The Seeds of a Crisis: A Theory of Bank Liquidity and Risk-Taking over the Business Cycle[1]

Viral Acharya[2]
New York University Stern School of Business, CEPR, NBER and ECGI.

Hassan Naqvi[3]
National University of Singapore and NUS Risk Management Institute.

May 3, 2010

[1]We thank Anjan Thakor (discussant) for useful suggestions. Comments by conference and seminar participants at the 2010 International Research Forum on Monetary Policy, 2010 AEA Annual Meeting, CEMFI, Lancaster University Management School, and National University of Singapore are also appreciated. We also acknowledge Hanh Le and Michelle Zemel for research assistance.

[2]Contact: Department of Finance, Stern School of Business, New York University, 44 West 4th Street, Room 9-84, New York, NY 10012. Tel: +1 212 998 0354. Fax: +1 212 995 4256. e-mail: vacharya@stern.nyu.edu

[3]Contact: Department of Finance, NUS Business School, Mochtar Riady Building, BIZ 1 #7-62, 15 Kent Ridge Drive, National University of Singapore, Singapore 119245. Tel: +65 6516 5552. Fax: +65 6779 2083. e-mail: naqvi@nus.edu.sg

Abstract

We examine how the banking sector may ignite the formation of asset price bubbles when there is access to abundant liquidity. Inside banks, given lack of observability of effort, loan officers (or risk takers) are compensated based on the volume of loans but are penalized if banks suffer a high enough liquidity shortfall. Outside banks, when there is heightened macroeconomic risk, investors reduce direct investment and hold more bank deposits. This 'flight to quality' leaves banks flush with liquidity, lowering the sensitivity of bankers' payoffs to downside risks of loans and inducing excessive credit volume and asset price bubbles. The seeds of a crisis are thus sown. We show that the optimal monetary policy involves a "leaning against liquidity" approach: A Central Bank should adopt a contractionary monetary policy in times of excessive bank liquidity in order to curb risk-taking incentives at banks, and conversely, follow an expansionary monetary policy in times of scarce liquidity so as to boost investment.

JEL Classifications: E32, E52, E58, G21

Keywords: Bubbles, flight to quality, Greenspan put, leaning against liquidity, leaning against the wind, monetary policy, moral hazard

"*For too long, the debate has got sidetracked. Into whether we can rely on monetary policy 'mopping up' after bubbles burst. Or into whether monetary policy could be used to control asset prices as well as doing its orthodox job of steering nominal trends in the economy...*" - Paul Tucker, Executive Director for Markets and Monetary Policy Committee (MPC) member at the Bank of England.[1]

1 Introduction

It is clear that in the period leading up to the global financial crisis of 2007-2009, credit and asset prices were growing at a ferocious pace.[2] In the United States, for example, in the five-year period from 2002 to 2007, the ratio of debt to national income went up from 3.75 to one, to 4.75 to one. It had taken the prior full decade to accomplish this feat, and fifteen years prior to that. During this same period, house prices grew at an unprecedented rate of 11% per year while there was no evidence of appreciating borrower quality over this period. The credit growth was across board, in mortgages, especially sub-prime ones, in financing of leveraged buy-out transactions (LBOs), and through increasing issuance of low-rated bonds. This rapid rise in asset volume and prices met with a precipitous fall. For instance, the median house price divided by rent in the United States[3] over the 1975 to 2003 period varied within a relatively tight band around its long-run mean. Yet starting in late 2003, this ratio increased at an alarming rate. In mid 2006, however, the ratio flattened and kept falling sharply until 2009 (See Figure 1).

What caused this tremendous asset growth and the subsequent puncture is likely to intrigue economists for years. Some have argued that the global economy was in a relatively benign low-volatility environment in the decade leading up to the ongoing crisis (the so-called "Great Moderation", see Stock

[1] Bank of England's Quarterly Bulletin 2008 Q2, Volume 48 No. 2.

[2] The series of facts to follow are borrowed from Acharya and Richardson (2009a).

[3] In particular, this is the ratio of the Office of Federal Housing Enterprise Oversight (OFHEO) repeat-sale house price index to the Bureau of Labor Statistics (BLS) shelter index (i.e., gross rent plus utilities components of the CPI).

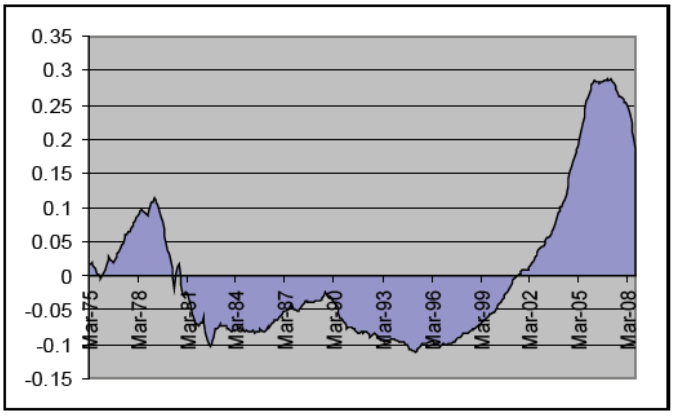

Figure 1: House Price to Rent Ratio. The Figure graphs the demeaned value of the ratio of the Office of Federal Housing Enterprise Oversight (OFHEO) repeat-sale house price index to the Bureau of Labor Statistics (BLS) shelter index (i.e., gross rent plus utilities components of the CPI). Because of demeaning, the average value of this ratio is zero.

and Watson, 2002). In departure from this view, we argue that it is likely not a coincidence that the phase of remarkable asset growth described above started at the turn of the global recession of 2001–2002. In fact, in response to the unprecedented rate of corporate defaults and heightened macroeconomic risk during that recession, the Federal Reserve lowered interest rates to 1%, the lowest level since 1958. A period of abundant availability of liquidity to the financial sector ensued, bank balance-sheets grew two-fold within four years, and when the "bubble burst", a number of agency problems within banks in those years came to the fore. Such problems were primarily concentrated in centers that were in charge of undertaking large risks and took the form of risk-takers being paid out huge bonuses based on the volume of assets they created rather than on (long-term) profits they generated.[4]

[4]See Rajan (2008) and Chapter 8 of Acharya and Richardson (2009b) for a discussion of bank-level principal-agent problem – the "fake alpha" problem when performance is measured based on short-term returns but risks are long-term or in other words in the

In this paper, we develop a theoretical model that combines these ingredients and explains why access to abundant liquidity aggravates the risk-taking moral hazard at banks, giving rise to asset price bubbles. Somewhat perversely, the seeds of crisis are sown when the macroeconomic risk is high and investors in the economy switch from investments to savings in the form of bank deposits. We argue that these bubbles can be counteracted by Central Banks with a contractionary monetary policy in such times, and conversely can in fact be exacerbated by expansionary monetary policy in such times. Expansionary monetary policy may be tempting to persist with when macroeconomic risk is high, but this may flush banks with (even more) liquidity, fueling credit booms and asset price bubbles and sowing seeds of the next crisis.

We first develop a benchmark model in Section 2.1 wherein the representative bank collects deposits from investors and then allocates a fraction of these deposits to investment projects. The bank faces random deposit withdrawals and in case of liquidity shortfalls suffers a penalty cost. The penalty cost could be interpreted as the cost of fire sales or alternatively the cost of raising external finance from markets. In order to avoid such costs the bank has an incentive to set aside some reserves (cash and marketable assets or other forms of ready liquidity). The rest of the deposits are invested in projects (e.g. houses) depending on the demand for loans (e.g. mortgages). The bank chooses the optimal lending rate that maximizes its expected profits subject to the depositors' participation constraint. We show in this benchmark model that the bank lending rate appropriately reflects the underlying risk of the project.

In Section 2.2 we enrich the model to study how agency problems within the bank affect the pricing of loans. In practice, bankers and loan officers ("bank managers") often have incentives to give out excessive loans since their payoffs are proportional to the amount of loans advanced.[5] We show

"tail" – and the role that this problem played in causing the financial crisis of 2007–2009.

[5]The Bureau of Labor Statistics reports that "Most (loan officers) are paid a commission based on the number of loans they originate." (See the Bureau of Labor Statistics' Occupational Outlook Handbook, 2008-09 Edition available at http://www.bls.gov/oco/ocos018.htm#earnings.)

that such incentives can arise as part of an optimal contracting outcome of a principal-agent problem when managerial action or effort is unobservable. We consider a setting where the principal can conduct a (costly) audit to verify whether or not the manager had acted over-aggressively by sanctioning excessive loans. Subsequent to an audit if it is inferred that the manager had acted over-aggressively the manager is penalized a fraction of the penalty costs incurred by the bank arising from liquidity shortfalls. We show that it is optimal to conduct an audit if the liquidity shortfall suffered by the bank is large enough. In this setup the optimal managerial compensation is increasing in the volume of loans but if the manager underprices the risk of the investments (in order to sanction excessive loans) he faces the risk of a penalty if the bank suffers a significant liquidity shortfall. However, we show that such mispricing of risk only occurs when the bank is awash with liquidity (deposits). This is because in presence of excessive liquidity the manager attaches too little weight to the scenario where the bank might later face liquidity shortfalls. In other words, excessive liquidity encourages managers to disregard downside risk, increase loan volume and underprice the risks of projects.

We then show in Section 2.3 that such behavior ultimately has an impact on asset prices. We assume that the demand for loans arises from investments by the household or corporate sector in underlying assets of the economy. To show how asset price "bubbles" are formed we first define the "fundamental" asset prices as those that arise in the absence of any agency frictions within banks. We construct the optimal demand function for assets by bank borrowers and then solve for the underlying asset price given the market clearing condition that the aggregate demand for assets should equal their finite supply. If the bank lending rate underprices risks, then there is an increase in aggregate borrowing from banks. This in turn fuels an excessive demand for assets in the real sector which leads to prices rising above their fundamental values. We interpret this asset price inflation as a "bubble". Importantly, such bubbles are formed only when bank liquidity is high enough as only then do bank managers underprice risk.

Next, in Section 3 we study when are asset price bubbles most likely to

be formed. We show that this is the case when the macroeconomic risk in the economy is high. When macroeconomic risk increases, depositors avoid direct entrepreneurial investments and prefer to save their money in bank deposits which are perceived to be safer. Gatev and Strahan (2006) offer direct empirical evidence consistent with this effect. In our model, such "flight to quality" results in excessive bank liquidity and induces bubble formation in line with our earlier results.

Finally, we study the implications of the results for optimal monetary policy. We show that if the Central Bank adopts a contractionary monetary policy in times of excessive bank liquidity, then it can counter the flight to quality by drawing out the increases in bank liquidity and avoiding the emergence of bubbles. On the contrary, if the Central Bank adopts an expansionary monetary policy then this accentuates the formation of bubbles. Intuitively, an increase in the money supply only serves to increase bank liquidity further when there is already a flight to quality of deposits. Our model can thus explain how lax monetary policy by the Scandinavian Central Banks in 1980's, Bank of Japan during 1986-1987, and the Federal Reserve in the United States during most of the Greenspan era culminated in housing and real estate bubbles in these countries. In contrast, in times of scarce bank liquidity, banks raise lending rates which adversely affects aggregate investment. We show that during these times if the Central Bank adopts an expansionary monetary policy then it can boost aggregate investment by effectively injecting liquidity into the banking system. We thus argue in Section 4 that the optimal monetary policy involves a "leaning against liquidity" approach, which recognizes that "leaning against macroeconomic risk" is not necessarily the desirable policy.

Proponents of the 'Greenspan camp' argue that monetary policy should not be geared towards avoiding the emergence of bubbles and should focus instead on targeting the natural interest rate and the natural rate of employment as has traditionally been the case. This is justified on the basis that central banks cannot pinpoint an asset price bubble. Nevertheless, we prove that targeting bank liquidity is optimal even if central banks are not aware of where the economy is in the *business cycle*. Since the asset price bubble

is intuitively tied to bank liquidity, we believe that the central banks' task in identifying times for employing a contractionary policy is not as onerous as is often suggested since its task is to track the extent of liquidity of the banking (more broadly, financial intermediation) sector.[6]

1.1 Related Literature

While Jensen and Meckling (1976) showed that leverage induces equityholders to prefer excessive risk, our point is concerned with risk-taking incentives inside banks as a function of liquidity. On this front, the mechanics of our model are similar to Myers and Rajan (1998) wherein access to liquidity allows financial firms to switch to riskier assets, and the anticipation of such behavior, renders them illiquid ex ante. The channel in our model is somewhat different in that when banks are flush with liquidity, managers are hedged from the downside risks they undertake, and this induces risk-taking incentives.

Allen and Gale (2000) show in a model of risk-shifting that uncertainty in monetary policy acts to exacerbate the risk-taking incentives ex ante and fosters an asset price bubble. Diamond and Rajan (2008) show that lowering interest rates ex post may be desirable for a central bank in order to avoid bank runs and fire sales, but that this can induce moral hazard and incentivize banks to hold more illiquid assets. It may thus be desirable for the Central Bank to commit to raising interest rates when they are low. In our model, the focus is also on the ex-ante effects of central bank

[6] In fact, a number of economists, including those who traditionally believed that monetary policy should not react to asset price bubbles, have revised their priors on its conduct. Some examples include: (i) "Given the events of the last eight months, it would be foolish not to reconsider the Greenspan doctrine," by Kenneth Rogoff, Financial Times, 16 May 2008; (ii) "I think I am still with the orthodoxy but I have to admit that recent events are sowing seeds of doubt," by Alan Blinder, Financial Times, 16 May 2008; (iii) "A Central Bank should bear in mind those long-run consequences of asset price bubbles and financial imbalances in the setting of current interest rates," by Charles Bean, Financial Times, 16 May 2008; and, (iv) "We need a new philosophical approach...which recognises that market liquidity is beneficial up to a point but not beyond that point..." by Lord Turner, Chairman of the Financial Services Authority, Financial Times, 18 March 2010.

intervention. The difference arises from the fact that our model embeds the principal-agent problem between dispersed shareholders of banks and the loan officers or the risk takers. The severity of the agency problem is affected by the availability of liquidity, and thus by central bank interventions. Importantly, in the absence of bank-level agency problems, there is no excess in credit growth and central bank interventions do not distort credit outcomes.

