker observes individual orders, but not the identity of the individual submitting the order, *i.e.* cannot identify whether a particular order comes from the noise trader or the informed trader.

Effectively, the market maker's information set is larger than in Kyle (1985), allowing him to refine his beliefs and set the price more accurately.[1] We believe that this formula-

[1]In our model, the market maker can distinguish between an order {d, d, -d} and {d} while in the

tion is closer to real world markets than the commonly used Kyle approach of aggregate order flow. Such models with call markets have been studied in Kyle (1989), Kyle and Vila (1991) and Rochet and Vila (1994). Corb (1993) and Bagnoli, Viswanathan, and Holden (2001) analyse equilibria in these types of models. In the appendix, we check for the robustness of our result and present an alternative version of our model with the classic Kyle (1985) assumptions.

Finally, the market maker, after observing demand sets the price and clears the market. Everyone is risk-neutral and the risk-free rate is zero. The informed traders maximise expected profit, taking into account the effect of their trades. The price determined by the market maker is the market maker's best conjecture of the liquidation value of the asset, given all his information.

4 Model Solution

We now solve the model for increasing levels of competition - for one informed trader, for two informed traders and finally, for N informed traders. This allows us to study the effect of competition on the trading behaviour of market participants.

4.1 One informed trader

As a first step, we solve the model for the '*baseline case*' of a single informed trader. The informed trader decides whether to trade early at period $t = 0$ based on the signal or to wait and verify the signal and trade later at $t = 1$. In this case, the informed trader faces no competitive pressures and trades with the market maker in the presence of the noise trader. We look for a symmetric, perfect Bayesian Nash equilibrium (PBE). We find that the optimal strategy of the informed trader is to wait and verify the signal before trading.

Proposition 19. *There exists a unique PBE, such that the informed trader will*

- *buy d assets at $t = 1$ if he confirmed a positive signal at $t = 0$*

Kyle (1985) model the aggregate demand would be the same, *i.e.* d.

- *sell d assets at $t = 1$ if he confirmed a negative signal at $t = 0$*
- *never trade at $t = 0$*

Proof. The market maker observes the demand from the noise trader and the informed trader. Based on the observed demand at $t = 0$, D_0, the market maker sets the price, P_0, equal to his best guess of the liquidation value of the asset.

$$P_0 = \begin{cases} \bar{V} & \text{if} \quad D_0 = \{d\} \\ \bar{V} & \text{if} \quad D_0 = \{-d\} \\ \lambda H + (1-\lambda)L & \text{if} \quad D_0 = \{d, d\} \\ \bar{V} & \text{if} \quad D_0 = \{d, -d\} \\ \lambda L + (1-\lambda)H & \text{if} \quad D_0 = \{-d, -d\} \end{cases}$$

Note that in this case price does not change at $t = 0$ (in equilibrium) since the market maker knows that the informed trader will only trade at $t = 1$. This is because the informed trader does not want to reveal his information to the market maker at $t = 0$. Hence the last three cases $\{d, d\}$, $\{d, -d\}$ and $\{-d, -d\}$ are not in equilibrium and the market maker attributes any demand at $t = 0$ to noise traders.

At $t = 1$, we get three cases. If the market maker observes two positive orders, *i.e.* $D_1 = \{d, d\}$, he knows that the informed trader has submitted a $+d$ order and since the informed trader has perfect information at $t = 1$, he sets price $P_1 = H$. Conversely, on observing two negative orders, $D_1 = \{-d, -d\}$, he sets price $P_1 = L$. Finally, if he observes one positive and one negative order, *i.e.* $D_1 = \{d, -d\}$, he does not know whether the informed trader submitted $+d$ or $-d$, and thus the price remains unchanged, $P_1 = \bar{V}$. Note that this is the only case where the informed trader can profit from the information.

$$P_1 = \begin{cases} H & \text{if} \quad D_1 = \{d, d\} \\ \bar{V} & \text{if} \quad D_1 = \{d, -d\} \\ L & \text{if} \quad D_1 = \{-d, -d\} \end{cases}$$

Given the market maker's belief (and the resulting prices), we evaluate the informed trader's strategy at $t = 0$. Assume that the informed trader gets a high signal s_H at $t = 0$, *i.e.* $V = H$ at $t = 2$ with probability λ. The low signal case is analogous.

If the informed trader submits demand $+d$ based on the high signal at $t = 0$, his profits depend on the actions of the noise trader. If the noise trader also submits demand $+d$ (with probability $\frac{1}{2}$), the market maker observes demand $D_0 = \{d, d\}$ and sets the price equal to the expected profits of the informed trader, $\lambda H + (1 - \lambda)L$, given the accuracy of the signal λ. The informed trader makes zero profit in this case. On the other hand, if the noise trader submits demand $-d$ (again with probability $\frac{1}{2}$), the market maker, unable to distinguish between the noise and informed trader demand, keeps the price at $\bar{V}$ and the informed trader's expected profit equals $\lambda H + (1 - \lambda)L - \bar{V}$. Thus, the informed trader profits from trading early at $t = 0$,

$$\begin{aligned}\Pi_0 &= \frac{1}{2} \cdot 0 + \frac{1}{2}\big(\lambda H + (1 - \lambda)L - \bar{V}\big) \\ &= \frac{1}{2}\big(\lambda H + (1 - \lambda)L - \bar{V}\big)\end{aligned}$$

On the other hand, suppose the informed trader waits to verify his signal and trades at $t = 1$. We now get four cases. Suppose the signal is correct (probability λ) and the value of the asset turns out to be H at $t = 2$. The informed trader submits demand $+d$. Like before, his profits are contingent on the actions of the noise trader. If the noise trader also submits $+d$ demand (with $\frac{1}{2}$ probability), the market maker becomes informed and the informed trader makes zero profit. If, on the other hand, the noise trader submits $-d$ demand, the market maker remains uninformed and the informed trader makes a profit $H - \bar{V}$. Conversely, if the signal turns out to be incorrect (probability $(1 - \lambda)$) and the asset value turns out to be L, the informed trader submits demand $-d$. If the noise trader also submits $-d$, the informed trader makes zero profit, while if the noise trader submits d, his profit equals $\bar{V} - L$.

Put differently, at $t = 1$, the informed trader knows V and makes a profit only if the noise trader trades in the opposite direction (with probability $\frac{1}{2}$). In this case the profit is $H - \bar{V} = \bar{V} - L$. Thus, the informed trader's profit from trading on the verified signal at $t = 1$,

$$
\begin{aligned}
\Pi_1 &= \lambda\Big(\frac{1}{2}\cdot 0 + \frac{1}{2}\big(H - \bar{V}\big)\Big) + (1-\lambda)\Big(\frac{1}{2}\cdot 0 + \frac{1}{2}\big(\bar{V} - L\big)\Big) \\
&= \frac{1}{2}\big(H - \bar{V}\big)
\end{aligned}
$$

We can see $\Pi_1 > \Pi_0$ given $\lambda < 1$. Therefore the informed trader will always trade at $t = 1$.

□

Proposition 1 implies that in the absence of competitive pressures, the dominant strategy for the informed trader is to wait and verify the signal. So there is no premature trading in this case. This makes intuitive sense since in the absence of competition, there is no incentive to trade earlier than necessary.

To set ideas, we can think of an example where a brokerage house has exclusive access to some unverified information on a looming merger and acquisition deal between two companies. The brokerage house is not yet aware of the details of the deal or even whether the deal would finally materialise. In such a situation, the brokerage house can either take a bet and trade on its information - in which case it would find it quite difficult to assess the precise valuation; or alternatively, it can spend some time assessing the details of the deal (and/or waiting for the terms of the deal to be finalised). Analogous to the one informed trader case, if the brokerage house was sure that the information would not become available to the wider market (or to its competitors), then it would prefer to wait.

The quality of information (or the noise embedded in the signal) plays a critical role in our model and is the key driver behind the strategic action of the informed trader. The extent of the advantage from waiting to verify the signal depends on the accuracy of signal. As the accuracy of information, λ increases, the difference between Π_1 and Π_0 and hence the advantage of waiting falls. In the extreme, when $\lambda = 1$, i.e. the correct information is available in the first place, there is no reason that the informed trader postpones his action.

4.2 Two informed traders

We now introduce competition in the form of a second informed trader. We assume that the two informed traders receive a common signal at $t = 0$. Let x denote the probability that each informed trader trades at period $t = 0$.