In another related paper Farhi and Tirole (2009) study how monetary policy has an affect on the collective behavior of banks. They argue that banks have an incentive to correlate their risk exposures because if everyone engages in maturity transformation (for example), authorities have little choice ex post other than facilitating refinancing. They thus argue that the optimal monetary policy is inherently time inconsistent. Acharya and Yorulmazer (2007) also study the incentives for banks to correlate their risks when regulatory forbearance (bailout policy) suffers from a time-consistency problem. In a model with two banks they show that banks have an incentive to reduce their correlations given the possibility that one bank may acquire the other. However, if the rents obtained from such an acquisition are outweighed by the bailout guarantee when both banks fail, then banks herd and seek to maximize their correlation. In contrast to these papers, our paper studies whether or not monetary policy should target asset prices given the agency problems inherent in banks.

In a recent paper, Agur and Demertzis (2010) also argue that monetary policy concerned with financial stability may have to be conservative and set higher rates on average, but in reaction to negative shocks, be more aggressive, i.e. engage in deeper but shorter-lived cuts (compared to traditional monetary policy concerned only with inflation or output) when there are negative macroeconomic shocks.

Adrian and Shin (2009) provide supporting empirical evidence for many of these papers, including ours. They show that the aggregate balance-sheet of financial intermediaries grows more rapidly in times of asset price booms and that such times are also coincident with easing of monetary policy. They argue that growth in financial sector balance-sheets might be the relevant

measure of liquidity to rein in the pro-cyclicality of its risk choices. This is consistent with the key channel of our model that banks expand lending when they experience a liquidity influx (and in the process make worse-quality loans).

Kashyap and Stein (1995) and (2000) find that the impact of monetary policy on lending behavior is stronger for banks with less liquid balance sheets. They interpret their results as support for a "bank lending channel" of monetary policy. Their results are in line with our paper whereby monetary policy has real effects on the economy via the bank lending channel.

A number of recent empirical papers find evidence directly in support of our theory. Jiminez, Ongena, Peydró and Saurina (2009) using twenty-two years of data from the credit register of Spain find that bank risk-taking increases following loose monetary policies. Using data from the Bolivian credit register, Ioannidou, Ongena and Peydró (2009) find that a reduction in monetary policy rates spurs the granting of new loans at lower spreads. Furthermore, they find that this effect is more pronounced for banks with more agency problems suggesting a link from low policy rates to excessive risk-taking. Maddaloni and Peydró (2009) find evidence that low monetary policy rates have resulted in a softening of the lending standards in Europe and USA and that these results are stronger when banking supervision is weak and when bank moral hazard problems are high. Finally, Berger and Bouwman (2010) test our theory and in confirmation with our results find that high liquidity creation is accompanied by a high likelihood of the occurrence of a crisis.

There are several alternative theories of bank lending over the cycle that are not directly related to liquidity and monetary policy. Rajan (1994) argues that it is easier for loan officers to share blame in bad times and this leads to herding and delay of loan-loss recognitions in good times, inducing pro-cyclicality to bank lending policies. He also provides supporting empirical evidence based on the real-estate banking crisis of early 90's.[7]

[7] Mei and Saunders (1997) show that the real-estate lending of U.S. financial institutions exhibits a "trend-chasing" pattern, lending more when real estate returns are expected to be low and vice-versa. They do not, however, relate the pattern to monetary policy.

Thakor (2005) argues that bank over-lending is due to banks permitting higher loan commitments and not invoking the MAC clause during booms given reputational concerns.[8] Ruckes (2004) shows that in expansions banks reduce their screening activity which results in loans being extended to lower quality borrowers. However, in economic downturns banks tighten credit standards. He argues that this explains the substantial variation in credit standards over the business cycle. Dell'Ariccia and Marquez (2006) show that as banks obtain private information about borrowers and information asymmetries across banks decrease, banks may loosen their lending standards resulting in lower profits and expanded aggregate credit. The boom in lending make banks more vulnerable to economic downturns. Matsuyama (2007) studies a model of credit market imperfections with heterogeneous investment projects, and analyses how a movement in borrower net worth causes the composition of the credit to switch between investment projects with different productivity levels, which in turn affect borrower net worth. He shows that these interactions can result in credit cycles (fluctuations in net worth) and credit traps (low borrower net worth). Acharya and Yorulmazer (2008) consider the collective limited liability of banks and show that it induces herding as when banks fail, they impose a negative externality on each other through information contagion. This limited liability effect is stronger in downturns. The view provided by our paper is complementary in that over-lending occurs in our model due to access to liquidity and easy monetary policy and in the form of higher spot lending by banks.

2 The model

2.1 Base case

We consider a three-period model of a bank wherein the bank at $t = 0$ receives deposits D from risk-neutral investors. Each investor deposits 1 unit of his endowment in the bank. The reservation utility of depositors is

[8] Material Adverse Change (MAC) clause permits the bank to decline to lend under the commitment if the borrower's financial condition has declined significantly since the commitment was sold.

given by $\bar{u}$. Hence in order to secure deposits the bank needs to set the rate of return on deposits, r_D, such that the depositors earn an expected payoff of at least $\bar{u}$.

The bank subsequently makes investments in projects while holding a fraction of the deposits as liquid reserves, R. The bank-funded projects either succeed or fail at $t = 2$. The probability of success of bank projects is given by θ and in the event the project is successful it pays off at $t = 2$. The project is illiquid in the sense that if it were to be liquidated prematurely at $t = 1$ the bank faces a penalty or a liquidation cost. The bank observes θ and sets r_L which is the (gross) rate of return on loans. When choosing the lending rate, the bank takes into account the demand function for loans (say, by the corporate sector) which is given by $L(r_L)$ where $L'(r_L) < 0$. Bank reserves are thus given by:

$$R = D - L(r_L)$$

The bank may experience withdrawals at $t = 1$ and for simplicity we assume that the fraction of depositors who experience a liquidity shock is a random variable given by $\tilde{x}$, where $x \in [0, 1]$.[9] The cumulative distribution function of $\tilde{x}$ is given by $F(x)$ while the probability distribution function is denoted by $f(x)$. Each depositor who withdraws early receives 1 unit of his endowment back at $t = 1$. Thus the total amount of withdrawals at $t = 1$ is given by $\tilde{x}D$. If the realization of $\tilde{x}D$ is greater than R, then the bank faces a liquidity shortage, and it incurs a penalty cost, given by $r_p(xD - R)$, which is proportional to the liquidity shortage, where $r_p > r_L > 1$.

The penalty cost can be justified in a number of ways. The bank may be forced to cover the shortfall in a costly manner by selling some of its assets prematurely at fire-sale prices. This is particularly likely when firms

[9] As in Allen and Gale (1998) and Naqvi (2007) we could have assumed that $\tilde{x}$ is correlated with asset quality news in the sense that depositors receive a noisy signal of θ on which they base their decision on whether or not to run. While this is more realistic, it does not affect our qualitative result but highly complicates the analysis. Hence similar to Diamond and Dybvig (1983) and Prisman, Slovin and Sushka (1986) we assume that $\tilde{x}$ is random.

in other industries are also facing difficulties.[10] Alternatively the bank can raise external financing via capital markets. However, this is also costly because raising equity leads to dilution of existing shareholders due to the debt overhang problem (Myers, 1977). Furthermore, raising external finance entails a price impact due to the adverse selection problem a la Myers and Majluf (1984). Capital-raising can also entail costs related to monitoring that the new financier must undertake. Finally, if the bank attempts to cover the shortfall by emergency borrowing from the central bank, this can also be costly as the central bank may charge a penalty rate. Finally, apart from pecuniary costs, the bank may also suffer non-pecuniary costs such as a reputational cost, e.g., the stigma associated with borrowing from the central bank's emergency facilities.

If the projects financed by bank borrowings are successful, then the bank is solvent and is able to repay the patient depositors the promised rate of return of r_D at $t = 2$, whilst the equityholders consume the residual returns from bank investments after these depositors have been paid off. However, in case of the failure of bank-funded projects, the surplus reserves, $R - \tilde{x}D$, if any, are divided amongst the patient depositors whilst the equityholders consume zero. The sequence of events is summarized in the timeline depicted in Figure 2.

Given this setup the risk neutral bank's problem is as follows:

$$\max_{r_L^*, r_D^*, R^*} \Pi = \pi - r_p E\left[\max\left(\tilde{x}D - R, 0\right)\right] \tag{1}$$

[10] Shleifer and Vishny (1992) argue that the price that distressed firms receive for their assets is based on industry conditions. In particular, the distressed firm is forced to sell assets for less than full value to industry outsiders when other industry firms are also experiencing difficulties. There is strong empirical support for this idea in the corporate-finance literature, as shown, for example, by Pulvino (1998) for the airline industry, and by Acharya, Bharath, and Srinivasan (2006) for the entire universe of defaulted firms in the US over the period 1981 to 1999 (see also Berger, Ofek, and Swary, 1996, and Stromberg, 2000). In the evidence of such specificity for banks and financial institutions, James (1991) shows that the liquidation value of a bank is typically lower than its market value as a going concern. In particular, his empirical analysis of the determinants of the losses from bank failures reveals a significant difference in the value of assets that are liquidated and similar assets that are assumed by acquiring banks.

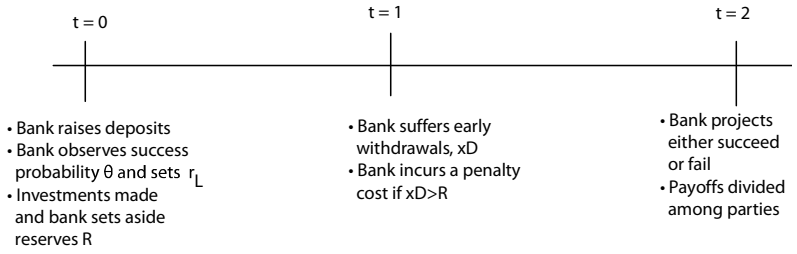

Figure 2: Benchmark model: Timeline of events

subject to

$$E(\tilde{x}) + (1 - E(\tilde{x})) \left[\theta r_D + (1 - \theta) \frac{E[\max(R - \tilde{x}D, 0)]}{(1 - E(\tilde{x}))D} \right] \geq \bar{u} \qquad (2)$$

where π is given by:

$$\pi = \theta \{ r_L L(r_L) - r_D D(1 - E(\tilde{x})) + E[\max(R - \tilde{x}D, 0)] \}. \qquad (3)$$

The above program says that the bank chooses deposit and lending rates as well as the level of bank reserves so as to maximize its expected profits, π, net of any penalty costs incurred in the case of liquidity shortage subject to the participation constraint of the depositors. Expression (2) represents the participation constraint of depositors. A depositor withdraws his funds early with a probability of $E(\tilde{x})$ in which case he receives a payoff of 1. With a probability of $1 - E(\tilde{x})$ the depositor does not experience a liquidity shock in which case he receives a promised payment of r_D if the bank projects succeed (which is with probability θ). In case of the failure of bank investments (which happens with probability $1 - \theta$) any surplus bank reserves are divided amongst the patient depositors. Thus expression (2) states that the depositors must on average receive at least their reservation utility. Equation (3) represents the expected profit of the bank exclusive of the penalty costs. With probability $(1 - \theta)$ bank profits are zero since the investment projects fail. With probability θ the bank projects succeed in which case

the bank's expected profit is given by the expected return from the loans ($r_L L(r_L)$) minus the expected cost of deposits ($r_D D[1 - E(\tilde{x})]$) plus the expected value of net reserve holdings at the end of the period (which is given by the last term of the equation).

Note that for simplicity we have considered a setup with a given penalty cost. In the appendix we consider a setup wherein the penalty costs are explicitly calculated in an environment where the bank finances the shortfall by selling its assets at fire-sale prices. We show that in this three-period environment, the objective function of the bank is analogous to equation (1) and is given by π minus a cost term which is proportional to the bank's liquidity shortfall. Hence our qualitative results are unchanged and thus we use the simpler setup given its parsimony and tractability.

We next solve the bank's optimization problem and derive the first-best lending and deposit rates and the optimal level of bank reserves. Subsequently we study how the riskiness of bank projects as well as bank liquidity affects the loan rate offered by the bank. The results are summarized in Proposition 1.

Proposition 1 *1. The optimal gross lending rate is given by*

$$r_L^* = \frac{1 + (r_p - 1)\Pr(\tilde{x}D \geq R^*)}{\theta\left(1 - \frac{1}{\eta_L}\right)} \tag{4}$$

where $\eta_L = -r_L L'(r_L)/L > 0$ is the elasticity of the demand for loans. The optimal gross deposit rate is given by

$$r_D^* = \frac{(\bar{u} - E(\tilde{x}))D - (1-\theta)E[\max(R^* - \tilde{x}D, 0)]}{\theta(1 - E(\tilde{x}))D}. \tag{5}$$

And, the optimal level of reserves is given by:

$$R^* = D - L(r_L^*)$$

2. *(Risk effect) $\frac{\partial r_L^*}{\partial \theta} < 0$, i.e., an increase in risk $(1-\theta)$, ceteris paribus, increases the equilibrium lending rate.*

3. *(Liquidity effect) $\frac{\partial r_L^*}{\partial D} < 0$, i.e., an increase in bank liquidity, ceteris paribus, decreases the equilibrium lending rate.*

Proof. See Appendix. ■

It is interesting to note that as the elasticity of demand for loans decreases, the lending rate increases and hence the spread between the loan rate and deposit rate increases. This result is consistent with the Monti-Klein (Klein, 1971 and Monti, 1972) model. As expected, the greater the market power of the bank, the higher will be the lending rate that it can charge and vice versa.

The second and third parts of the proposition are also intuitive. The lending rate prices both risk and liquidity. An increase in liquidity lowers the expected cost of liquidity shortage and the bank passes some of this benefit to the borrowers via a lower loan rate.

2.2 Agency problem and over-lending

2.2.1 Setting of the problem

So far we have abstracted from any agency issues between the bank equity-holders and the bank manager. Let us now consider a setting where such issues are relevant.

A study by OCC (1988) found that "Management-driven weaknesses played a significant role in the decline of 90 percent of the failed and problem banks the OCC evaluated... directors' or managements' overly aggressive behavior resulted in imprudent lending practices and excessive loan growth." They also found that 73% of the failed banks had indulged in over-lending. This indicates that principal-agent problems within banks have been one of the key reasons for bank failures and that bank managers tend to engage in 'overly aggressive risk-taking behavior'.[11] Perhaps even more striking evidence is presented by the financial crisis of 2007-2009 which has revealed that in the period preceding the crisis, traders and large profit/risk centers at a large number of financial institutions were paying themselves bonuses based on the size of their risky positions rather than their long-run profitability.

[11] The OCC's study is based on an analysis of banks that failed, became problems and recovered, or remained healthy during the period 1979-1987. The study analysed 171 failed banks to identify characteristics and conditions present when the banks deteriorated.

Moreover, in many cases, it was a conscious choice of senior management to silence the risk management groups that had spotted weaknesses in the portfolio of building risks.[12]

To study how such managerial agency problems can have an effect on bank policies, we model the agency problem within banks explicitly. Let e denote the unobservable effort level of the manager, such that $e \in \{e_L, e_H\}$, where $e_H > e_L$. We assume that although the loans are affected by effort, they are not fully determined by it. The stochastic relationship is necessary to ensure that effort level remains unobservable. We assume that the distribution of loan demand $L(r_L)$ conditional on e_H first-order stochastically dominates the distribution conditional on e_L. In other words, for a given level of lending rate, the manager on average makes a higher volume of loans when he exerts high effort relative to the case where he exerts lower effort, i.e. $E[L(r_L)|e_H] > E[L(r_L)|e_L]$. Furthermore, we suppose that $\Pi|e_H - \Pi|e_L > E[w|e_H] - E[w|e_L]$. This means that the incremental increase in the expected profit from implementing a high effort is greater than the increase in the expected wage costs from implementing the high effort. In other words the principal has an incentive to implement the high effort level since the gains from doing so are greater than the associated costs.

The principal can impose a penalty, ψ, on the manager if it is 'inferred' that the manager had acted over-aggressively. However, in order to impose a penalty on the manager the principal must conduct an audit to verify whether or not the manager had acted over-aggressively. Audits are costly and the cost of an audit is given by $z\left(\hat{\Pi}\right)$, where $\hat{\Pi}$ represents the realized value of the bank's profit net of penalty costs. Following Dye (1986) who presumes that audit costs are increasing in output we assume $z'\left(\hat{\Pi}\right) > 0$.[13]

[12] See Chapter 8 of Acharya and Richardson (2009b), which contains a detailed account of governance and management failures at a large number of financial institutions. The most detailed evidence is for UBS based on its "Shareholder Report on UBS's Write Downs" prepared for the Swiss Federal Banking Commission.

[13] Dye (1986) implicitly assumes that an increase in output due to the manager's efforts translates to an increase in profits. Moreover, even if the pecuniary audit costs do not vary with output, the non-pecuniary audit costs are likely to be higher when profits are high since during these times it is more difficult to justify an audit. Furthermore, there

Let ϕ denote the probability of conducting an audit and let ζ denote the probability that the manager will be penalized if an audit is carried out. The audit technology is imperfect but correlated to the manager's choice of r_L, so that $\zeta > 1/2$ if $r_L < r_L^f$ but $\zeta < 1/2$ if $r_L = r_L^f$, where r_L^f denotes the first-best loan rate.

The manager is an expected utility maximizer with a Bernoulli utility function $u(w, \psi, e)$ over his wages w, potential penalty ψ and effort e. The utility function satisfies $u_w(w, \psi, e) > 0$, $u_{ww}(w, \psi, e) < 0$, $u_\psi(w, \psi, e) < 0$, $u_{\psi\psi}(w, \psi, e) > 0$, and $u_e(w, \psi, e) < 0$ (where the subscripts denote the partial derivatives). This implies that the manager prefers more wealth to less, he is risk averse, dislikes penalties and dislikes high effort. More specifically we assume that the utility function is additively separable and is given by $u(w, \psi, e) = v(w) - c(\psi) - e$, where $v'(w) > 0$, $v''(w) < 0$, $c'(\psi) > 0$ and $c''(\psi) > 0$. The manager's reservation utility is given by u^o.

We assume that the principal observes the distribution of bank liquidity (rather than its exact level) and that liquidity is non-verifiable. This is plausible given that liquidity is not even well-defined as it can take several forms. Furthermore, bank liquidity may be lent out to other banks via the interbank market or conversely it may be the excess liquidity of other banks that makes it way to the bank in question. It is also particularly difficult to verify off-balance sheet liquidity which may take the form of unused loan commitments.

The time line is as follows: At $t = 0$ the principal offers a contract to the manager (such that e_H is chosen). Subsequently, the manager receives deposits, chooses effort and sets the loan rate, r_L. At $t = 0.5$, for a given level of r_L the volume of loans will be realized. As before, at $t = 1$ there may be early withdrawals and finally at $t = 2$ the payoffs are realized and divided between the parties given the contractual terms. The timeline is depicted in Figure 3.

is significant empirical evidence that bigger firms incur higher audit costs. To the extent that highly profitable firms also have a high market capitalization, such firms would have higher audit costs relative to less profitable firms.

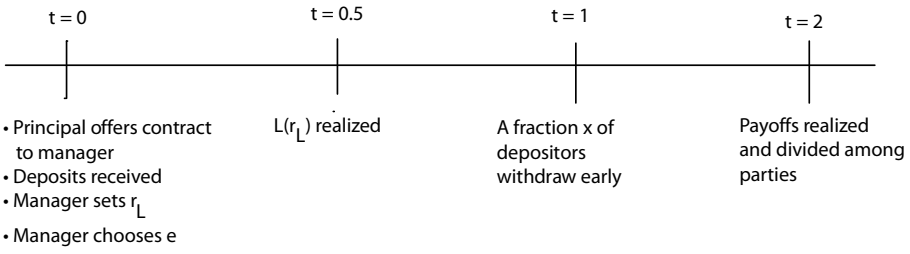

Figure 3: Timeline of events.

2.2.2 Symmetric-information problem

In the presence of symmetric information, there is no agency problem and the bank's problem is analogous to that of Section 2.1 with the bank maximizing

$$\Pi = \pi - r_p E\left[\max\left(\tilde{x}D - R, 0\right)|e = e_H\right] \tag{6}$$

subject to the following participation constraint

$$E\left(\tilde{x}\right) + \left(1 - E\left(\tilde{x}\right)\right)\left[\theta r_D + \left(1-\theta\right)\frac{E\left[\max\left(R - \tilde{x}D, 0\right)|e = e_H\right]}{\left(1 - E\left(\tilde{x}\right)\right)D}\right] \geq \bar{u} \tag{7}$$

where π is given by

$$\pi = \theta\left\{r_L E\left[L\left(r_L\right)|e_H\right] - r_D D\left(1 - E\left(\tilde{x}\right)\right) + E\left[\max\left(R - \tilde{x}D, 0\right)|e = e_H\right]\right\} \tag{8}$$

The first-best lending rate analogous to equation (4) is given by

$$r_L^f = \frac{1 + \left(r_p - 1\right)\Pr\left[\left(\tilde{x}D \geq \bar{R}_f\right)|e = e_H\right]}{\theta\left(1 - \frac{1}{\bar{\eta}_L}\right)} \tag{9}$$

where $\bar{\eta}_L = -r_L\frac{\partial E[L(r_L)|e_H]/\partial r_L}{E[L(r_L)|e_H]} > 0$ and $\bar{R}_f = D - E\left[L\left(r_L\right)\right]$.

2.2.3 Contractual problem under asymmetric information

Next, we introduce the agency problem. The manager can observe the quality of the project, θ, and also the specific level of bank liquidity, D, at the time it is setting the loan rate. However, this information is not

available to the principal and hence it cannot 'infer' the first-best loan rate and thus verify whether or not the manager had acted over-aggressively unless it conducts an audit. In this asymmetric information setting, the contract that the principal offers the manager specifies the compensation of the manager in the form of wages, w, penalties, ψ, as well as the timing of an audit, ϕ. More specifically, the principal needs to solve the following program:

$$\max_{w(L),\psi(S),\phi(S)} \Pi - (E[w(L)] - E[\psi(S)]) - E(z) \tag{10}$$

subject to

$$E[v(w(L))] - E[c(\psi(S))] - e \geq u^o \tag{11}$$

$$E[v(w(L)|e_H)] - e_H \geq E[v(w(L)|e_L)] - e_L \tag{12}$$

$$E\left[c\left(\psi(S)|r_L = r_L^f\right)\right] \leq E\left[c\left(\psi(S)|r_L < r_L^f\right)\right] \tag{13}$$

where $S = \max(xD - R, 0)$ represents the liquidity shortfall of the bank, if any. The above program says that the principal chooses the compensation schedule so as to maximize his expected profits minus the expected compensation of the manager minus the expected audit costs subject to a number of constraints. Constraint (11) is the participation constraint which says that the manager's utility must be at least equal to his reservation utility. Constraint (12) is the incentive compatibility constraint for effort which says that the manager's expected utility from exerting high effort should be at least equal to his expected utility from exerting low effort. Constraint (13) is the incentive constraint for setting the first-best rate. The constraint says that the expected managerial penalty is higher when the manager acts over-aggressively compared to the case where he sets the first-best lending rate.

We can then prove the following proposition.

Proposition 2 *Managerial wages, w, are increasing in loan volume, L. However, if an audit is conducted and it is inferred that the manager had acted over-aggressively then he is penalized where the managerial penalty, ψ, is such that it is increasing in the bank's liquidity shortfall, S.*

Proof. See Appendix. ■

Since the bank's penalty cost is proportional to the liquidity shortfall, the above proposition implies that the managerial penalty is increasing in the bank's penalty cost. The intuition behind the above proposition is straightforward. If managerial compensation only depends on the volume of loans, then the manager will be incentivized to lower lending rates as much as possible because he will not be penalized when low lending rates adversely impact net profits. In this case, the volume of loans will be excessive, reserves will be too low and hence liquidity shortages very likely. However, the presence of a penalty which is increasing in the bank's liquidity shortfall creates a trade-off for the manager. The manager can increase his payoffs by setting a low loan rate and consequently increase the loan volume. But, an increase in loan volume can trigger a liquidity shortfall and subsequently the manager faces a risk of being penalized.

Note that in the absence of constraint (13) there would have been no trade-off and thus the manager would have had no incentive to avoid a shortfall whatsoever. Thus the constraint is important to the extent that it influences managerial decisions. We resolve this trade-off in Proposition 4 where we show that once the manager receives deposits, he will take into account the level of bank liquidity when deciding whether or not to underprice risk.

We assume for simplicity that contracts are linear[14] so that the manager receives a proportion β of loans in the form of bonuses or wages, but is penalized a proportion γ of the bank's penalty costs. Thus the expected penalty cost of the manager is given by $E\left[\left(\psi\left(S\right)\right)\right] = \phi\zeta E\left[\gamma r_p S\right]$, while the expected disutility from penalty is given by $E\left[\left(\psi\left(S\right)\right)\right] = \phi\zeta E\left[c\left(\gamma r_p S\right)\right]$.

Next, we solve for the optimal audit timing, ϕ, as well as the proportion of the penalty, γ, that the manager is charged in the event that he is punished.

[14] The assumption of linear contracts is for simplicity and has no bearing on our results. Furthermore, Holmstrom and Milgrom (1987) argue that real-world compensation schemes involve linear contracts as more complex contracts are subject to costly "gaming" by agents. Holmstrom and Milgrom show that in a dynamic environment with CARA preferences and (binomial) i.i.d. increments in output, optimal linear contracts can be obtained.

In order to do this we add the following constraints to the problem faced by the principal.

$$\phi \in [0, 1] \tag{14}$$

$$\gamma \in [0, \bar{\gamma}] \tag{15}$$

where $\bar{\gamma} \leq 1$. Constraint (14) says that the probability of an audit varies between 0 and 1. Along the lines of Baron and Besanko (1984) constraint (15) says that there is an upper bound on the punishment that can be imposed on the manager and that at most the manager incurs the entire penalty costs. We can then prove the following proposition.

Proposition 3 *The principal will conduct an audit if and only if the liquidity shortfall suffered by the bank exceeds some threshold S^*. Thus the optimal audit timing as defined by the audit probability, ϕ, is given by*

$$\phi = \begin{cases} 1 & \text{if } S > S^* \\ 0 & \text{otherwise} \end{cases} . \tag{16}$$

Furthermore, subsequent to an audit if it is inferred that the manager had acted over-aggressively, the principal will charge the maximum penalty to the manager and thus the manager will bear a proportion, $\bar{\gamma}$, of the bank's penalty costs.

Proof. See Appendix. ■

The intuition behind the above proposition is as follows. By verifying whether or not the agent had acted over-aggressively when liquidity shortfalls are substantial and punishing him with the maximum penalty if it is inferred that he had underpriced risk, the principal discourages the agent from setting a loan rate that is below first-best. Furthermore, if there are no liquidity shortfalls or liquidity shortfalls are minimal then that sends a signal to the principal that the manager was less likely to have acted over-aggressively. Moreover, in the case of liquidity shortfalls if it is inferred that the manager had underpriced risk he is asked to contribute a proportion of the losses. In the absence of liquidity shortfalls there is no such return from a costly audit and hence there is no incentive ex post to conduct an audit.

As in Baron and Besanko (1984) the maximum possible penalty is optimal since the benefit of an audit rises with the managerial penalty.[15]

2.2.4 Liquidity-induced agency problem

The above results can be summarized as follows: Managerial wages are increasing in loan volume. However, an audit is triggered if the liquidity shortfall exceeds a threshold, S^*. In the event of an audit if it is inferred that the manager had acted over-aggressively he is penalized a fraction $\bar{\gamma}$ of the total penalty cost suffered by the bank. One can then prove the following proposition.

Proposition 4 *The manager will engage in overly-aggressive behavior if and only if bank liquidity, D, is sufficiently high.*

Proof. See Appendix. ■

The above proposition says that for high enough bank liquidity the manager has an incentive to engage in overly-aggressive behavior by mispricing in the loan rate the underlying risk of loans. It should be noted that the agency problem is only actuated when bank liquidity is high enough. This is because even though the manager bears a proportion of the penalty costs, in the presence of excessive liquidity, the probability of experiencing a liquidity shortage as well as the probability of an audit is low and this encourages the manager to indulge in excessive lending. Because of this 'insurance effect' of liquidity, the manager's compensation becomes more sensitive to loan volume. The manager will thus indulge in excessive lending as long as his expected utility from doing so exceeds his expected utility conditional on setting the first-best loan rate. On the other hand, for low enough liquidity the agency problem is immaterial since in this case the manager's expected utility conditional on setting the first-best loan rate exceeds his expected utility from acting over-aggressively. Hence for low enough liquidity the

[15] It is not difficult to show that in the presence of an additional limited liability constraint whereby the manager's total penalty cannot exceed $\bar{\psi}$, the managerial penalty would be given by $\psi = \min\left[\bar{\gamma} r_p S, \bar{\psi}\right]$. It is thus in the interest of the principal to impose the maximum possible penalty as long as the limited liability constraint is not violated.

manager does not sanction excessive loans for the fear of incurring a penalty in the event of a liquidity shortfall.