Proposition 20. *There exists a unique symmetric PBE, such that the informed traders will*

- *with probability x, buy d assets at $t = 0$ if they receive a positive signal at $t = 0$*
- *with probability $(1-x)$, buy d assets at $t = 1$ if the positive signal is confirmed, else sell d assets*
- *with probability x, sell d assets at $t = 0$ if they receive a negative signal at $t = 0$*
- *with probability $(1 - x)$, sell d assets at $t = 1$ if the negative signal is confirmed, else buy d assets*

where x is the mixed strategy (symmetric) equilibrium,

$$x = \frac{2\lambda - 1}{4\lambda - 2\lambda^2 - \frac{1}{2}}$$

Proof. As before, the market maker observes the demand from the two informed traders and the noise trader and sets the price. At $t = 0$,

$$P_0 = \begin{cases} \lambda H + (1-\lambda)L & \text{if} \quad D_0 = \{d, d, d\} \\ \lambda H + (1-\lambda)L & \text{if} \quad D_0 = \{d, d\} \\ \lambda H + (1-\lambda)L & \text{if} \quad D_0 = \{d, d, -d\} \\ \bar{V} & \text{if} \quad D_0 = \{d\} \\ \bar{V} & \text{if} \quad D_0 = \{d, -d\} \\ \bar{V} & \text{if} \quad D_0 = \{-d\} \\ \lambda L + (1-\lambda)H & \text{if} \quad D_0 = \{d, -d, -d\} \\ \lambda L + (1-\lambda)H & \text{if} \quad D_0 = \{-d, -d\} \\ \lambda L + (1-\lambda)H & \text{if} \quad D_0 = \{-d, -d, -d\} \end{cases}$$

Note that the only profit-relevant case is $D_0 = \{d, -d\}$, *i.e.* when only one informed trader trades and the noise trader trades against the informed trader.

In setting the price at $t = 1$, the market maker takes into account demand at $t = 0$. Since the informed traders can trade only once, the trades that take place in period $t = 0$ give the market maker valuable information about who is potentially trading in period $t = 1$. If the market maker observes three orders in period $t = 0$, say $D_0 = \{d, d, d\}$ or $D_0 = \{d, d, -d\}$,[2] he knows that both informed traders traded at $t = 0$. Therefore, any demand the market maker observes at $t = 1$ is attributable to the noise trader and contains no new information and hence there is no price impact.

- If $D_0 = \{d, d, d\}$ or $D_0 = \{d, d, -d\}$

$$P_1 = P_0 = \lambda H + (1-\lambda)L$$

- If $D_0 = \{-d, -d, d\}$ or $D_0 = \{-d, -d, -d\}$

$$P_1 = P_0 = \lambda L + (1-\lambda)H$$

On the other hand, if the market maker observes only one order at $t = 0$, he knows

[2]Since the signals are correlated, we cannot have the case where one informed trader submits $+d$ while the other submits $-d$.

that both informed traders will be trading at $t = 1$. In this case, irrespective of the actions of the noise trader, the informed traders signal is revealed to the market maker. The market maker sets the price accordingly and hence the informed traders cannot profit from their information.

- If $D_0 = \{d\}$

$$P_1 = \begin{cases} H & \text{if} \quad D_1 = \{d, d, d\} \\ H & \text{if} \quad D_1 = \{d, d, -d\} \\ L & \text{if} \quad D_1 = \{-d, -d, d\} \\ L & \text{if} \quad D_1 = \{-d, -d, -d\} \end{cases}$$

- If $D_0 = \{-d\}$

$$P_1 = \begin{cases} H & \text{if} \quad D_1 = \{d, d, d\} \\ H & \text{if} \quad D_1 = \{d, d, -d\} \\ L & \text{if} \quad D_1 = \{-d, -d, d\} \\ L & \text{if} \quad D_1 = \{-d, -d, -d\} \end{cases}$$

Finally, when the market maker observes two orders in $t = 0$, only one informed trader trades at $t = 1$ and the actions of the noise trader becomes important. If the market maker observes $D_1 = \{d, d\}$, he knows that one of the informed traders has submitted $+d$ in this period, and since informed traders in period $t = 1$ have perfect information, he sets $P_1 = H$. Conversely, if he observes demand $D_1 = \{-d, -d\}$, he sets $P_1 = L$. However, in case $D_1 = \{d, -d\}$, the market maker does not know whether the informed trader submitted $+d$ or $-d$ and the price remains the same, *i.e.* $P_1 = P_0$.

- If $D_0 = \{d, d\}$

$$P_1 = \begin{cases} H & \text{if} \quad D_1 = \{d, d\} \\ P_0 = \lambda H + (1 - \lambda) L & \text{if} \quad D_1 = \{d, -d\} \\ L & \text{if} \quad D_1 = \{-d, -d\} \end{cases}$$

- If $D_0 = \{d, -d\}$

$$P_1 = \begin{cases} H & \text{if} \quad D_1 = \{d, d\} \\ \bar{V} & \text{if} \quad D_1 = \{d, -d\} \\ L & \text{if} \quad D_1 = \{-d, -d\} \end{cases}$$

- If $D_0 = \{-d, -d\}$

$$P_1 = \begin{cases} H & \text{if} \quad D_1 = \{d, d\} \\ P_0 = \lambda L + (1-\lambda)H & \text{if} \quad D_1 = \{d, -d\} \\ L & \text{if} \quad D_1 = \{-d, -d\} \end{cases}$$

Given the prices, we can evaluate each informed trader's strategy at $t = 0$. The intuition behind these equations is the same as in the one informed trader case, only more complicated because of the additional cases. As before, the informed traders profit from trading at $t = 0$,

$$\begin{aligned} \Pi_0 &= (1-x) \cdot \frac{1}{2} \cdot \big((\lambda H + (1-\lambda)L - \frac{1}{2}(H+L)\big) \\ &= \frac{1}{2}(1-x)(\lambda - \frac{1}{2})(H-L) \end{aligned}$$

On the other hand, if he waits to verify his signal and trades at $t = 1$,

$$\begin{aligned} \Pi_1 = x \cdot \Big(&\frac{1}{2}\big(\lambda \cdot \frac{1}{2}(H - \frac{1}{2}(H+L)) + (1-\lambda) \cdot \frac{1}{2}\frac{1}{2}(\frac{1}{2}(H+L) - L)\big) \\ &+ \frac{1}{2}\big(\lambda \cdot \frac{1}{2}(H - \lambda H - (1-\lambda)L) + (1-\lambda) \cdot \frac{1}{2}(\lambda H + (1-\lambda)L - L)\big)\Big) \\ = &\frac{1}{4}(H-L)x\big(\frac{1}{2} + 2\lambda(1-\lambda)\big) \end{aligned}$$

Solving for a mixed strategy equilibrium, we find that the informed trader will trade at $t = 0$ with probability x where

$$x = \frac{2\lambda - 1}{4\lambda - 2\lambda^2 - \frac{1}{2}}$$

□

Proposition 2 states that with two informed traders, the symmetric equilibrium is a mixed strategy where the informed traders find it optimal to trade prematurely on unverified information. In other words, instead of waiting for the information to be verified at $t = 1$, informed traders find it optimal to use a mixed strategy and trade early at $t = 0$ based on their signal. In our model, the extent of such premature trading depends on the level of accuracy of information (λ). This reflects the clear trade-off between taking an early action and postponing the action which can be quantified by the quality of information. Unlike the case of a single informed trader, informed traders are no longer able to play a waiting game since the other informed trader can potentially usurp the benefits from the information by trading early at $t = 0$.

Our results show that the extent of premature trading is a monotonically increasing in λ, *i.e.* $\frac{\partial x}{\partial \lambda} > 0$ for $\lambda \in (0.5, 1]$. This is along expected lines. As λ increases, the initial signal becomes more and more accurate and the benefits from trading late falls (as the gains from verification are lower) and thus x is larger. However, note that even if the $\lambda = 1$ and the incentive to trade late disappears, in order to profit from the private signal, the informed traders still have to utilize a mixed strategy. This is because if both informed traders trade at $t = 0$, then the private signal is revealed to the market maker.

It is clear from Propositions 19 and 20 that competition results in premature trading. In case of only one informed trader, the informed trader finds it optimal to wait and trades on verified information at $t = 1$. However, with competition, the informed traders play a mixed strategy and with probability x trades prematurely at $t = 0$.