Given the result of Proposition 4 we can then study how a reduction in the precision of the audit technology affects managerial incentives to act over-aggressively. The result is summarized in Proposition 5.

Proposition 5 *The liquidity threshold, D^*, above which an agency problem is actuated, decreases as the precision of the audit technology worsens, ceteris paribus. Consequently the manager is more likely to act over-aggressively.*

Proof. See Appendix. ■

A worsening of the precision of the audit technology makes the auditor more prone to making 'Type I' and 'Type II' errors. In other words, an increase in the likelihood of making incorrect inferences regarding managerial behavior implies that the principal is more likely to penalize the manager even if he had not acted over-aggressively (Type I error) and at the same time is less likely to penalize the manager when he had acted over-aggressively (Type II error). Thus the manager's expected utility from acting over-aggressively increases relative to the case where he does not misprice risk. This exacerbates the distortion in managerial incentives and consequently the manager's incentive to under-price risk increases, ceteris paribus.

2.3 Asset pricing and bubbles

Next we consider the asset pricing implications of our model. We define the fundamental asset price as the price that holds in the absence of any distortions created by agency problems. We then compare the fundamental asset price with the actual asset price which may or may not be distorted depending on whether or not agency problems have been actuated within the banking system. To facilitate this comparison we first model the asset demand by bank borrowers which was so far taken as given. We assume that there exists a continuum, n, of risk-neutral borrowers who have no wealth and hence need to borrow in order to finance investments. Since there is a

continuum of borrowers and loans cannot be conditioned on their size, they can borrow as much as they like at the going rate of interest.

Although there is a continuum of borrowers, we shall analyze the behavior of a representative borrower. This implies that the equilibrium is symmetric and that all borrowers choose the same portfolio. The fact that all borrowers are identical ex post means that the bank cannot discriminate between borrowers by conditioning the terms of the loan on the amount borrowed or any other characteristic.

Let X_d denote the number of units of the asset demanded by the representative borrower and X_s denote the total supply of the risky asset. We assume for simplicity that the supply of the asset is finite and fixed. The asset returns a cash flow of C per unit with a probability of θ. Let P denote the price of one unit of the asset.

As in Allen and Gale (2000) we assume the borrowers face a non-pecuniary cost of investing in the risky asset $b(X_d)$ such that it satisfies the usual neoclassical properties: $b(0) = b'(0)$, $b'(X_d) > 0$ and $b''(X_d) > 0$ for all $X_d > 0$. The purpose of the investment cost is to restrict the size of the individual portfolios and to ensure the concavity of the borrower's objective function.

The optimization problem faced by the representative borrower is to choose the amount of borrowing so as to maximize expected profits. That is

$$\max_{X_d} \theta [CX_d - r_L P X_d] - b(X_d). \tag{17}$$

Note that the borrower has to pay an interest of r_L on its borrowing. The market-clearing condition for the asset is:

$$nX_d = X_s. \tag{18}$$

The first-order condition of the above problem is as follows:

$$\theta [C - r_L P] - b'(X_d) = 0$$

Setting $X_d = 1$ in the first order condition and letting $\tau(X_d) = b'(X_d)$ denote the marginal investment cost, the per unit asset price can be written as:

$$P = \frac{\theta C - \tau(1)}{\theta r_L}. \tag{19}$$

As expected, the asset price is the discounted value of the expected cash flows net of the investment cost. It is also clear that there is a one-to-one mapping from the (gross) lending rate, r_L, to the asset price, P. Let r_L^f denote the fundamental (gross) lending rate which is the rate obtained in the absence of any distortions such as agency problems. Recall that r_L^f is the rate which maximizes the equity value of the bank. Since depositors always earn $\bar{u}$ on average, r_L^f is also the rate which maximizes the value of the bank. Hence r_L^f is given by expression (9). Then the fundamental asset price is given by:

$$P^f = \frac{\theta C - \tau(1)}{\theta r_L^f}. \tag{20}$$

Having derived the fundamental asset price we can next define an asset price bubble. An asset price *bubble* is formed whenever $P > P^f$ since the asset is overpriced. Note that $P > P^f$ as long as $r_L < r_L^f$. Intuitively, a lending rate lower than the fundamental rate creates a high demand for the asset which in turn leads to an increase in asset prices over and above the fundamental values.

From Proposition 4 we know that for high enough liquidity an agency problem is actuated and as a result the loan rate set by the manager is lower than the fundamental rate. Given that the asset price is given by expression (19) and the fundamental asset price is defined by (20) we immediately have the following corollary to Propositions 4.

Corollary 1 *In the presence of an agency problem between the bank manager and the equityholders, an asset price bubble is formed for high enough bank liquidity.*

To better understand the mechanics behind the formation of a bubble, the four-quadrant diagram in figure 4 is useful. Quadrant I in the figure depicts the relationship between the risk of project failure, $(1-\theta)$, and the loan rate, r_L, charged by the bank. In general the higher the riskiness the higher would be the equilibrium lending rate as is captured by the line AA. The loan rate in turn determines the demand for loans and the volume of credit in the economy. For any given lending rate, the expected amount of

bank loans is given by $E[L(r_L)|e_H]$. Since $L'(r_L) < 0$ we know that the lower the loan rate the higher is the amount of expected investment in the economy as is captured by the line NN in quadrant II. From expression (19) we know that a reduction in the loan rate increases asset prices. This is because a decrease in the loan rate increases expected investment and hence pushes up the asset demand. The increase in asset demand in turn pushes up asset prices. This relationship between the demand for the asset and the asset price is captured by the line YY in quadrant III. Finally quadrant IV derives the relationship between the asset price and risk. In general, the higher is the underlying risk the lower will be the asset price as is depicted by the line ZZ. The equilibrium relationship between asset price and risk is derived by tracing the effect of risk on the loan rate, which in turn has an effect on the amount of investment which subsequently determines the asset price.

Let the line AA represent the fundamental relationship between risk and the bank loan rate, i.e. the relationship that would be obtained in the absence of agency issues. Then for any given level of risk, the fundamental asset price would be represented by the line ZZ. However, as we showed in Proposition 4 an agency problem is actuated for sufficiently high bank liquidity levels where the bank loan rate is lower for any given level of risk. This in turn shifts the AA line to A^1A^1. From quadrant II we know that the expected volume of credit in the economy increases following lower loan rates. Consequently asset prices increase as a result of market-clearing as is shown in quadrant III. The final relationship between asset prices and risk is shown in quadrant IV and it is obvious that the actuation of the principal-agent problem shifts the ZZ line to Z^1Z^1. In the end we see that the asset price is higher for the same level of risk leading to the formation of a bubble.

It is also interesting to note that our model implies that the size of the bubble is monotonic in the leverage of bank borrowers. This is because bank borrowers in the model borrow more the lower the lending rates offered by the banks. The greater the severity of the agency problems, the lower are the lending rates, which encourages excessive bank borrowing. This increase

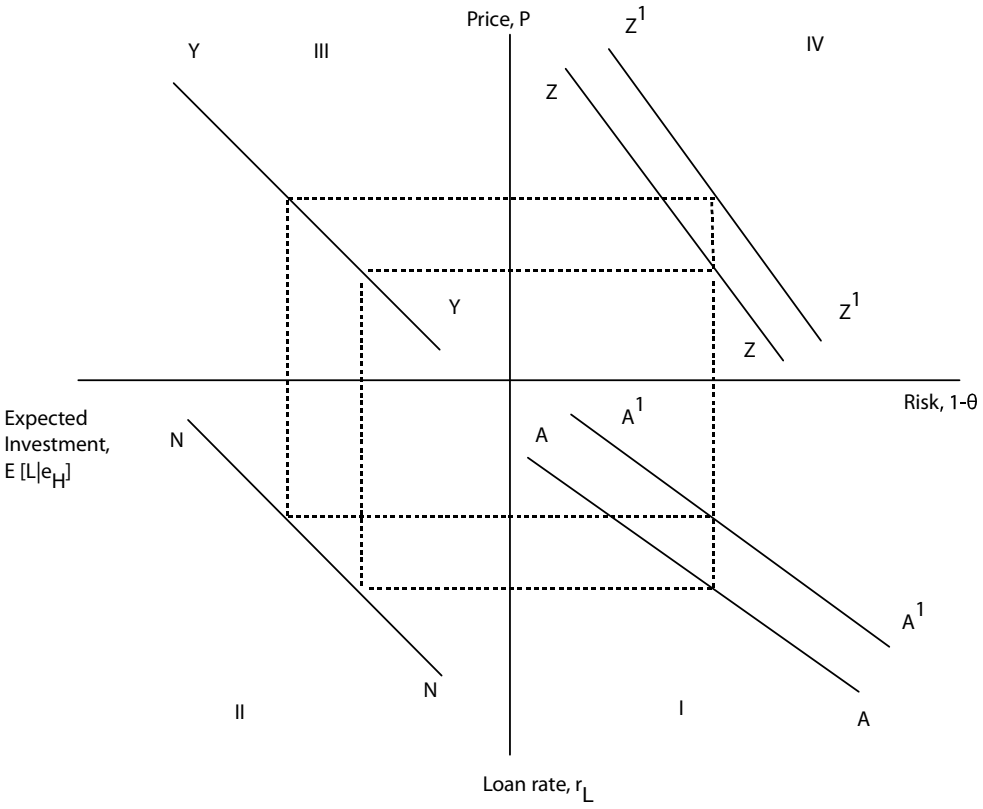

Figure 4: The mechanics of the formation of asset price bubbles.

in leverage in turn increases asset demand which pushes up asset prices.

Our model helps explain how agency problems in the banking sector can induce the formation of asset price bubbles. In terms of the four-quadrant diagram we would be reducing our attention to quadrant IV alone if we fail to consider the role of the banking sector. Embedding of the banking sector in a pricing framework gives us a fuller picture of how the banking sector contributes to equilibrium asset or investment demand in the economy and to the formation of asset price bubbles.

3 When are bubbles likely to be formed?

3.1 High macroeconomic risk

We have established so far that asset price bubbles are formed when bank liquidity is substantially high. So the question that arises is when are banks most likely to be flushed with liquidity. In an empirical study, Gatev and Strahan (2006) find that as spreads in the commercial paper market increase, bank deposits increase while bank asset (loan) growth also increases. The spreads on commercial paper are a measure of the investors' perception of risk in the real economy. Intuitively, when investors are apprehensive of the risk in the entrepreneurial sector they are more likely to deposit their investments in banks rather than make direct investments.[16]

To formalize the above intuition we integrate with the model the entrepreneurial sector, endogenize the quantity of deposits in the economy and show that bank deposits will increase at a time when the underlying risk increases. Consider an economy where entrepreneurs have access to projects

[16] The flight of depositors to banks may be due to banks having greater expertise in screening borrowers during stress times, inducing a natural negative correlation between the usage of lines of credit and deposit withdrawals as argued by Kashyap, Rajan and Stein (2002). Alternatively, the flight may simply be due to the fact that bank deposits are insured (up to a threshold) by the Federal Deposit Insurance Corporation (FDIC) whereas commercial paper and money market funds are uninsured, at least until the extraordinary actions taken by the Federal Reserve during 2008 and 2009. Pennacchi (2006) finds evidence supportive of this latter hypothesis by examining lending behavior of banks during crises prior to the creation of the FDIC.

that yield a terminal cash flow C if it succeeds and 0 otherwise. The probability of success depends partly on the realization of the state variable, $\tilde{\theta}$, and partly on the entrepreneurs' effort decision, ϵ, which identifies whether the entrepreneur is diligent ($\epsilon = 1$) or shirks ($\epsilon = 0$) in which case, entrepreneurs extract a private benefit of B. If the entrepreneur is diligent, the probability of success is θ but in the presence of shirking the probability of success is $\varphi\theta$, where $\varphi \in (0, 1)$. The realization of the state variable θ is observable to the entrepreneurs.

Entrepreneurs promise to pay the risk-neutral investors who invest directly in their projects a face value of y. Thus the expected payoff of the risk-neutral entrepreneurs whose projects are financed by direct investment is given by:

$$E(p) = \begin{cases} \theta(C - y) & \text{if } \epsilon = 1 \\ \varphi\theta(C - y) + B & \text{if } \epsilon = 0 \end{cases}. \tag{21}$$

One can then prove the following proposition.

Proposition 6 *Entrepreneurs will shirk (i.e. $\epsilon = 0$) if and only if $\theta < \theta^*$ where $\theta^* = \frac{B}{(1-\varphi)(C-y)}$.*

Proof. See Appendix. ■

The above proposition says that entrepreneurs are more likely to shirk when macroeconomic risk is high. Intuitively, an increase in macroeconomic risk lowers the success probability of entrepreneurial projects which encourages entrepreneurs to shirk so as to consume their private benefits.

We know that investors on average earn $\bar{u}$ from bank investment. Let $E(u_e|\theta)$ denote the expected utility of the investor from entrepreneurial investment. We assume there exists θ^* such that:

$$E(u_e|\theta < \theta^*) < \bar{u} < E(u_e|\theta \geq \theta^*).$$

That is, the expected utility of investors from direct investment will be higher than that from bank investment as long as $\tilde{\theta} \geq \theta^*$ and conversely the expected utility from bank investment will be higher than that from direct investment as long as $\tilde{\theta} < \theta^*$. This will be the case in the presence of deposit insurance since investors would then perceive bank deposits to be safer and

hence in bad times would prefer to invest in bank deposits rather than make direct entrepreneurial investments. Alternatively it can be argued that in bad times banks are better monitors of entrepreneurial activity and thus entrepreneurs cannot shirk if financed by banks.

If investors perfectly observe θ before deciding whether to invest directly in projects or in banks, all investments will either be channeled directly into projects or into banks depending on the realization of θ. Hence in order to make a more realistic distribution of investments between entrepreneurs and deposits we assume that depositors do not observe θ but each depositor receives an imperfect signal, s, on the basis of which they decide whether to allocate their endowments to direct entrepreneurial investment or bank deposits. A signal $s_j = g$ received by investor j is a good signal which implies that $\theta \geq \theta^*$ but a signal $s_j = b$ is a bad signal which would be an indication to the investor that $\theta < \theta^*$. Thus an investor allocates his capital to entrepreneurial projects only if he receives a good signal.[17]

The probability distribution of the signals is assumed to be identical and independent across depositors and given as: $\Pr(s = g) = \nu\theta$ and $\Pr(s = b) = 1 - \nu\theta$, where $\nu \in (0, 1)$. Investors only observe their own signals and are not aware of the probability distribution of the signals. The simple formulation of the probability distribution implies that a proportion $\nu\tilde{\theta}$ of the investors will allocate their endowments to entrepreneurial projects while a proportion $1 - \nu\tilde{\theta}$ will allocate their endowments to bank deposits. This formulation implies that the better the macroeconomic state θ, the higher would be the amount of direct investments. However a low realization of θ would imply that a larger amount of endowments will be tunneled to bank deposits. Hence, bank deposits are given by $D = (1 - \nu\theta) I$ assuming that there are a total of I risk neutral investors.

It is clear from the formulation of D that bank liquidity increases when

[17] This will have implications for the participation constraint (2). In the constraint, θ_i would need to be replaced by a Bayesian expectation of θ which will slightly alter the equations in Proposition 1. However since all qualitative results are unchanged we can assume for simplicity that when offered a banking contract depositors are able to ascertain whether r_D is high enough to satisfy their reservation utility. This will be the case if the required risk premium on deposits, defined as $\kappa^* \equiv r_D^* - \bar{u}$, is public information.

the underlying macroeconomic risk is high. This is consistent with the findings of Gatev and Strahan (2006) cited earlier. However, we hypothesize that for extremely high macroeconomic risk (very low θ) bank liquidity may be adversely affected. This is because casual observation of crises suggests that once macroeconomic risk crosses a certain threshold the bubble bursts subsequently, adversely affecting banks' liquidity. Nevertheless since this paper focuses on the formation of bubbles rather than the burst of bubbles we do not model this phenomenon but simply capture this by assuming that $D = (1 - \nu\theta)\, I \quad \forall\theta \in [\underline{\theta}, 1]$ where $\underline{\theta}$ is the threshold below which bank liquidity decreases as θ decreases.

We can then prove the following proposition.

Proposition 7 *A bubble is formed in the economy when the macroeconomic risk is high enough. More formally, there exists a threshold* θ^c *such that* $P > P^f$ *if* $\theta < \theta^c$ *where* $\theta^c \in [\underline{\theta}, 1]$.

Proof. See Appendix. ■

As macroeconomic risk increases, there is a flight to quality whereby investors prefer to invest in bank deposits rather than engage in direct lending. Subsequently, banks find themselves flushed with liquidity during times when spreads in the commercial paper market (i.e. the direct funding costs of entrepreneurs by investors) are high. Excessive liquidity encourages bank managers to increase the volume of credit in the economy by mispricing of underlying risk. This in turn fuels a bubble in asset prices.

3.2 Loose monetary policy

Before we turn to the implications for central bank's monetary policy, we briefly discuss how monetary policy has a direct effect on bank's liquidity. When embarking on an expansionary monetary policy via open market operations, central banks buy government securities from primary dealers who have accounts with depository institutions. The way this transaction works in practice is that the central bank directly credits the reserves which commercial banks have with the central bank, hence effectively increasing the

deposit base of the bank. On the other hand, in order to implement a contractionary monetary policy, the central banks sell government securities to primary dealers and at the same time debit their accounts which effectively reduces the deposit base of banks. Hence bank deposits are a function of both macroeconomic risk (θ) as well as monetary policy (M):

$$D = D(\theta; M) \tag{22}$$

The above relationship is depicted in Figure 5. As discussed in the previous section, as macroeconomic risk increases there is a flight to quality whereby bank deposits increase and this continues until risk crosses the threshold $1-\underline{\theta}$ after which more and more investors prefer to just consume their reservation utility. In the absence of an active monetary policy, the relationship between bank liquidity and risk is given by DD. However, following an expansionary monetary policy, bank liquidity increases for the same level of risk and the DD line shifts upwards to D^+D^+. Conversely, subsequent to a contractionary monetary policy, bank liquidity decreases for the same level of risk and consequently the DD line moves downwards to D^-D^-.

In the figure, D^* is the liquidity threshold above which asset price bubbles are formed. When macroeconomic risk increases above $1-\theta^c$ to say $1-\theta^1$, bank liquidity crosses the threshold D^* to D^1 leading to the formation of a bubble. However the central bank can offset this effect via a contractionary monetary policy which will shift the DD line downwards. The magnitude of the contractionary monetary policy should be such that the DD line moves downwards to at least D^-D^-. As can be seen from the figure this is the minimum shift that is required to ensure that for the new level of risk $1-\theta^1$, bank liquidity is at or below D^*.

It is interesting to note from Figure 5 that even if the macroeconomic risk level is below $1-\theta^c$, central banks can fuel asset price bubbles by adopting loose monetary policies thereby shifting the DD line upwards such that the liquidity level crosses the threshold D^*. Indeed Kindleberger (2005) in his study on the history of financial crises notes that: "Speculative manias gather speed through expansion of money and credit."

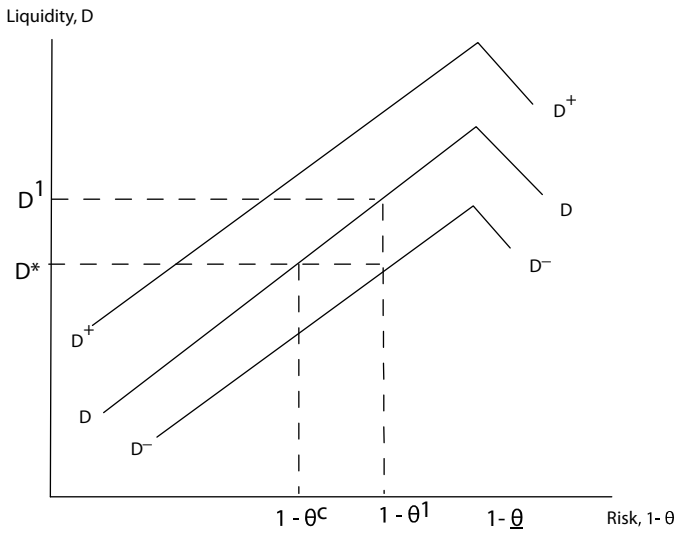

Figure 5: The interplay between bank liquidity, macroeconomic risk and monetary policy.

4 Optimal monetary policy

We next formalize the argument in Section 3.2 and study implications for optimal monetary policy in the presence of the following trade-off faced by the central bank: An increase in money supply increases aggregate investment, but an increase in money supply also increases bank liquidity and we know from our earlier results that excessive bank liquidity can induce bubbles in asset prices. As discussed formally in the extension (section A.2 of the Appendix) bubbles are costly given that aggressive behavior of managers and underpriced loan rates result in a deterioration in the quality of bank investments which in turn increases the average default risk and hence the expected default costs. Let the expected cost of the bubble, conditional on $P > P^f$, be denoted by $\Omega(\Delta)$, where $\Delta \equiv E\left[\left(P - P^f\right) | P > P^f\right]$ denotes the expected size of the bubble. We make the plausible assumption that the cost of the bubble increases with the size of the bubble, i.e. $\Omega'(\Delta) > 0$. (This is also shown formally in extension A.2 of the Appendix).

The trade-off faced by the central bank can be expressed by the following

expected welfare (objective) function:

$$\max_{M^*} W = \bar{L}_H(r_L) - [\Pr(D(\theta) > D^*)\,\Omega(\Delta)] \tag{23}$$

where M^* denotes the optimal money supply and $\bar{L}_H(r_L) \equiv E[L(r_L)|e_H]$ represents the expected demand for bank loans conditional on high effort exerted by bank managers. Since entrepreneurs have zero wealth, $\bar{L}_H$ also represents the expected investment made by entrepreneurs. The second term denotes the expected cost of a bubble since a bubble is formed when deposits cross the threshold D^*.

Taking the derivative of (23) with respect to M we get the first-order condition:

$$\bar{L}'_H \frac{dr_L}{dM} = \Omega(\Delta)\frac{\partial \Pr(D > D^*)}{\partial M} + \Pr(D > D^*)\frac{\partial \Omega(\Delta)}{\partial M} \tag{24}$$

where $\bar{L}'_H = \partial \bar{L}'_H / \partial r_L$. The LHS in (24) represents the marginal benefits (MB) of expansion. The RHS represents the marginal costs (MC) of expansion. Thus the central bank expands money supply up to the point where the marginal benefits of expansion (in terms of increased investment) just equal the marginal costs of expansion (in terms of a higher likelihood of a bubble and the associated costs).

We assume that the SOC is satisfied, namely: $\frac{\partial MB}{\partial M} < \frac{\partial MC}{\partial M}$. This will be the case if there exist diminishing returns on investment and if the marginal cost of a bubble is a non-decreasing function of the money supply. The assumption of diminishing returns to investment implies that as money supply increases, the marginal benefits (in terms of higher investment) increase but at a decreasing rate. As discussed in section A.2 of the appendix, managers progressively make worse quality loans, which can also explain the diminishing returns on investment. The assumption that the marginal cost of a bubble is a non-decreasing function of money supply implies that as bank liquidity and subsequently the expected size of the bubble increases, the incremental cost of the bubble does not decrease. This is also plausible because if anything we expect the marginal cost of a bubble to be an increasing function of the size of the bubble.

We can now prove the following proposition.

Proposition 8 *The optimal monetary policy dictates that the central bank decrease the money supply as macroeconomic risk,* $(1-\theta)$*, increases as long as the size of the bubble is increasing in macroeconomic risk, i.e.,* $d\Delta/d\theta < 0$*.*

However, if the size of the bubble is decreasing in macroeconomic risk, i.e., $d\Delta/d\theta > 0$*, then the optimal monetary policy dictates that the central bank increase the money supply as macroeconomic risk increases. More formally,*

$$\frac{dM^*}{d\theta} \begin{cases} > 0 & \text{if } \frac{d\Delta}{d\theta} < 0 \\ < 0 & \text{if } \frac{d\Delta}{d\theta} > 0 \end{cases}.$$

Proof. See Appendix. ■

The intuition behind the above proposition is as follows: If the expected size of the bubble is increasing in macroeconomic risk, i.e. $\frac{d\Delta}{d\theta} < 0$, say for instance due to a flight to quality effect which increases bank liquidity, then this raises the cost of bubbles. The central bank can counter this effect by decreasing the money supply and hence draining out liquidity from the banking system. If, on the other hand, bank liquidity is decreasing in macroeconomic risk and consequently the expected size of the bubble decreases as the underlying risk increases, i.e. $\frac{d\Delta}{d\theta} > 0$, say for instance, due to investors' loss of confidence in times of a crisis, then the central bank can offset this effect by increasing the money supply. In other words, the central bank should *lean against macroeconomic risk as long as the expected cost of a bubble is increasing with risk, but should lean with macroeconomic risk as long as the expected cost of a bubble is decreasing with risk.*

Proponents of the Greenspan camp may argue that the central bank may not be aware where we are in the business cycle and hence whether bank liquidity is increasing or decreasing in macroeconomic risk. Nevertheless, it can be shown that a much simpler policy recommendation is to *lean against bank liquidity* regardless of where we are in the business cycle. The optimality of this policy is stated in the following proposition.

Proposition 9 *The optimal monetary policy implies a leaning against liquidity approach, i.e., tightening monetary policy in times of excessive bank liquidity and loosening monetary policy in times of falling bank liquidity. More formally,* $\frac{dM^*}{dD} < 0 \ \forall\theta$*.*

Proof. See Appendix. ■

The above proposition is intuitive. In times of excessive bank liquidity, bubbles are likely to be formed and the central bank can avoid the formation of bubbles by tightening monetary policy. On the other hand, in times of scarce liquidity, banks raise loan rates and hence aggregate investment is adversely affected. The central bank can prevent the fall in investment by loosening monetary policy.

We thus argue that the 'Greenspan put' should be employed in times of falling bank liquidity. However, in times when banks are flush with liquidity, a loose monetary policy only enhances the liquidity insurance enjoyed by banks, and thus aggravates their risk-taking incentives. This in turn increases the likelihood of bubbles in asset prices.

4.1 Discussion

Traditionally, as suggested by the Taylor rule, monetary policy has targeted interest rates and employment. However, in the light of our results, we argue that monetary policy should also target asset prices. Our results suggest that asset prices can be targeted if the monetary authorities adopt a "leaning against liquidity" approach. In fact we showed that a "leaning against liquidity" policy performs a twofold purpose: In times of abundant liquidity it counters the surge in asset prices, whilst in times of scarce liquidity it performs the role of quantitative easing and subsequently encourages investment.

Allen and Gale in their book "*Understanding financial crises*" document the following: "In Finland an expansionary budget in 1987 resulted in massive credit expansion. The ratio of bank loans to nominal GDP increased from 55 percent in 1984 to 90 percent in 1990. Housing prices rose by a total of 68 percent in 1987 and 1988... In Sweden a steady credit expansion through the late 1980's led to a property boom." These observations are perfectly in line with our model. Loose monetary policies can potentially lead to excessive liquidity in the banking system which in turn encourages bank mangers to underprice the underlying risk and thereby increase the volume of credit in the economy. This in turn creates an asset price bubble.

Our model can also explain how lax monetary policy in Japan during the mid 1980s led to asset price inflation. Bank of Japan (BOJ) reduced the official discount rate five times between January, 1986 and February, 1987, leaving it finally at 2.5 percent. It is widely accepted that the easy credit policies adopted by BOJ created excess liquidity in the Japanese economy, as also acknowledged by Goyal and Yamada (2004). The sequence of events started with the Plaza Accord (1985), in which the G5 countries agreed on a stronger yen so as to lower the U.S. trade deficit. However, BOJ's intervention in foreign exchange markets appreciated the yen rapidly. Responding to the strengthening yen and seeking to avert deflationary effects in the domestic economy, Bank of Japan lowered interest rates and consequently increased liquidity in the economy. In the subsequent years a large real estate bubble was formed.