Revisiting our brokerage house example from earlier, we can think of a more realistic case of two brokerage houses getting involved in the deal. As before, the initial information available to both the brokerage houses is noisy. However, unlike the one trader case where the brokerage house has the luxury of waiting for the details of the deal to emerge, competition incentivises the brokerage house to act on the information quickly before its competitor gets a chance to do so. The brokerage houses face a clear tradeoff. As before they can wait for the details to emerge and ensure that they trade on correct information, *i.e.* verify that the deal would actually materialise and confirm its terms. However, now this waiting entails the risk of the other brokerage house would act first, thereby cashing in on the opportunity. Therefore, competition incentivises the brokerage house to take on more risk and go for it, *i.e.* trade on the information before it can verify all the details. This premature trading behaviour is captured in our model by the mixed strategy symmetric equilibrium x.

This finding is compatible with the aggressive trading result found in Admati and Pfleiderer (1988) and Holden and Subrahmanyam (1992) in Kyle type models. However, the novelty of our approach lies in the fact that our results are driven by the information structure of the signal. This allows us to model phenomena such as high volatility and aggressive trading from the perspective of competition among informed traders triggered by imperfect information.

Note that in deriving our results, we have focussed on the symmetric equilibrium and ignored the possibility that the two informed traders might be able to coordinate their actions. In other words, there exists non-symmetric equilibria where one trader always trades early and the other always trades late. However, since it is difficult to justify the coordination between traders *ex-ante*, particularly when we move to a more generalised model with N traders, following other papers in the literature (e.g. Brunnermeier and Pedersen, 2005), we choose to focus on the *symmetric* equilibrium.

4.3 N informed trader

We can now generalise our model to $N > 2$ informed traders, each receiving a common signal at $t = 0$. This allows us to look at the impact of increasing competition on the trading behaviour of market participants.

Proposition 21. *There exists a unique symmetric PBE, such that the informed traders will*

- *with probability x, buy d assets at $t = 0$ if they receive a positive signal at $t = 0$*
- *with probability $(1 - x)$, buy d assets at $t = 1$ if the positive signal is confirmed, else sell d assets*
- *with probability x, sell d assets at $t = 0$ if they receive a negative signal at $t = 0$*
- *with probability $(1 - x)$, sell d assets at $t = 1$ if the negative signal is confirmed, else buy d assets*

where

$$x = 1 - \frac{1}{1+a} \ , \ and \ a = \left(\frac{2\lambda - 1}{4\lambda(1-\lambda)}\right)^{\frac{1}{N-1}}$$

Proof. The proof is analogous to Proposition 20. We start with prices at $t = 0$,

$$P_0 = \begin{cases} \lambda H + (1-\lambda)L & \text{if the market maker (MM) observes at least two positive orders} \\ \lambda L + (1-\lambda)H & \text{if the MM observes at least two negative orders} \\ \bar{V} & \text{otherwise} \end{cases}$$

At $t = 1$, once again the price depends on the history at $t = 0$. We get the following cases,

- If the MM observed at least two positive orders at $t = 0$

$$P_1 = \begin{cases} H & \text{if MM observes at least two positive orders at } t = 1 \\ L & \text{if MM observes at least two negative orders at } t = 1 \\ P_0 = \lambda H + (1-\lambda)L & \text{otherwise} \end{cases}$$

- If the MM observed at least two negative orders at $t = 0$

$$P_1 = \begin{cases} H & \text{if MM observes at least two positive orders at } t = 1 \\ L & \text{if MM observes at least two negative orders at } t = 1 \\ P_0 = \lambda L + (1-\lambda)H & \text{otherwise} \end{cases}$$

- Otherwise

$$P_1 = \begin{cases} H & \text{if MM observes at least two positive orders at } t = 1 \\ L & \text{if MM observes at least two negative orders at } t = 1 \\ P_0 = \bar{V} & \text{otherwise} \end{cases}$$

Given the market maker's beliefs (and the resulting prices), we evaluate the informed traders strategy at $t = 0$. Assuming that the informed trader gets a high signal ($V = H$)

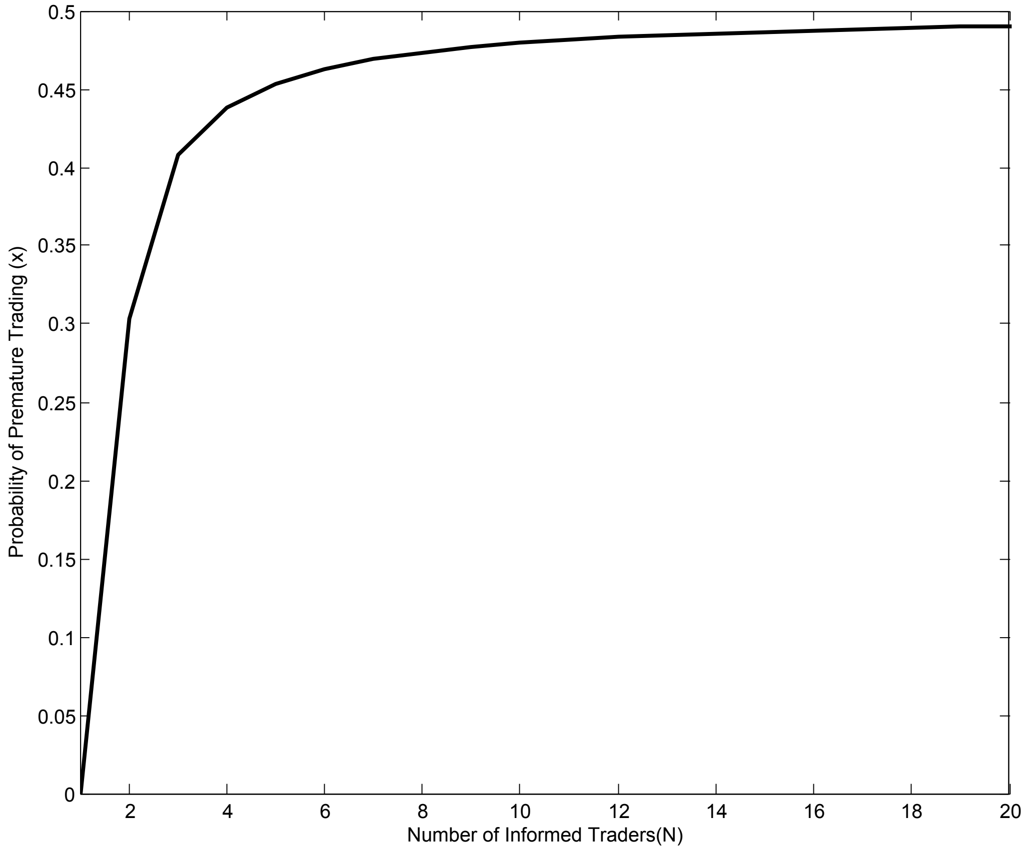


Figure 1: Premature trading increases with competition ($\lambda = 0.7$)

at $t = 0$, his profit from trading at $t = 0$,

$$\begin{aligned}\Pi_0 &= (1-x)^{N-1} \cdot \frac{1}{2} \cdot \big((\lambda H + (1-\lambda)L - \frac{1}{2}(H+L)\big) \\ &= \frac{1}{2}(1-x)^{N-1}(\lambda - \frac{1}{2})(H-L)\end{aligned}$$

On the other hand, if he waits to verify his signal and trades at $t = 1$,

$$\Pi_1 = x^{N-1}\lambda(1-\lambda)(H-L)$$

Hence we have

$$x = 1 - \frac{1}{1+a}$$

where $a = \left(\frac{2\lambda-1}{4\lambda(1-\lambda)}\right)^{\frac{1}{N-1}}$ □

The main point to note here is that x is increasing in N as long as the initial infor-

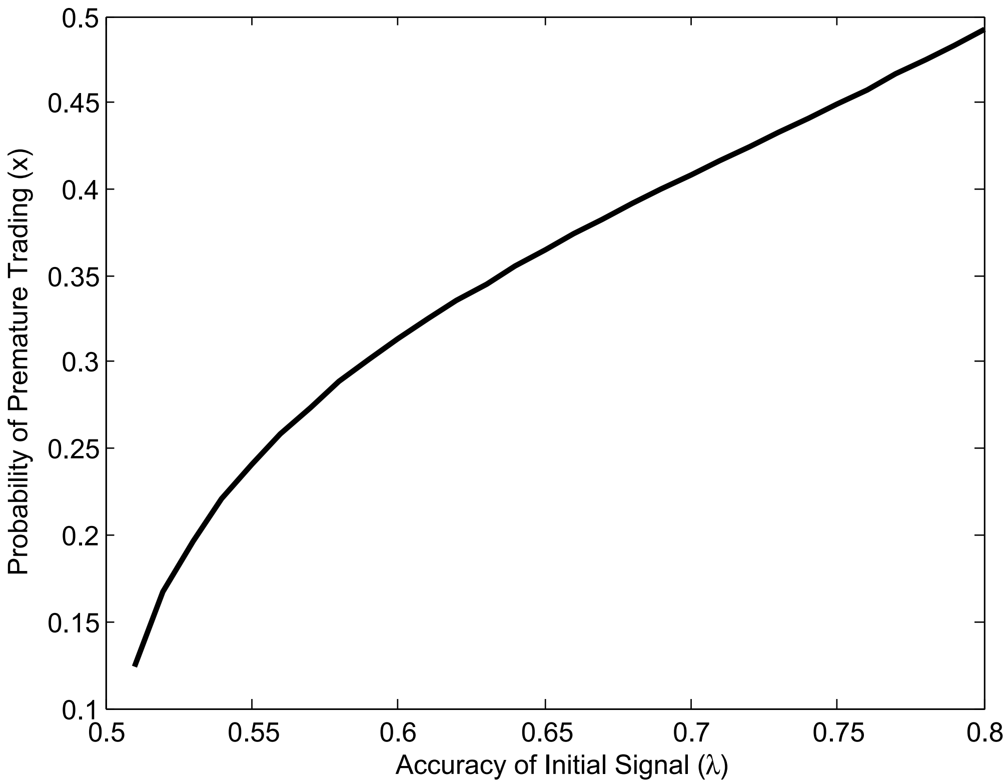


Figure 2: Premature trading increases with accuracy of signal ($N = 3$)

mation is not too accurate, *i.e.* $\lambda < \frac{1+\sqrt{5}}{4}$. Figure 1 shows that the premature trading increases with the intensity of competition. Intuitively, as the number of informed agents increases, the probability that one of them will trade early also increases, increasing the potential loss from trading later at $t = 1$. On the other hand, the benefits, in the form of verified signal, remain the same. Thus, premature trading rises with increased competition.

Also note that x converge towards $\frac{1}{2}$ as N goes to infinity. This is because $x = \frac{1}{2}$ minimises the probability that informed agents trade at the same time.[3] Figure 2 shows that premature trading increases with the accuracy of the signal. This is along expected lines since the benefits from trading late diminish as the accuracy of the signal increases.

[3] $\frac{1}{2}$ achieves maximum diversification and would be the solution when there are no signal verification benefits. However, in our case, since trading late is beneficial, as $N \to \infty$, the optimal strategy converges to $\frac{1}{2}$, *i.e.* $x \to \frac{1}{2}$, but is strictly below.

5 Implications

Propositions 19, 20 and 21 outline the trading behaviour of informed traders under increasing levels of competition. Our results suggest that premature trading rises with competition as informed traders find it optimal to trade prematurely. In this section, we look at the implications of this result across three dimensions - the expected profits of informed traders;[4] the impact of premature trading on market volatility and price discovery; and finally the behaviour of informed traders when we introduce market closing and the window in which they can profitably trade diminishes.

5.1 Expected Profits

Using Propositions 19, 20 and 21, we can calculate the total expected profits of the informed traders.

Proposition 22. *Total expected profits are decreasing in N*

Proof. We assume that the informed trader(s) receives a positive signal (the negative signal case is analogous). Let $E(\Pi^i)$ be the expected profit of informed traders when the number of informed traders is i.

One informed trader Proposition 19 implies that the informed trader will only trade at $t = 2$. Furthermore, he makes a profit only when the noise trader trades against him. Expected profit,

$$E(\Pi^1) = \frac{1}{2}(H - \bar{V}) = \frac{1}{4}(H - L)$$

Two informed traders With two informed traders, each trader has a mixed strategy symmetric equilibrium and trades at $t = 0$ with probability x (Proposition 20). The informed traders profit only when they trade in different periods and the noise

[4]Note that the market maker always makes zero profit in equilibrium and the noise trader trades randomly.

trader trades against them. The expected profit becomes,

$$
\begin{aligned}
E(\Pi^2) =& x^2 \cdot 0 + (1-x)^2 \cdot 0 + 2x(1-x)\Big(\frac{1}{2}(\lambda - \frac{1}{2})(H-L) + \frac{1}{4}(H-L)\big(\frac{1}{2} + 2\lambda(1-\lambda)\big)\Big) \\
=& x(1-x)(H-L)\big(2\lambda - \lambda^2 - \frac{1}{4}\big)
\end{aligned}
$$

N **informed traders** Finally, using Proposition 21, with N informed traders, expected profit,

$$
\begin{aligned}
E(\Pi^N) =& \frac{1}{2}x(1-x)^{N-1}(\lambda - \frac{1}{2})(H-L) + (1-x)x^{N-1}\lambda(1-\lambda)(H-L) \\
=& x(1-x)(H-L)\big(\frac{1}{2}(1-x)^{N-2}(\lambda - \frac{1}{2}) + x^{N-2}\lambda(1-\lambda)\big).
\end{aligned}
$$

Since by assumption,

$$
\lambda \in (0.5, 1], \lambda - \lambda^2 \leq \frac{1}{4} \implies 2\lambda - \lambda^2 - \frac{1}{4} \leq 1 \implies x(1-x)\big(2\lambda - \lambda^2 - \frac{1}{4}\big) \leq \frac{1}{4}
$$

Since $E(\Pi^2) \leq E(\Pi^1)$, the total expected profit with two informed traders is less than that with one informed trader.

Moreover, the total expected profit with three informed traders is

$$
\begin{aligned}
& x(1-x)(H-L)\big(\frac{1}{2}(\lambda - \frac{1}{2}) + \lambda(1-\lambda)\big) \\
&= x(1-x)(H-L)\big(\frac{3}{2}\lambda - \lambda^2 - \frac{1}{4}\big) \\
&< x(1-x)(H-L)\big(2\lambda - \lambda^2 - \frac{1}{4}\big)
\end{aligned}
$$

Since $E(\Pi^3) \leq E(\Pi^2)$, the total expected profits with three informed traders are less than that with two informed trader.

Finally, $x(1-x)(H-L)\big(\frac{1}{2}(1-x)^{N-2}(\lambda - \frac{1}{2}) + x^{N-2}\lambda(1-\lambda)\big)$ is decreasing in N. Therefore the total expected profits are decreasing with the number of traders. □

Proposition 22 and Figure 3 clearly show that total expected profits are decreasing in N. Propositions 19, 20 and 21 show that premature trading increases with competition. As competition rises, informed traders, fearful of losing out on the trading opportunity, trade prematurely on unverified information. This leads to lower profits in equilibrium.

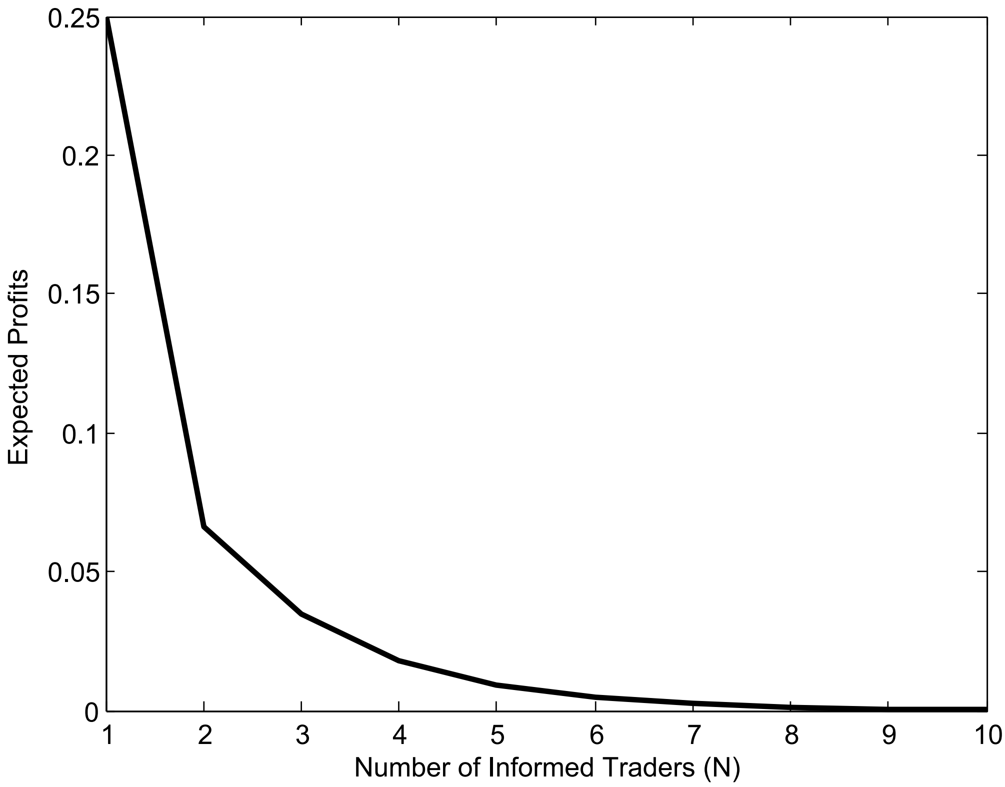


Figure 3: Expected profits are decreasing in N ($\lambda = 0.7$)

Furthermore, as competition increases, traders find it harder and harder to conceal their identity in equilibrium. Thus the market maker can learn the informed traders information more easily and set the price accordingly.