One of the causes of the current subprime crisis has been suggested to be the loose monetary policy adopted by the Federal Reserve in the United States. In 2003, the Fed lowered the federal funds rate to 1% - a level that at that time was last seen only in 1958. Subsequently banks mispriced risk and indulged in over-lending which finally culminated in the subprime crisis. In fact the world was awash with liquidity prior to the crisis. We would thus argue that this excess of liquidity contributed significantly in causing the crisis.

The issue of when a central bank should tighten monetary policy following a crisis has resurfaced in the aftermath of the rescue packages administered to recover from the crisis of 2007-09. For instance, the Federal Reserve in the United States has discussed raising the interest paid to banks on their reserves holdings and selling its inventory of mortgage-backed assets as potential tools. The Federal Reserve Chairman Bernanke has however assessed that "The economy continues to require the support of accommodative monetary policies. However, we have been working to ensure that we have the tools to reverse, at the appropriate time, the currently very high degree of monetary stimulus" (Financial Times, February 11 2010). Broadly though, Chairman Bernanke has stated a preference for "leaving rates low for an extended period."

In contrast, some other countries have already started the monetary tightening process. China, in particular, has "ordered its commercial banks to increase the reserves (by 50 basis points from February 25) they hold, as an effort to control rapid lending, rather than significantly tighten monetary policy" (Financial Times, February 13 2010). The Chinese economy expanded by 10.7 per cent in the fourth quarter of 2009 and Chinese banks issued a record Rmb9,600bn in new loans in 2009, about double the amount from the previous year, which fueled a rapid increase in asset prices, especially in Chinese stock markets. House prices in China had increased by 7.8 per cent in December 2009 from the same month a year earlier (Financial Times, January 14 2010). Not surprisingly, given the results of our model, the liquidity of Chinese banks also soared during this period. In fact, household and corporate deposits in the Chinese banking system are now equivalent to a record 150 per cent of gross domestic product (Financial Times, March 3 2010).

Both of these examples get at the heart of our policy discussion. Our model highlights that the key parameter to examine is the extent of bank lending in the economy, as in the discussion about Chinese lending and asset prices above, and that the risk of the Federal Reserve not tightening monetary policy sufficiently soon is precisely that lending may take off by several multiples given the high levels of bank liquidity (reserves) and force the central bank to either tighten excessively ex post or be mopping up after the asset prices have been inflated too high.

4.1.1 Other policy tools

Another policy tool that central banks can employ to mitigate the formation of asset price bubbles is the imposition of minimum liquidity requirements. Suppose banks are required to maintain a minimum liquidity requirement but are penalized whenever their liquidity falls below this level. In the absence of a minimum liquidity requirement a shortfall was induced whenever liquidity was insufficient to service withdrawals. However, in the presence of a minimum liquidity requirement the bank suffers a liquidity shortfall whenever its liquidity drops below the minimum requirement, following which it

suffers a penalty. Such a regulatory requirement will reduce the incentives of bank managers to act over-aggressively given the potential penalty they will suffer in the event of a liquidity shortfall below the minimum level. Nevertheless, as before, if bank liquidity is high enough bank managers will indulge in risk-taking. However, the important difference is that the liquidity threshold, D^*, above which agency problems are actuated will increase in the presence of minimum liquidity requirements. This will reduce the probability of the formation of bubbles. Given that bubbles are still formed in the presence of high enough liquidity, minimum liquidity requirements are a complement but not a substitute to our recommended policy tool of "leaning against liquidity".

Finally, Naqvi (2007) shows that the central bank's lender of last resort operations need to be complemented ex ante by an efficient supervisory framework so as to avoid the moral hazard repercussions of bail-outs. What we learn from our paper is that such supervision is even more essential during times when the banking system is flushed with liquidity. This is because during such times bank managers are more likely to under-price risk and hence over-invest. Thus adequate supervision in times of abundant liquidity might be another possible tool to mitigate the risk-taking appetite of banks.

5 Conclusion

We develop a theory of bank lending explaining how the seeds of a crisis may be sown when banks are flush with liquidity. The main empirical implication of our model is that excessive liquidity induces risk-taking behavior on the part of bank managers. In summary, we obtain the following results: (a) bank managers will behave in an overly-aggressive manner by mispricing risk when bank liquidity is sufficiently high; (b) asset price bubbles are formed for high enough bank liquidity; (c) bubbles are more likely to be formed when the underlying macroeconomic risk is high as it induces investors to save with banks rather than make direct entrepreneurial investments; and, finally (d) bubbles are more likely to be formed following loose monetary

policies adopted by the central bank.

We also show that the optimal monetary policy involves a "leaning against liquidity" approach, i.e. a central bank should adopt a contractionary monetary policy at times when banks are awash with liquidity so as to draw out their reserves; and it should adopt an expansionary monetary policy at times when banks have scarce liquidity so as to boost investment.

Some, most notably Alan Greenspan, have argued that "we are never certain where we are in the cycle"[18] and hence monetary policy should not be used to target asset prices. Nevertheless, our model showed that even if this is the case a "leaning against liquidity" policy can be rationalized. Thus we argue that monetary policy should target not just interest rates and employment but also asset prices as they are reflections of the risk appetite of the financial intermediation sector.

Finally, it should be noted that an increase in macroeconomic risk can also increase bank liquidity of developed economies due to global imbalances. For instance, Caballero (2009) argues that as a result of the NASDAQ crash there was an increased global demand for safe securities and the U.S. financial system catered to this demand by creating collateralized debt obligations (CDOs). This in turn was conducive to global imbalances whereby there was an influx of liquidity in the U.S. financial system from emerging economies. Inevitably this increased the liquidity of the U.S. banking system.

More broadly speaking the rise in deposits could also be interpreted as capital inflows which find its way in the U.S. financial system. For instance, Jagannathan et al. (2009) argue that after the stock market crash of 2000, savings from China flowed into the U.S. debt market. The flow of money into securitized mortgage pools drove down the cost of borrowing resulting in a housing bubble. Reinhart and Rogoff (2008) also show that standard indicators for a financial crisis include rising leverage and asset price inflation. Furthermore they find that leverage booms and asset price inflations are preceded by large capital inflows. In fact, as suggested by Jagannathan et al. (2009) an increase in global (rather than just domestic) macroeconomic risk can also lead to capital inflows into the U.S. given that U.S. Dollar is the

[18] Alan Greenspan, Financial Times, 27 May 08.

reserve currency of the world and most commodities are traded in Dollars. The U.S. is thus a natural recipient of liquidity from developing and emerging countries wishing to build up their reserves as a buffer against increasing macroeconomic risk. To the extent that such liquidity finds its way into the banking system, global imbalances can also actuate agency problems within banks resulting in a mispricing of risk and bubble formation. We aim to explore these linkages further in our future work.

Appendix

A.1 Extension: Bank's objective function in the presence of fire-sale prices for assets liquidated prematurely. In this appendix we will derive the bank's objective function in a setup where instead of taking the penalty cost term as exogenous we derive the term in an environment where the bank can finance the liquidity shortfall by liquidating assets at fire-sale prices. We show that the objective function in this setup has the same generic form as that in the main body of the paper and hence the qualitative results are unaffected.

The timing of the game is as follows. At time $t = 0$ the bank receives deposits and then invests a fraction of these deposits in investment projects via loans to entrepreneurs. A proportion of the total loans, $\mathcal{L}$, are of very good quality such that they repay early with probability 1 at $t = 1$. These loans are indexed by L^e and the rate of return on these loans is r_L^e. However such investments are scarce so the bank cannot make its entire investments in such assets. The rest of the investments are made in assets which repay at $t = 2$ with a probability of θ. These investments are indexed by L and the rate of return on these investments is given by r_L. Thus the total investment is given by $\mathcal{L}= L + L^e$. The rest of the deposits are retained as reserves, R.

At time $t = 1$ the bank distributes a fraction of its 'early returns', $r_L^e L^e$, to its equityholders in the form of dividends where the dividend payout ratio is given by d. For simplicity we normalize $d = 1$, which implies that the early returns are accrued by the equityholders. Furthermore, at $t = 1$ the bank experiences a random liquidity shock such that there are withdrawals of $\tilde{x}D$. If the bank reserves at $t = 1$ are insufficient to service

these withdrawals then the bank will have to inefficiently liquidate a fraction, ξ, of its portfolio. The value of the bank's portfolio at $t = 1$ is given by $r_L^e L^e + r_L \theta L$. However, the fire-sale value of the portfolio is $\ell\left(r_L^e L^e + r_L \theta L\right)$, where $\ell \in (0,1)$. Thus the fraction of premature liquidation is given by $\xi = \left(xD - R\right) / \left[\ell\left(r_L^e L^e + r_L \theta L\right)\right]$. We assume for simplicity that the value of the bank's portfolio is high enough so that it can service its withdrawals without going bankrupt. This is just for simplicity and allowing for bankruptcy does not change the results.

Finally at time $t = 2$ the returns from bank investments, if any, are divided amongst the depositors and the bank equityholders. With probability θ, the return from the 'late' projects will be realized and the equityholders will consume $r_L L - r_D D\left(1 - \tilde{x}\right) + \left(R - \tilde{x}D\right)_+$ where $\left(R - \tilde{x}D\right)_+ = R - \tilde{x}D$ if $\tilde{x}D < R$ and 0 otherwise. Also note that the equityholders need to make the promised payment $r_D D$ to the patient depositors. With probability $1-\theta$ the late projects fail and the value of reserves is divided amongst the depositors. In this case the equityholders only consume their dividends at $t = 1$.

Given this setup the expected profits of the bank's equityholders is given by:

$$\begin{aligned}\Pi_1 \;=\; & \Pr\left(\tilde{x}D < R\right)\left[\theta\left\{r_L L - r_D D\left(1 - E\left(\tilde{x}\right)\right) + E\left[\left(R - \tilde{x}D\right)|\tilde{x}D < R\right]\right\} + r_L^e L^e\right] \\ & + \Pr\left(\tilde{x}D \geq R\right)\left[\theta\left\{r_L\left(1-\xi\right)L - r_D D\left(1 - E\left(\tilde{x}\right)\right)\right\} + r_L^e\left(1-\xi\right)L^e\right].\end{aligned}$$

Since $\Pr\left(\tilde{x}D < R\right) = 1 - \Pr\left(\tilde{x}D \geq R\right)$ and defining

$\pi_{NF} = \theta\left\{r_L L - r_D D\left(1 - E\left(\tilde{x}\right)\right) + E\left[\left(R - \tilde{x}D\right)|\tilde{x}D < R\right]\right\}$ and

$\pi_F \;=\; \theta\left\{r_L\left(1-\xi\right)L - r_D D\left(1 - E\left(\tilde{x}\right)\right)\right\}$ we can rewrite the above as follows:

$$\Pi_1 = \left(\pi_{NF} + r_L^e L^e\right) - \Pr\left(\tilde{x}D \geq R\right)\left[\left(\pi_{NF} - \pi_F\right) + \xi r_L^e L^e\right]$$

where $\left(\pi_{NF} - \pi_F\right)$ denotes the difference in the profits from the late projects with and without fire sales while $\xi r_L^e L^e$ denotes the loss in value from the early projects due to fire sales. Thus the expected profit of the bank is given by the expected profit in the absence of any fire sales, $\left(\pi_{NF} + r_L^e L^e\right)$, minus the probability of a liquidity shortage times the loss in value due to fire sales.

To show the qualitative equivalence of the above dynamic setup with that of the static setup, we first note that $\pi_{NF} = \pi$ as defined in equation (3). Thus,

$$\Pi_1 = (\pi + r_L^e L^e) - \Pr(\tilde{x} \geq R)\left[(\pi_{NF} - \pi_F) + \xi r_L^e L^e\right]$$

Note that $r_L^e L^e$ is a constant and hence maximizing Π_1 w.r.t. r_L and r_D is equivalent to maximizing the following

$$\Pi_2 = \pi - \Pr(\tilde{x}D \geq R)\left[(\pi_{NF} - \pi_F) + \xi r_L^e L^e\right] \tag{25}$$

Also note that $[(\pi_{NF} - \pi_F) + \xi r_L^e L^e]$ is increasing in $(\tilde{x}D - R)$ given that $\frac{\partial \pi_F}{\partial \xi} < 0$ and $\frac{\partial \xi}{\partial(\tilde{x}D - R)} > 0$. Thus $[(\pi_{NF} - \pi_F) + \xi r_L^e L^e]$ is increasing in $(\tilde{x}D - R)$. Note the stark similarity of maximizing Π as in equation (1) with that of equation (25) above. In both cases, the expected profit of the bank is given by π minus a cost term that is proportional to the shortfall. Thus in the main body of the paper, we use the simpler setup with an exogenous penalty cost given that the qualitative results are unaffected.

A.2 Extension: Cost of a bubble

Suppose that bank borrowers have an outside option given by $\bar{u}_B$. In other words, if they do not borrow from banks to invest in projects they can consume their outside option. Furthermore, bank borrowers are heterogenous in the sense that the project of borrower i succeeds with probability θ^i. Banks cannot observe the success probability of individual entrepreneurs and hence under a pooling equilibrium they set a loan rate which cannot be conditioned on the individual θ^i's but is based on the average risk of the participating entrepreneurs. The projects require a capital of K and if the project succeeds it generates a cash flow CK, where $C > 1$. For simplicity we assume that borrowers have limited wealth and hence need to borrow k from banks in order to become entrepreneurs. Given limited liability and a loan rate, r_L, charged by banks, borrower i will invest in his project and hence borrow from the bank if and only if

$$\theta^i (CK - r_L K) \geq \bar{u}_B$$

or if and only if

$$r_L \leq \bar{r}_L^i = \frac{CK - \bar{u}_B/\theta^i}{K}.$$

Hence as long as $r_L > \bar{r}_L^i$ the agent would prefer to consume his outside option and would invest in risky projects only if the loan rate is low enough.