This result is compatible with the semi-strong form Efficient Market Hypothesis (EMH), which suggests that market participants, trading on public information, make zero profit in equilibrium. Intuitively, when the number of informed traders (N) is small, the assumption of semi-strong form market efficiency is violated and market participants make positive profits in expectation. However as N increases, we get closer to the assumptions of semi-strong form market efficiency as the informed traders' private signal is available to larger number of market participants and as a result they can no longer profit from their information.

5.2 Excess Volatility

Section 5.1 shows us that total expected profits decline with competition. This is along expected lines. However, the more interesting consequence of increasing competition amongst market participants and the resultant premature trading is its impact on market volatility. Here we show that increased competition leads to higher market volatility.

In the absence of competitive pressure, the informed trader would wait to verify this signal and trade only at the end of the period. At this time, the informed trader would have perfect information. On the other hand, with competition, the informed trader, fearful of losing out on the trading opportunity, might decide to trade on the unverified information prematurely at the beginning of the period. In doing so, the informed trader would run the risk of getting the trade wrong in case the initial signal turns out to be incorrect. This in turn would lead to higher market volatility.

Proposition 23. *Competition increases market volatility.*

Proof. **In the absence of premature trading, *i.e.* one informed trader case**, since the informed trader always trades on verified information at $t = 1$, there is no price movement at $t = 0$. When the informed trader trades at $t = 1$, the price movement depends upon the actions of the noise trader. If the noise trader trades against the informed trader, there is no price movement at $t = 1$. On the other hand, if the noise trader trades in the same direction as the informed trader, price will move at $t = 1$. Therefore price volatility would be

$$\begin{aligned} E\big(Var(P)\big) &= E\big((P-\bar{V})^2\big) \\ &= \frac{1}{2}(H-\bar{V})^2 = \frac{1}{8}(H-L)^2 \end{aligned}$$

In the presence of premature trading, *i.e.* at least two informed traders, price will move at $t = 0$ with probability $x(1-x)+x^2 = x$, and will not move with probability $1-x$. The price may then move to H or L depending on whether the signal is correct or not. Therefore the volatility of the return will be

$$\begin{aligned} E\left[Var(P)\right] = {} & (1-x)^N\frac{1}{2}(H-\bar{V})^2 + x^N\frac{1}{2}(\lambda H+(1-\lambda)L-\bar{V})^2 \\ & + Nx(1-x)^{N-1}\left\{\frac{1}{4}(H-\bar{V})^2+\frac{1}{4}\left[(\lambda H+(1-\lambda)L-\bar{V})^2+(H-\bar{V})^2\right]\right\} \\ & + Nx^{N-1}(1-x)\left\{\frac{1}{4}(\lambda H+(1-\lambda)L-\bar{V})^2+\frac{1}{4}\left[(\lambda H+(1-\lambda)L-\bar{V})^2+(H-\bar{V})^2\right]\right\} \\ & + \Big\{\left[1-(1-x)^N-Nx(1-x)^{N-1}-Nx^{N-1}(1-x)-x^N\right] \end{aligned}$$

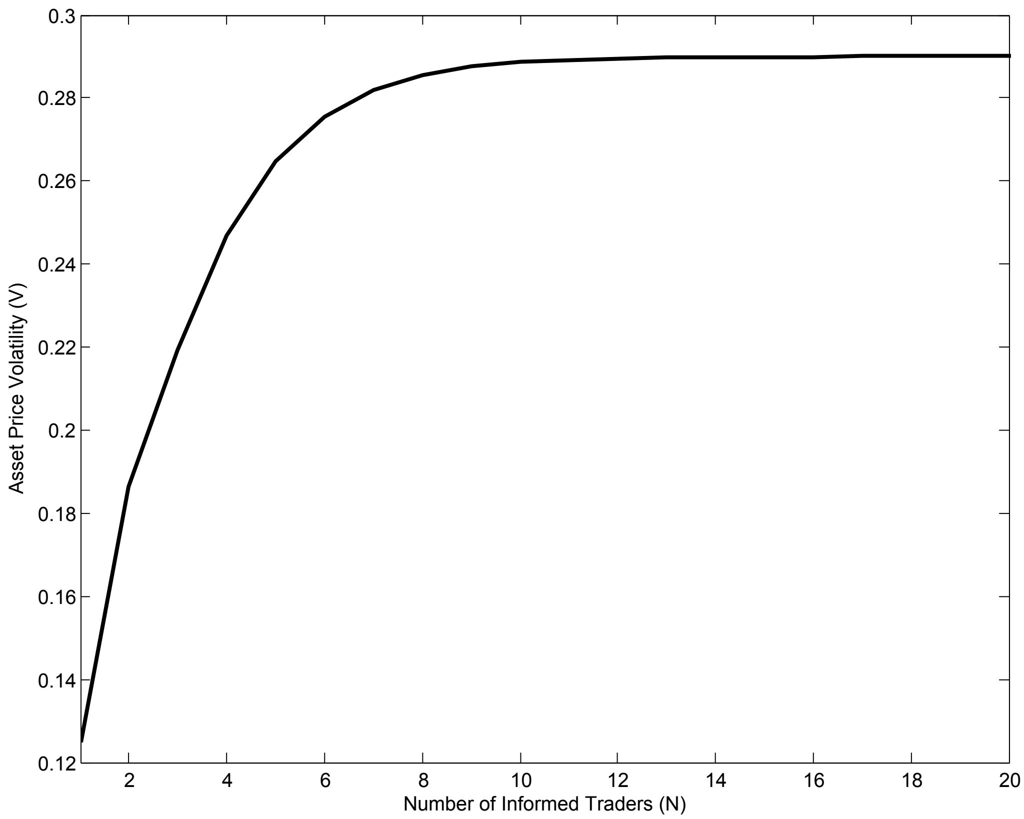


Figure 4: Volatility is increasing in N ($\lambda = 0.7$)

$$
\begin{aligned}
&\times \frac{1}{2}\left[(\lambda H + (1-\lambda)L - \bar{V})^2 + (H - \bar{V})^2\right]\Big\} \\
&= \frac{1}{2}\left\{\left[1 - (1-x)^N - \frac{1}{2}Nx(1-x)^{N-1}\right](\lambda H + (1-\lambda)L - \bar{V})^2\right. \\
&\qquad \left. + \left[1 - x^N - \frac{1}{2}Nx^{N-1}(1-x)\right](H - \bar{V})^2\right\} \\
&= \frac{1}{2}(H-L)^2\left\{\left[1 - (1-x)^N - \frac{1}{2}Nx(1-x)^{N-1}\right](\lambda - \frac{1}{2})^2\right. \\
&\qquad \left. + \left[1 - x^N - \frac{1}{2}Nx^{N-1}(1-x)\right]\frac{1}{4}\right\}
\end{aligned}
$$

Clearly, $E\big(Var(P)\big)$ is increasing in N.

$\square$

Figure 4 shows that market volatility increases with competition. Intuitively, as the number of informed traders increases, each trader, fearful of losing out on the trading opportunity finds it optimal to use a mixed strategy and trade prematurely with increasing probability. This increase in premature trading in turn leads to incorrect trading and higher market volatility, hindering price discovery.

5.3 Spike in the market closing volatility

It is widely documented that returns volatility varies systemically over the trading day, with Wood, McInish, and Ord (1985) and Harris (1986) documenting the existence of a distinct 'U-shaped' pattern in return volatility over the trading day. In particular, these papers document spikes in market volatility as the market closing time approaches.

Our results suggest that the spike in volatility at the closing time of the market can be modelled as a direct outcome of premature trading. Intuitively, as the market closing time draws near, informed traders tend to overreact further in fear of not being able to use the information at hand as the window over which they can profitably trade narrows. This is because if informed traders wait till the next trading day, the information is revealed to the wider market and ceases to be a profitable trading opportunity. As a result, informed traders tend to trade even more on unverified and noisy initial information, resulting in the spike in market volatility during the end of the trading day. Note that our approach to the spike in the volatility at the market closing via premature trading is different from that of Admati and Pfleiderer (1988), which approaches this phenomenon via concentrated trading arising from the strategic behaviour of noisy traders.

Formally, we extend our three-period model with two traders outlined in Section 4.2 by allowing for uncertainty of timing with respect to the arrival of unverified signal. In particular, we now assume that the signal can arrive at $t = 0$ or $t = 1$ with equal probability (as opposed to $t = 0$ with certainty in Section 4.2). The arrival of signal itself is common knowledge, but only the informed traders can observe the signal.

As before, the signal is initially correct with probability λ and it takes one period for informed traders to verify the signal. The asset pays off two periods after the signal is received. We assume that the trading day starts at $t = 0$ and the market closes at $t = 2$. In this setting, when the signal arrives at $t = 0$ (and can therefore be verified at $t = 1$), the model remains the same as in Section 4.2. However, when the signal arrives at $t = 1$, the signal can only be verified at $t = 2$ when the market has closed. Therefore, the informed traders have to decide whether to trade on the information immediately or give up on the trading opportunity. This is because if they wait until the next trading day to verify the signal, the information is revealed to the wider market overnight and therefore ceases to be a profitable trading opportunity. This allows us to study the behaviour of market participants when their window of profitable trading is curtailed.

Proposition 24. *There exists a unique symmetric PBE, such that the informed traders will*

- *act exactly as describe in Proposition 20 if they receive a signal at* $t = 0$
- *trade according to the signal with probability* 1 *if they receive a signal at* $t = 1$

Proof. If the signal arrives at $t = 0$, the proof is identical to that of Proposition 20. On the other hand, if the signal arrives at $t = 1$, it is straightforward to show that the only symmetric equilibrium is to trade immediately on the unverified information, given that the informed traders cannot profit at all at $t = 2$ when the information is revealed. □

Since there is no volatility in the absence of price movement, we focus on the premature trading with price impact. Because the signal arrives at $t = 0$ with 1/2 probability, the probability of premature trading with price impact at $t = 0$ is $\frac{1}{2}[x(1-x)+x^2] = \frac{x}{2}$. On the other hand, the probability of premature trading with price impact at $t = 1$ is $\frac{1}{2}$. Therefore, when the arrival time of the signal is uncertain, traders are more likely to trade prematurely based on a wrong signal at the market closing, resulting in a spike in volatility.

6 Conclusion

We develop a model of trading under competition. In our setup, informed traders receive an initial noisy signal, which they need to verify. They can do this costlessly by waiting for one period. Our model shows that with competition, informed traders, fearful of losing out on the trading opportunity, may choose not to verify their signal and instead trade prematurely. This premature trading takes place in a completely rational REE framework and is a direct result of competition. Our results show that this premature trading increases with competition.

We go on to look at the impact of such premature trading. We show that not only does this lead to lower overall profits in equilibrium, but also results in higher market volatility. We believe that our model can shed some light on the widely documented excess volatility puzzle in financial markets. Our model also suggests that the strong evidence suggesting that market volatility spikes as the market closing time approaches may also be a consequence of premature trading.

References

Abreu, D., and M. Brunnermeier, 2003, "Bubbles and crashes," *Econometrica*, 71(1), 173–204.

Admati, A., and P. Pfleiderer, 1988, "A theory of intraday patterns: Volume and price variability," *Review of Financial Studies*, 1(1), 3.

Aitken, M., P. Brown, C. Buckland, H. Izan, and T. Walter, 1996, "Price clustering on the Australian stock exchange," *Pacific-Basin Finance Journal*, 4(2-3), 297–314.

Andersen, T., T. Bollerslev, and J. Cai, 2000, "Intraday and interday volatility in the Japanese stock market," *Journal of International Financial Markets, Institutions and Money*, 10(2), 107–130.

Andersen, T., T. Bollerslev, F. Diebold, and H. Ebens, 2001, "The distribution of realized stock return volatility* 1," *Journal of Financial Economics*, 61(1), 43–76.

Avramov, D., T. Chordia, and A. Goyal, 2006, "The impact of trades on daily volatility," *Review of Financial Studies*, 19(4), 1241.

Bagehot, W., 1971, "The Only Game in Town," *Financial Analysts Journal*, pp. 12–22.

Bagnoli, M., S. Viswanathan, and C. Holden, 2001, "On the existence of linear equilibria in models of market making," *Mathematical Finance*, 11(1), 1–31.

Barron, O. E., and J. Karpoff, 2004, "Information precision, transaction costs, and trading volume," *Journal of Banking & Finance*, 28, 1207–1223.

Barsky, R., and J. De Long, 1993, "Why does the stock market fluctuate?," *The Quarterly Journal of Economics*, 108(2), 291.

Biais, B., and J. Rochet, 1997, "Risk sharing, adverse selection and market structure," *Financial mathematics*, pp. 1–51.

Black, F., 1976, "Studies of Stock Price Volatility Changes. Proceedings of the 1976 Meetings of the American Statistical Association," *Business and Economical Statistics Section*, pp. 177–181.

Brock, W., and A. Kleidon, 1992, "Periodic market closure and trading volume:: A model of intraday bids and asks," *Journal of Economic Dynamics and Control*, 16(3-4), 451–489.

Brunnermeier, M., and L. H. Pedersen, 2005, "Predatory Trading," *Journal of Finance*, 60(4), 1825–1863.

Bulkley, G., and R. Harris, 1997, "Irrational analysts' expectations as a cause of excess volatility in stock prices," *The Economic Journal*, 107(441), 359–371.

Campbell, J., and L. Hentschel, 1992, "No news is good news:: An asymmetric model of changing volatility in stock returns," *Journal of financial Economics*, 31(3), 281–318.

Campbell, J. Y., and A. S. Kyle, 1993, "Smart Money, Noise Trading and Stock Price Behaviour," *The Review of Economic Studies*, 60(1), 1–34.

Chan, K., and W. Fong, 2000, "Trade size, order imbalance, and the volatility-volume relation* 1," *Journal of Financial Economics*, 57(2), 247–273.

Chen, Y., and H. Zhao, 2012, "Informed trading, information uncertainty, and price momentum," *Journal of Banking & Finance*, 36, 2095–2019.

Choe, H., and H. Sik Shin, 1993, "An analysis of interday and intraday return volatility-evidence from the Korea Stock Exchange," *Pacific-Basin Finance Journal*, 1(2), 175–188.

Corb, H., 1993, "Informed speculation and asset prices when the order flow is disaggregated," working paper, Working paper, Stanford University.

Cutler, D., J. Poterba, and L. Summers, 1989, "What moves stock prices?," *Journal of Portfolio Management*, 15, 3.

——— , 1990, "Speculative Dynamics and the Role of Feedback Traders," *The American Economic Review*.

De Bondt, W., and R. Thaler, 1985, "Does the stock market overreact?," *Journal of finance*, pp. 793–805.

De Long, J., A. Shleifer, L. Summers, and R. Waldmann, 1990, "Noise Trader Risk in Financial Markets," *Journal of Political Economy*, 98(4).

de Long, J., A. Shleifer, L. Summers, and R. Waldmann, 1990, "Positive Feedback Investment Strategies and Destabilizing Rational Speculation," *The Journal of Finance*, 45(2), 379–395.

Du, J., and S. Wei, 2004, "Does Insider Trading Raise Market Volatility?," *The Economic Journal*, 114(498), 916–942.

Easley, D., and M. O'Hara, 1987, "Price, trade size, and information in securities markets," *Journal of Financial economics*, 19(1), 69–90.

Ebrahim, M., and I. Mathur, 2001, "Investor heterogeneity, market segmentation, leverage and the equity premium puzzle," *Journal of Banking & Finance*, 25(10), 1897–1919.

Epstein, L., and M. Schneider, 2008, "Ambiguity, information quality, and asset pricing," *Journal of Finance*, 62, 197–228.

Foster, F., and S. Viswanathan, 1990, "A theory of the interday variations in volume, variance, and trading costs in securities markets," *Review of Financial Studies*, 3(4), 593–624.

French, K., and R. Roll, 1986, "Stock return variances: The arrival of information and the reaction of traders," *Journal of Financial Economics*, 17(1), 5–26.

French, K., G. Schwert, and R. Stambaugh, 1987, "Expected stock returns and volatility," *Journal of Financial Economics*, 19, 3–29.

Friedman, M., 1953, *Essays in positive economics*. University of Chicago Press.