Suppose that there are many entrepreneurs each with a unique success probability θ^i such that θ is uniformly distributed between $[0, 1]$. As long as there are no agency problems between bank equityholders and bank managers, the loan rate is given by r_L^f. Thus in the no-agency world, all entrepreneurs with $\bar{r}_L^i > r_L^f$ will invest whilst the infra-marginal entrepreneur will be indifferent between investing and consuming his outside option. Thus the marginal entrepreneur's success probability, denoted by θ^m, satisfies the following:

$$r_L^f = \frac{CK - \bar{u}_B/\theta^m}{K}$$

or

$$\theta^m = \frac{\bar{u}_B}{K\left(C - r_L^f\right)}.$$

Hence the average success probability of the active entrepreneurs in the no-agency world is given by:

$$\bar{\theta} = \frac{1 + \theta^m}{2}.$$

Next we introduce a cost of default in the economy denoted by Ψ. If the projects of the (active) entrepreneurs fail then a default occurs and the economy suffers a cost Ψ. This cost can take several forms. It could be a cost suffered by the taxpayers who eventually foot the bill for bank bailouts. It could be a cost suffered by the banking system: if banks are not bailed out it would be the cost of bank failures; it can represent a reputational cost for banks; banks can also face a cost in terms of tougher regulations (for instance limits on the size of banks, limits on proprietary trading and/or higher regulatory taxes). Finally it could represent a political cost borne by regulators. Hence in the absence of agency problems, the average default cost of an active entrepreneur is given by

$$\bar{\Psi} = \left(1 - \bar{\theta}\right)\Psi.$$

We have shown that in the presence of an agency problem, bank managers act over-aggressively and set a loan rate, r_L^a, such that $r_L^a < r_L^f$. When

the loan rate is lowered, this encourages over-investment and some agents who were previously consuming their outside options will have an incentive to indulge in bank borrowing. In the presence of an agency problem, the marginal entrepreneur's success probability is given by:

$$\theta^{m\prime} = \frac{\bar{u}_B}{k\left(C - r_L^a\right)}$$

where $\theta^{m\prime} < \theta^m$ since $r_L^a < r_L^f$. Hence the average success probability of the active entrepreneurs in the presence of an agency problem is given by:

$$\bar{\theta}' = \frac{1 + \theta^{m\prime}}{2}.$$

It is clear that the average success probability of projects falls in the presence of an agency problem. In other words, the average quality of loans deteriorates when managers behave over-aggressively. The average default cost is now given by

$$\bar{\Psi}' = \left(1 - \bar{\theta}'\right)\Psi$$

where $\bar{\Psi}' > \bar{\Psi}$.

Intuitively when managers act over-aggressively they set a loan rate which encourages excessive borrowing. This leads to a deterioration of the average quality of loans and hence increases the economy's cost of default. Thus in bubble periods there is over-investment due to a worsening of the quality of loans. This in turn is conducive to higher default costs on average.

Thus the cost of the bubble can be defined as

$$\Omega \equiv \bar{\Psi}' - \bar{\Psi}.$$

Let $\Delta \equiv P - P^f$ denote the size of the bubble. Then note that $\Omega'(\Delta) > 0$. This is because a reduction in lending rates (which implies an increase in asset prices given equation (19)) lowers $\theta^{m\prime}$ which in turn lowers $\bar{\theta}'$ and hence increases the average default cost $\bar{\Psi}'$. In other words, the cost of the bubble is increasing in the magnitude of the bubble.

A.3 Proofs

Proof of Proposition 1. The participation constraint of the bank will be binding because otherwise the bank can increase its expected profits by

slightly reducing r_D. Thus, r_D^* is given by the solution to the following:

$$E(\tilde{x}) + (1 - E(\tilde{x})) \left[\theta r_D + (1 - \theta) \frac{E[\max(R - \tilde{x}D, 0)]}{D} \right] = \bar{u}$$

Solving for r_D^* we get (5).

We can then substitute r_D^* in the bank's objective function and hence r_L^* will be the solution to the following unconstrained maximization problem:

$$\begin{aligned} \max_{r_L^*} \Pi &= \theta \{ r_L L(r_L) - r_D^* D(1 - E(\tilde{x})) + E[\max(R - \tilde{x}D, 0)] \} \\ &\quad - r_p E[\max(\tilde{x}D - R, 0)]. \end{aligned}$$

Assuming that Π is quasi-concave in r_L and noting that $R = D - L$, the maximum is characterized by the following first order condition:

$$\frac{\partial \Pi}{\partial r_L} = \theta L(r_L) - \theta \Pr[\tilde{x}D < R] L'(r_L) + \theta r_L L'(r_L)$$

$$- r_p \Pr[\tilde{x}D \geq R] L'(r_L) - \theta D(1 - E(\tilde{x})) \frac{\partial r_D^*}{\partial r_L} = 0. \tag{26}$$

Noting that $\partial r_D^* / \partial r_L = (1 - \theta) \Pr[\tilde{x}D < R] L'(r_L) / \theta D(1 - E(\tilde{x}))$ and solving for r_L after some simplification we get (4). Thus the optimal reserve level is given by $R^* = D - L(r_L^*)$ which proves the first part of the proposition.

From the FOC (26), if we solve for r_L^* directly without exploiting the definition of η_L we get the following expression for the return on loans:

$$r_L^* = \frac{1}{\theta} - \frac{L}{L'} + \frac{(r_p - 1) \Pr(\tilde{x}D \geq R^*)}{\theta} \tag{27}$$

Taking the partial derivative of the above expression w.r.t. θ we get:

$$\frac{\partial r_L^*}{\partial \theta} = - \frac{1 + (r_p - 1) \Pr(\tilde{x}D \geq R^*)}{\theta^2} < 0 \tag{28}$$

since $r_p > r_L > 1$, which proves the second part of the proposition.

Next note that $\partial \Pr(\tilde{x}D \geq R) / \partial D < 0$, i.e. an increase in bank liquidity (deposits) lowers the probability of liquidity shortfalls since $R = D - L$. Then taking the partial derivative of (27) w.r.t. $1 - F(R) = \Pr(\tilde{x}D \geq R)$ we get:

$$\frac{\partial r_L^*}{\partial [1 - F(R)]} = \frac{r_p - 1}{\theta} > 0 \tag{29}$$

Hence $\frac{\partial r_L^*}{\partial D} = \frac{\partial r_L^*}{\partial [1-F(R)]} \frac{\partial [1-F(R)]}{\partial D} < 0$, which proves the third part of the proposition. Q.E.D.

Proof of Proposition 2. Let μ_1, μ_2, μ_3^i denote the Lagrange multipliers for constraints (11), (12) and (13) respectively. Taking the FOC wrt $w(L)$ we get

$$\frac{1}{v'\left(w\left(L\right)\right)} = \mu_1 + \mu_2 \left[1 - \frac{g\left(L\left(r_L\right)|e_L\right)}{g\left(L\left(r_L\right)|e_H\right)}\right]$$

where $g\left(L\left(r_L\right)|e\right)$ is the density function of loans conditional on effort. As is common in the literature, we then invoke the monotone likelihood ratio property (MLRP), i.e. $\left[g\left(L\left(r_L\right)|e_L\right)/g\left(L\left(r_L\right)|e_H\right)\right]$ is decreasing in L. In words, this means that as bank loans increase, the likelihood of getting a given level of loans and profits if effort is e_H, relative to the likelihood if effort is e_L must increase. Given that $v'' < 0$, this implies that the manager's wages are monotonically increasing in L. Similarly, taking the FOC wrt $\psi\left(S\right)$ we get

$$\frac{1}{c'\left(\psi\left(S\right)\right)} = \mu_1 + \sum_{i \neq f} \mu_3^i \left[1 - \frac{h\left(\psi\left(S\right)|r_L^i < r_L^f\right)\left(\zeta|r_L^i < r_L^f\right)}{h\left(\psi\left(S\right)|r_L = r_L^f\right)\left(\zeta|r_L = r_L^f\right)}\right] \tag{30}$$

where $h\left(\cdot|\cdot\right)$ represents the conditional density function of liquidity shortfalls and $\left(\zeta|\cdot\right)$ denotes the probability of the manager being penalized following an audit conditional on whether or not the manager had acted over-aggressively. Assuming that the MLRP holds, $h\left(\psi\left(S\right)|r_L^i < r_L^f\right)/h\left(\psi\left(S\right)|r_L = r_L^f\right)$ is increasing in S, which implies that $1/c'$ is decreasing in S. Noting that $c''\left(\psi\right) > 0$, it follows that $\psi\left(S\right)$ is increasing in liquidity shortfalls. Q.E.D.

Proof of Proposition 3. Taking the FOC of the bank's maximization problem with respect to ϕ we get

$$\begin{aligned} &\zeta E\left[\gamma r_p S\right] - E\left(z|\phi\left(S\right)\right) - \mu_1 \zeta E\left[c\left(\gamma r_p S\right)\right] \\ &+ \sum_{i \neq f} \mu_3^i \left\{E\left[\zeta c\left(\gamma r_p S|r_L^i < r_L^f\right)\right] - E\left[\zeta c\left(\gamma r_p S|r_L = r_L^f\right)\right]\right\} \\ &+ \left(\lambda_1 - \lambda_2\right) = 0 \end{aligned} \tag{31}$$

where λ_1 and λ_2 correspond to the lagrange multipliers for the constraints

$\phi \geq 0$ and $\phi \leq 1$ respectively. An audit will take place if and only if

$$k(S) = \zeta E[\gamma r_p S] - E(z|\phi(S)) - \mu_1 \zeta E[c(\gamma r_p S)]$$
$$+ \sum_{i \neq f} \mu_3^i \left\{ E\left[\zeta c\left(\gamma r_p S | r_L^i < r_L^f\right)\right] - E\left[\zeta c\left(\gamma r_p S | r_L = r_L^f\right)\right] \right\} > 0.$$

This is because if $k(S) > 0$ it implies that that $\lambda_2 > \lambda_1$. But $\lambda_2 > \lambda_1$ if and only if the constraint $\phi \leq 1$ is binding as a binding constraint implies that $\lambda_2 > 0$ but $\lambda_1 = 0$. This would be the case if and only if $\phi = 1$. It follows that $\phi = 1$ if $k(S) > 0$ and $\phi = 0$ otherwise. Let S^* denote the threshold such that $k(S^*) = 0$. In order to prove that it is optimal to audit if and only if $S > S^*$, it would suffice to show that $k'(S)$ is strictly increasing in S.

Analogous to Dye (1985) it can be shown that after some simplification the derivative of $k(S)$ is given by $h\left(\psi(S)|r_L = r_L^f\right) dS$ times the following expression:

$$\int \psi'(S) \left(1 - c'(\psi(S)) \left(\mu_1 + \sum_{i \neq f} \mu_3^i \left[1 - \frac{h\left(\psi(S)|r_L^i < r_L^f\right)\left(\zeta|r_L^i < r_L^f\right)}{h\left(\psi(S)|r_L = r_L^f\right)\left(\zeta|r_L = r_L^f\right)}\right]\right)\right)$$
$$\cdot \left(\zeta|r_L = r_L^f\right) - \int z'(S).$$

where $\psi'(S) = \frac{dE[\gamma r_p S]}{dS}$. Substituting the value of $1/c'$ from equation (30) in the above expression it is clear that the integrand of the first term is zero. We know that $z'\left(\hat{\Pi}\right) > 0$. Since profits are decreasing in the amount of liquidity shortfalls, S, this implies that $z'(S) < 0$. It follows that $k'(S) > 0$. This proves the first part of the proposition.

Taking the FOC of the bank's maximization problem with respect to γ we get

$$J(S) + (\lambda_3 - \lambda_4) = 0$$

where $J(S)$ represents the derivative of the Lagrangian exclusive of constraint (15) with respect to γ and λ_3 and λ_4 correspond to the lagrange multipliers for the constraints $\gamma \geq 0$ and $\gamma \leq \bar{\gamma}$ respectively. If $J(S) < 0$ that would imply $(\lambda_3 - \lambda_4) > 0$ which in turn corresponds to $\lambda_3 > 0$ and

$\lambda_4 = 0$. But if this were the case γ would equal zero and hence $\psi = 0$ which would violate incentive compatibility as given by (30). Similarly, $J(S) = 0$ would imply that the principal is indifferent between any feasible value of γ including $\gamma = 0$ which again violates incentive compatibility. Hence, by contradiction, $J(S) > 0$ and $\lambda_3 < \lambda_4$. But $\lambda_3 < \lambda_4$ implies that $\lambda_4 > 0$ and hence the constraint $\gamma \leq \bar{\gamma}$ is binding, whilst the constraint $\gamma \geq 0$ is slack. It follows that $\gamma = \bar{\gamma}$ if an audit is carried out and it is inferred that the manager had acted over-aggressively. This proves the second part of the proposition. Q.E.D.

Proof of Proposition 4. If the manager engages in overly-aggressive behavior, his expected utility is given by the following expression:

$$\begin{aligned} & E\left[v\left(\beta L^a | e = e_H\right)\right] - \left(\zeta | r_L < r_L^f\right) E\left[c\left(\bar{\gamma} r_p \max\left(S - S^*, 0\right) | e = e_H\right)\right] - e_H \\ = \; & E\left[v\left(\beta L^a | e = e_H\right)\right] - \left(\zeta | r_L < r_L^f\right) E\left[c\left(\bar{\gamma} r_p \max\left(\tilde{x}D - R^a - S^*, 0\right) | e = e_H\right)\right] \\ & - e_H \end{aligned}$$

where $R^a = D - L^a$. (Note that choosing an expected loan volume is equivalent to choosing a lending rate since there is a one-one mapping from the lending rate to the expected loan volume).

However, in the absence of agency problems, the expected loan volume is given by $\bar{L}_H^f = E\left[L^f | e_H\right]$ which denotes the expected loan volume in the first-best world conditional on the manager exerting high effort. Thus the expected utility of the manager in the absence of an agency problem is given by

$$\Pi_m^{na} = E\left[v\left(\beta L^a | e = e_H\right)\right] - \left(\zeta | r_L = r_L^f\right) E\left[c\left(\bar{\gamma} r_p \max\left(S - S^*, 0\right) | e = e_H\right)\right] - e_H. \tag{32}$$

Since the manager can consume this expected utility in the absence of agency problems, Π_m^{na} can be treated as an opportunity cost in the agency world. In other words, the manager has to earn at least Π_m^{na} in the agency world as otherwise the manager would be better off by not acting over-aggressively.