Froot, K., D. Scharftstein, and J. Stein, 1992, "Herd on the Street: Informational Inefficiencies in a Market with Short-Term Speculation," *Journal of Finance*, 47(4), 1461–84.

Gilles, C., and S. LeRoy, 1991, "Econometric aspects of the variance-bounds tests: A survey," *Review of Financial Studies*, 4(4), 753.

Glosten, L., and P. Milgrom, 1985, "Bid, ask and transaction prices in a specialist market with heterogeneously informed traders," *Journal of financial economics*, 14(1), 71–100.

Harris, L., 1986, "A transaction data study of weekly and intradaily patterns in stock returns," *Journal of financial economics*, 16(1), 99–117.

Hellwig, M., 1980, "On the aggregation of information in competitive markets," *Journal of economic theory*, 22(3), 477–498.

Holden, C., and A. Subrahmanyam, 1992, "Long-lived private information and imperfect competition," *Journal of Finance*, pp. 247–270.

Jones, C., G. Kaul, and M. Lipson, 1994, "Transactions, volume, and volatility," *Review of Financial Studies*, 7(4), 631.

Kayahan, B., T. Stengos, and B. Saltoğlu, 2002, "Intra-day features of realized volatility: evidence from an emerging market," *International Journal of Business*, 1(1), 17–24.

Kleidon, A., 1986, "Variance bounds tests and stock price valuation models," *The Journal of Political Economy*, 94(5), 953–1001.

Kyle, A., 1985, "Continuous Auctions and Insider Trading," *Econometrica*, 53(6), 1315–35.

——— , 1989, "Informed speculation with imperfect competition," *The Review of Economic Studies*, 56(3), 317.

Kyle, A., and J. Vila, 1991, "Noise trading and takeovers," *The RAND Journal of Economics*, pp. 54–71.

LeRoy, S., and R. Porter, 1981, "The present-value relation: Tests based on implied variance bounds," *Econometrica: Journal of the Econometric Society*, pp. 555–574.

Li, T., 2007, "Heterogeneous beliefs, asset prices, and volatility in a pure exchange economy," *Journal of Economic Dynamics and Control*, 31(5), 1697–1727.

Lu, C., T. Chen, and H. Liao, 2010, "Information uncertainty, information asymmetry and corporate bond yield spreads," *Journal of Banking & Finance*, 34, 2265–2279.

Mankiw, N., D. Romer, and M. Shapiro, 1985, "An unbiased reexamination of stock market volatility," *The Journal of Finance*, 40(3), 677–687.

McInish, T., and R. Wood, 1990a, "An analysis of transactions data for the Toronto Stock Exchange: Return patterns and end-of-the-day effect," *Journal of Banking & Finance*, 14(2), 441–458.

——— , 1990b, "A transactions data analysis of the variability of common stock returns during 1980–1984," *Journal of Banking & Finance*, 14(1), 99–112.

Niemeyer, J., and P. Sandas, 1994, "An empirical analysis of the trading structure at the Stockholm Stock Exchange," *Journal of Multinational Finance Management*, 3(3-4), 63–101.

Rochet, J., and J. Vila, 1994, "Insider trading without normality," *The Review of Economic Studies*, 61(1), 131.

Roll, R., 1984, "Orange juice and weather," *The American Economic Review*, 74(5), 861–880.

——— , 1988, "Presidential address: R2," *Journal of Finance*, 43(3), 541–566.

Schwert, G., 1989, "Why does stock market volatility change over time?," *Journal of Finance*, 44, 1115–1153.

——— , 1990, "Stock Volatility and the Crash of'87," *Review of Financial studies*, 3(1), 77.

Shiller, R. J., 1981, "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?," *American Economic Review*, 71(3), 421–36.

Shleifer, A., and L. Summers, 1990, "The noise trader approach to finance," *The Journal of Economic Perspectives*, 4(2), 19–33.

Slezak, S., 1994, "A theory of the dynamics of security returns around market closures," *Journal of Finance*, pp. 1163–1211.

Spiegel, M., and A. Subrahmanyam, 1992, "Informed speculation and hedging in a non-competitive securities market," *Review of Financial Studies*, 5(2), 307–329.

——— , 1995, "On intraday risk premia," *Journal of Finance*, pp. 319–339.

Stoll, H., and R. Whaley, 1990, "Stock market structure and volatility," *Review of Financial Studies*, 3(1), 37.

Timmermann, A., 1993, "How learning in financial markets generates excess volatility and predictability in stock prices," *The Quarterly Journal of Economics*, 108(4), 1135.

Vives, X., 1995, "The speed of information revelation in a financial market mechanism," *Journal of Economic Theory*, 67(1), 178–204.

Wang, J., 1993, "A model of intertemporal asset prices under asymmetric information," *The Review of Economic Studies*, 60(2), 249.

——— , 1994, "A model of competitive stock trading volume," *The Journal of Political Economy*, 102(1), 127–168.

Wang, M., and L. Zu, 2011, "Trade Timing, Price Volatility and Serial Correlation," *European Financial Management.*

West, K., 1988, "Dividend Innovations and Stock Price Volatility," *Econometrica*, 56(1), 37–61.

Wood, R., T. McInish, and J. Ord, 1985, "An investigation of transactions data for NYSE stocks," *Journal of Finance*, pp. 723–739.

Zhang, X., 2006, "Information uncertainty and stock returns," *Journal of Finance*, 61, 105–137.

A Robustness Check

Alternative Information Structure: Aggregate Order Flow

While the setup of our model closely resembles Kyle (1985), the information structure in our model differs slightly by assuming that the market maker observes every order, but cannot distinguish between noise and informed traders. In the Kyle (1985) model, the market maker only observes aggregate order flow. In this appendix, we check the robustness of our results by assuming an alternative Kyle (1985) type information structure with aggregate order flows.

The setup of our model remains the same as outlined in section 3. The only difference is that now the market maker only observes the aggregate order flow. In making this assumption, we restrict the information set of the market maker. For example, in our original setup, the market maker could distinguish between an order {d, d, -d} and {d}. However, with the aggregate order flow setting of the Kyle (1985) model, the market maker would observe the same aggregate demand, +d.

While this makes our model more complicated, our core result go through. Proposition 19 and its proof remains the same as the information structure is identical in the two cases. We therefore omit it from the appendix. The core result of Proposition 20, *premature trading*, continues to hold, but the informed traders strategy is more complicated. This is because now the market maker has less information and the informed traders information is revealed more slowly. Proposition 20 now becomes,

Proposition 25. *There exists a unique symmetric PBE, such that the informed traders will*

- *with probability x, buy d assets at $t = 0$ if they receive a positive signal at $t = 0$*
- *with probability $(1 - x)$, buy d assets at $t = 1$ if the positive signal is confirmed, else sell d assets*
- *with probability x, sell d assets at $t = 0$ if they receive a negative signal at $t = 0$*
- *with probability $(1 - x)$, sell d assets at $t = 1$ if the negative signal is confirmed, else buy d assets*

where x is the mixed strategy (symmetric) equilibrium defined by

$$\frac{(1-x)^2}{x^2+2(1-x)^2}(\lambda-\frac{1}{2})=\frac{x}{2}\left[\frac{-\lambda^2+\lambda}{x^2+(1-x)^2}+\frac{1}{4}\right]$$

Proof. Without loss of generality, assume that the informed traders receive a high signal S_H at $t=0$. The market maker observes aggregate demand and sets the price. For notational simplicity, let $V^H=\lambda H+(1-\lambda)L$ and $V^L=\lambda L+(1-\lambda)H$.

At $t=0$,

$$P_0=\begin{cases} V^H & \text{if} \quad D_0=3d \\ V^H & \text{if} \quad D_0=2d \\ \frac{x^2}{x^2+2(1-x^2)}V^H+\frac{2(1-x^2)}{x^2+2(1-x^2)}\bar{V} & \text{if} \quad D_0=d \\ \bar{V} & \text{if} \quad D_0=0 \\ \frac{x^2}{x^2+2(1-x^2)}V^L+\frac{2(1-x^2)}{x^2+2(1-x^2)}\bar{V} & \text{if} \quad D_0=-d \\ V^L & \text{if} \quad D_0=-2d \\ V^L & \text{if} \quad D_0=-3d \end{cases}$$

In setting the price at $t=1$, the market maker takes into account demand at $t=0$. However, the market maker's information set is now smaller.

- If $D_0=3d$

$$P_1=P_0=V^H$$

- If $D_0=2d$

$$P_1=\begin{cases} H & \text{if} \quad D_1=2d \\ P_0=V^H & \text{if} \quad D_1=0 \\ L & \text{if} \quad D_1=-2d \end{cases}$$

- If $D_0=0$

$$P_1=\begin{cases} H & \text{if} \quad D_1=2d \\ \bar{V} & \text{if} \quad D_1=0 \\ L & \text{if} \quad D_1=-2d \end{cases}$$

- If $D_0 = -2d$

$$P_1 = \begin{cases} H & \text{if} \quad D_1 = 2d \\ P_0 = V^L & \text{if} \quad D_1 = 0 \\ L & \text{if} \quad D_1 = -2d \end{cases}$$

- If $D_0 = d$

$$P_1 = \begin{cases} H & \text{if} \quad D_1 = 3d \\ H & \text{if} \quad D_1 = 2d \\ \frac{x^2}{x^2+(1-x^2)}V^H + \frac{(1-x^2)}{x^2+(1-x^2)}H & \text{if} \quad D_1 = d \\ P_0 = \frac{x^2}{x^2+2(1-x^2)}V^H + \frac{2(1-x^2)}{x^2+2(1-x^2)}\bar{V} & \text{if} \quad D_1 = 0 \\ \frac{x^2}{x^2+(1-x^2)}V^H + \frac{(1-x^2)}{x^2+(1-x^2)}L & \text{if} \quad D_1 = -d \\ L & \text{if} \quad D_1 = -2d \\ L & \text{if} \quad D_1 = -3d \end{cases}$$

- If $D_0 = -d$

$$P_1 = \begin{cases} H & \text{if} \quad D_1 = 3d \\ H & \text{if} \quad D_1 = 2d \\ \frac{x^2}{x^2+(1-x^2)}V^H + \frac{(1-x^2)}{x^2+(1-x^2)}H & \text{if} \quad D_1 = d \\ P_0 = \frac{x^2}{x^2+2(1-x^2)}V^H + \frac{2(1-x^2)}{x^2+2(1-x^2)}\bar{V} & \text{if} \quad D_1 = 0 \\ \frac{x^2}{x^2+(1-x^2)}V^H + \frac{(1-x^2)}{x^2+(1-x^2)}L & \text{if} \quad D_1 = -d \\ L & \text{if} \quad D_1 = -2d \\ L & \text{if} \quad D_1 = -3d \end{cases}$$

- If $D_0 = -3d$

$$P_1 = P_0 = \lambda L + (1-\lambda)H$$

Given the market maker's beliefs (and the resulting prices), we evaluate the informed traders profits at $t=0$ and $t=1$.

$$\begin{aligned}
\Pi_0 &= x\frac{1}{2}\left(V^H - \left[\frac{x^2}{x^2+2(1-x)^2}V^H + \frac{2(1-x)^2}{x^2+2(1-x)^2}\bar{V}\right]\right) \\
&\quad +(1-x)\frac{1}{2}(V^H-\bar{V})
\end{aligned}$$

Using the expression for P_0, it is clear that the informed trader makes non-zero profit in only two cases - when the other informed trader also trades in the first period, the noise trader trades against both the informed traders ($D_0 = d$); and when the other informed trader trades at $t=1$ and the noise trader trades against the informed trader ($D_0 = 0$). In all other cases, the informed trader makes 0 profit.

$$\begin{aligned}
\Pi_1 &= x\lambda\frac{1}{2}\left(H - \frac{1}{2}\left(\bar{V}+V^H\right)\right) \\
&\quad +x\left(1-\lambda\right)\frac{1}{2}\left(\frac{1}{2}\left(\bar{V}+V^H\right) - L\right) \\
&\quad +\left(1-x\right)\lambda\frac{1}{2}\left(H - \left[\frac{x^2}{x^2+\left(1-x\right)^2}V^H + \frac{\left(1-x\right)^2}{x^2+\left(1-x\right)^2}H\right]\right) \\
&\quad +\left(1-x\right)\left(1-\lambda\right)\frac{1}{2}\left(\left[\frac{x^2}{x^2+\left(1-x\right)^2}V^H + \frac{\left(1-x\right)^2}{x^2+\left(1-x\right)^2}L\right] - L\right)
\end{aligned}$$

Here the first line represents the case where the signal turns out to be correct, the other informed trader trades in the first period and the noise trader trades against the informed trader of interest in the second period. The second line represents the case where everything is the same except that the information turns out to be wrong in the second period. The third line represents the case where the information is correct, the other informed trader trades in the second period and the noise trader trades against both the informed traders and the fourth line represents the case where everything is the same except that the information turns out to be wrong in the second place. Π_0 can be simplified as

$$\Pi_0 = \frac{\left(1-x\right)^2}{x^2+2(1-x)^2}(\lambda - \frac{1}{2})\left(H-L\right),$$

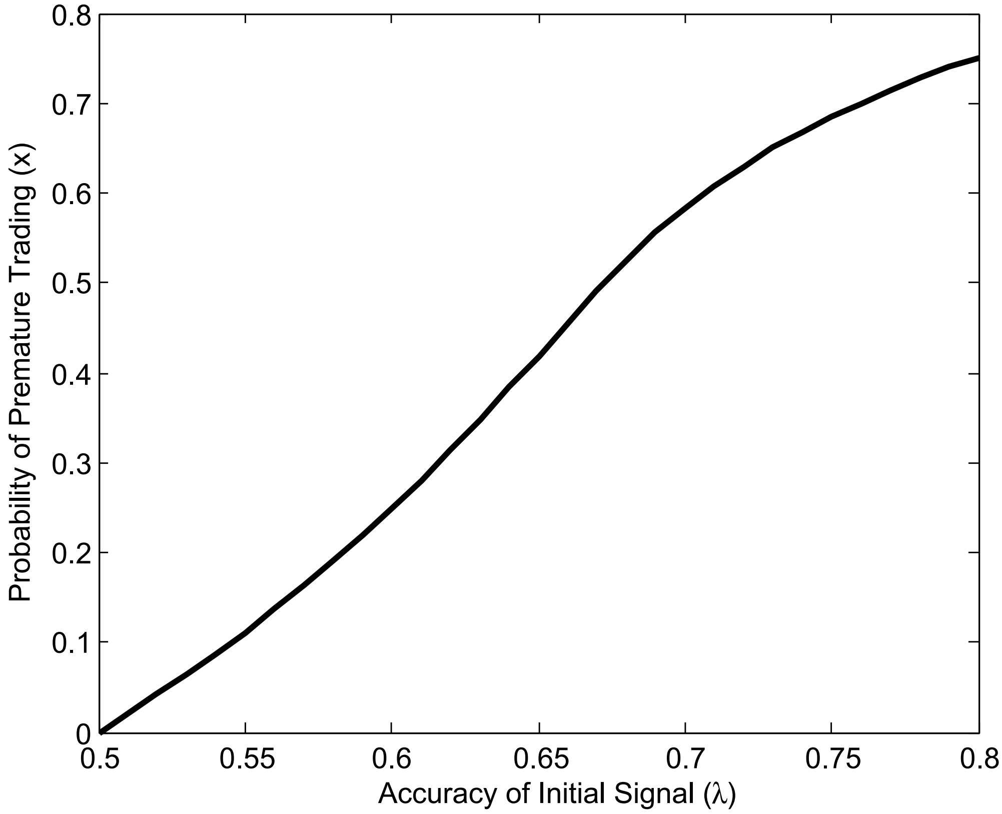


Figure 5: Numerical solution for different values of λ

Also, Π_1 can be simplified as

$$\Pi_1 = \frac{x}{2}\left[\frac{-\lambda^2+\lambda}{x^2+(1-x)^2}+\frac{1}{4}\right](H-L).$$

Since the equilibrium condition, $\Pi_0 = \Pi_1$ leads to the following high order polynomial equation (1), the closed form solution for the mixed strategy, x is not obtainable. Rather, we turn to seeking for the numerical solution.

$$\frac{(1-x)^2}{x^2+2(1-x)^2}(\lambda-\frac{1}{2}) = \frac{x}{2}\left[\frac{-\lambda^2+\lambda}{x^2+(1-x)^2}+\frac{1}{4}\right] \quad (1)$$

□

Figure 5 shows that even under the new set of assumptions, once competition is introduced, informed traders find it optimal to trade early. Thus our main result, *premature trading*, continues to hold and as before, premature trading increases with λ. The intuition behind this result remains the same, however note that x no longer converges to $\frac{1}{2}$.

This is because under these new set of assumptions, the market maker's information set is smaller and thus informed traders gain more from trading early as it is less likely that the market maker will be able to identify them.