Thus the problem of the manager is to maximize

$$\begin{aligned}\max_{L^a} \Pi_m^a &= E\left[v\left(\beta L^a | e = e_H\right)\right] \\ &\quad - \left(\zeta | r_L < r_L^f\right) E\left[c\left(\bar{\gamma} r_p \max\left(S - S^*, 0\right) | e = e_H\right)\right] - e_H - \Pi_m^{na}\end{aligned} \tag{33}$$

The first order condition is given by:

$$\int \beta v'\left(\cdot\right) dF\left(L | e = e_H\right) - \left(\zeta | r_L < r_L^f\right) \bar{\gamma} r_p \int_{S^*} c'\left(\cdot\right) dH\left(S\right) = 0. \tag{34}$$

where $F\left(\cdot\right)$ and $H\left(\cdot\right)$ represent the distribution functions for loan volume and liquidity shortfalls respectively. The second order condition is given by:

$$\int \beta^2 v''\left(\cdot\right) dF\left(L | e = e_H\right) - \left(\zeta | r_L < r_L^f\right) \left(\bar{\gamma} r_p\right)^2 \int_{S^*} c''\left(\cdot\right) dH\left(S\right) < 0.$$

The SOC above is satisfied since $v''\left(\cdot\right) < 0$ and $c''\left(\cdot\right) > 0$.

Differentiating the total derivative of the FOC wrt D we get:

$$\frac{\partial^2 \Pi_m^a}{\partial L^{a2}} \frac{dL^{a*}}{dD} + \frac{\partial^2 \Pi_m^a}{\partial L^a \partial D} = 0.$$

where L^{a*} denotes the optimal loan volume in the agency world.

Since $\frac{\partial^2 \Pi_m^a}{\partial L_H^{a2}} < 0$ given the SOC, it follows that

$$sign\left(\frac{dL^{a*}}{dD}\right) = sign\left(\frac{\partial^2 \Pi_m^a}{\partial L^a \partial D}\right)$$

Thus one needs to prove that $\frac{\partial^2 \Pi_m^a}{\partial L^a \partial D} > 0$ as this would imply that an increase in deposits would increase the optimal loan volume and hence overall the manager would be better off.

Taking the partial derivative of (34) wrt D we obtain:

$$\begin{aligned}&- \left(\zeta | r_L < r_L^f\right) \left(\bar{\gamma} r_p\right)^2 \int_{S^*} \left(-\left(1 - x\right)\right) c''\left(\cdot\right) dH\left(S\right) \\ = \quad &\left(\zeta | r_L < r_L^f\right) \left(\bar{\gamma} r_p\right)^2 \int_{S^*} \left(1 - x\right) c''\left(\cdot\right) dH\left(S\right) > 0.\end{aligned}$$

The above inequality holds since $c''(\cdot) > 0$. It follows that $\frac{dL^{a*}}{dD} > 0$.

Hence as deposits increase, manager's have an incentive to increase the loan volume (via setting a lower lending rate) since this increases their expected utility. Let D^* denote the deposit threshold such that expression (33) is zero. It follows that for high enough deposits (i.e. $D > D^*$) managerial utility from acting over-aggressively exceeds their utility conditional on setting the first-best rate (i.e. their opportunity cost). Conversely, for low enough liquidity (i.e. $D < D^*$) there will be no agency problems as in this case the manager would be better off by setting the first best lending rate. Q.E.D.

Proof of Proposition 5. If the precision of the audit technology worsens then $\left(\zeta | r_L < r_L^f\right)$ decreases while at the same time $\left(\zeta | r_L = r_L^f\right)$ increases. From expressions (32) and (33) it is clear that a reduction in $\left(\zeta | r_L < r_L^f\right)$ and an increase in $\left(\zeta | r_L = r_L^f\right)$ increases Π_m^a ceteris paribus, and thus lowers the threshold D^*. This in turn increases the manager's incentive to act over-aggressively. Q.E.D.

Proof of Proposition 6. We know from (21) that entrepreneurs will choose $\epsilon = 0$ if and only if $\varphi\theta(C - y) + B > \theta(C - y)$. This implies that entrepreneurs will shirk if and only if $\theta < \frac{B}{(1-\varphi)(C-y)}$. Q.E.D.

Proof of Proposition 7. Comparing (19) with (20) we know that $P > P^f$ if and only if $r_L < r_L^f$. From the proof to Proposition **??** we know that $r_L < r_L^f$ for sufficiently high D. Let D^* denote the threshold below which $r_L < r_L^f$ and assume the plausible that the number of investors I is big enough so that D^* exists. Hence all we need to show is that $\frac{dD}{d\theta} < 0$ $\forall \theta \in [\underline{\theta}, 1]$. Since $D = (1 - \nu\theta) I \ \ \forall \theta \in [\underline{\theta}, 1]$ it follows that $\frac{dD}{d\theta} = -\nu I < 0$ $\forall \theta \in [\underline{\theta}, 1]$. Since D is monotonically decreasing in θ for all $\theta \in [\underline{\theta}, 1]$ it follows that there exists a threshold θ^c below which $D > D^*$ and hence $P > P^f$, where θ^c is such that it solves $D^* = (1 - \nu\theta^c) I$. Q.E.D.

Proof of Proposition 8. Differentiating the total derivative of the FOC (24) wrt θ we get:

$$\frac{\partial^2 W}{\partial M^2} \frac{dM^*}{d\theta} + \frac{\partial^2 W}{\partial M \partial \theta} = 0.$$

Since $\frac{\partial^2 W}{\partial M^2} < 0$ given the SOC, it follows that

$$sign\left(\frac{dM^*}{d\theta}\right) = sign\left(\frac{\partial^2 W}{\partial M \partial \theta}\right).$$

Taking the partial derivative of (24) wrt θ we obtain after some simplification:

$$\frac{\partial^2 W}{\partial M \partial \theta} = -\frac{\partial \Pr\left(D > D^*\right)}{\partial M}\frac{\partial \Omega\left(\Delta\right)}{\partial \theta} - \frac{\partial \Omega\left(\Delta\right)}{\partial M}\frac{\partial \Pr\left(D > D^*\right)}{\partial \theta}. \tag{35}$$

We know that $\frac{\partial \Pr(D>D^*)}{\partial M} > 0$ since an increase in the money supply increases bank liquidity. Also,

$$\frac{\partial \Omega\left(\Delta\right)}{\partial \theta} = \Omega'\left(\Delta\right)\frac{d\Delta}{d\theta}$$

Given that $\Omega'\left(\Delta\right) > 0$, it follows that the first term in (35) is positive if $\frac{d\Delta}{d\theta} < 0$ and vice versa. Similarly,

$$\frac{\partial \Omega\left(\Delta\right)}{\partial M} = \Omega'\left(\Delta\right)\frac{d\Delta}{dM}$$

where $\frac{d\Delta}{dM} > 0$ since an increase in money supply increases bank liquidity, lowers the loan rate and thus increases Δ. This implies that the expected marginal cost of monetary expansion is positive.

Next note that $\frac{\partial D}{\partial \theta} < 0$ is a necessary condition for $\frac{d\Delta}{d\theta} < 0$ since the expected size of the bubble increases as liquidity increases. Conversely, $\frac{\partial D}{\partial \theta} > 0$ is a necessary condition for $\frac{d\Delta}{d\theta} > 0$. It follows that,

$$\frac{\partial \Pr\left(D > D^*\right)}{\partial \theta}\begin{cases} < 0 & \text{if } \frac{d\Delta}{d\theta} < 0 \\ > 0 & \text{if } \frac{d\Delta}{d\theta} > 0 \end{cases}$$

Thus, the second term in (35) is positive if $\frac{d\Delta}{d\theta} < 0$ and vice versa. Q.E.D.

Proof of Proposition 9. Differentiating the total derivative of the FOC (24) wrt D we get:

$$\frac{\partial^2 W}{\partial M^2}\frac{dM^*}{dD} + \frac{\partial^2 W}{\partial M \partial D} = 0.$$

Since $\frac{\partial^2 W}{\partial M^2} < 0$ given the SOC, it follows that

$$sign\left(\frac{dM^*}{dD}\right) = sign\left(\frac{\partial^2 W}{\partial M \partial D}\right).$$

Taking the partial derivative of (24) wrt D we obtain after some simplification:

$$\frac{\partial^2 W}{\partial M \partial D} = -\frac{\partial \Pr(D > D^*)}{\partial M}\frac{\partial \Omega(\Delta)}{\partial D} - \frac{\partial \Omega(\Delta)}{\partial M}\frac{\partial \Pr(D > D^*)}{\partial D} \tag{36}$$

We know that $\frac{\partial \Pr(D>D^*)}{\partial M} > 0$. Also $\frac{\partial \Omega(\Delta)}{\partial D} > 0$ since an increase in bank liquidity lowers the loan rate, increases expected asset prices and thus increases Δ. Thus the first term is negative. Similarly, we know $\frac{\partial \Omega(\Delta)}{\partial M} > 0$ and $\frac{\partial \Pr(D>D^*)}{\partial D} > 0$. Thus the second term is also negative. Q.E.D.

References

[1] Acharya, Viral V., Sreedhar T. Bharath and Anand Srinivasan, 2006, Does industry-wide distress affect defaulted firms? - Evidence from creditor recoveries, *Journal of Financial Economics 85(3)*, 787-821.

[2] Acharya, Viral V., and Matthew Richardson, 2009a, Causes of the Financial Crisis, *Critical Review*, 21(2:3), 195–210.

[3] Acharya, Viral V., and Matthew Richardson (editors), 2009b, *Restoring Financial Stability: How to Repair a Failed System*, New York University Stern School of Business, John Wiley.

[4] Acharya, Viral V., and Tanju Yorulmazer, 2007, Too-Many-To-Fail – An Analysis of Time-inconsistency in Bank Closure Policies, *Journal of Financial Intermediation*, 16(1), 1-31.

[5] Acharya, Viral V., and Tanju Yorulmazer, 2008, Information Contagion and Bank Herding, *Journal of Money, Credit and Banking*, 40(1), 215–231.

[6] Adrian, Tobias and Hyun-Song Shin, 2009, Money, liquidity and monetary policy, *American Economic Review 99(2)*, 600–605.

[7] Agur, Itai, and Maria Dermetzis, 2010, Monetary policy and excessive bank risk taking, Working Paper, Dutch Central Bank.

[8] Allen, Franklin, and Douglas Gale, 1998, Optimal Financial Crises, *Journal of Finance 53(4)*, 1245-1284.

[9] Allen, Franklin, and Douglas Gale, 2000, Bubbles and crises, *The Economic Journal 110*, 236-255.

[10] Allen, Franklin, and Douglas Gale, 2007, Understanding financial crises, New York: Oxford University Press.

[11] Baron, David P., and David Besanko, 1984, Regulation, asymmetric information, and auditing, *Rand Journal of Economics 15*, 447-470.

[12] Berger, Allen N., and Christa H.S. Bouwman, 2010, Bank liquidity creation, monetary policy, and financial crises, Wharton Financial Institutions Center working paper.

[13] Berger, Philip, Eli Ofek and Itzhak Swary, 1996, Investor Valuation of the abandonment option, *Journal of Financial Economics 42*, 257-287.

[14] Caballero, Ricardo J., 2010, The "other" imbalance and the financial crisis, NBER Working Paper No. 15636.

[15] Dell'Ariccia, Giovanni, and Robert Marquez, 2006, Lending booms and lending standards, *Journal of Finance 61*, 2511-2546.

[16] Diamond, Douglas, and Philip Dybvig, 1983, Bank runs, deposit insurance and liquidity, *Journal of Political Economy 91*, 401-419.

[17] Diamond, Douglas and Raghuram G. Rajan, 2008, Illiquidity and Interest Rate Policy, Working Paper, University of Chicago.

[18] Dye, Ronald A., 1986, Optimal monitoring policies in agencies, *Rand Journal of Economics 17(3)*, 339-350.

[19] Farhi, Emmanuel and Jean Tirole, 2009, Collective Moral Hazard, Maturity Mismatch and Systemic Bailouts, NBER working paper 15138.

[20] Freixas, Xavier, and Jean-Charles Rochet, 1997, *Microeconomics of Banking*, Cambridge: MIT Press.

[21] Gatev, Evan, and Philip E. Strahan, 2006, Banks' advantage in hedging liquidity risk: Theory and evidence from the commercial paper market, *Journal of Finance 61(2)*, 867-892.

[22] Goyal, Vidhan K., and Takeshi Yamada, 2004, Asset price shocks, financial constraints, and investment: Evidence from Japan, *Journal of Business 77(1)*, 175-199.

[23] Holmstrom, Bengt, 1979, Moral hazard and observability, *Bell Journal of Economics 10*, 74-91.

[24] Holmstrom, Bengt, and Paul Milgrom, 1987, Aggregation and linearity in the provision of intertemporal incentives, *Econometrica 55*, 303-28.

[25] Ioannidou, Vasso, Steven Ongena and José-Luis Peydró, 2009, Monetary policy, risk-taking and pricing: Evidence from a quasi-natural experiment, European Banking Center Discussion Paper No. 2009-04S.

[26] Jagannathan, Ravi, Mudit Kapoor, and Ernst Schaumburg, 2009, Why are we in a recession? The financial crisis is the symptom not the disease!, NBER working paper 15404.

[27] James, Christopher, 1991, The losses realized in bank failures, *Journal of Finance 46(4)*, 1223-1242.

[28] Jensen, Michael and William Meckling, 1976, Theory of the firm: Managerial behavior, agency costs, and ownership structure, *Journal of Financial Economics 3*, 305-360.

[29] Jiminez, Gabriel, Steven Ongena, José-Luis Peydró, and Jesus Saurina, 2009, Hazardous times for monetary policy: What do 23 million loans say about the impact of monetary policy on credit risk, AFA 2009 San Francisco meetings paper.

[30] Kashyap, Anil K., Raghuram G. Rajan and Jeremy C. Stein, 2002, Banks as liquidity providers: an explanation for the co-existence of lending and deposit-taking, *Journal of Finance 57(1)*, 33—74.

[31] Kashyap, Anil K., and Jeremy C. Stein, 1995, The impact of monetary policy on bank balance sheets, Carnegie-Rochester Conference Series on Public Policy 42, 151-95.

[32] Kashyap, Anil K., and Jeremy C. Stein, 2000, What do a million observations on banks say about the transmission of monetary policy?, *American Economic Review 90(3)*, 407-428.

[33] Kindleberger, Charles P., 2005, Manias, panics, and crashes: A history of financial crises, New Jersey: Wiley.

[34] Klein, M. A., 1971, A theory of the banking firm, *Journal of Money, Credit and Banking 3*, 205-218.

[35] Maddaloni, Angela and José-Luis Peydró, 2009, Bank risk-taking, securitization, supervision, and low interest rates: Evidence from lending standards, Working Paper, European Central Bank.

[36] Matsuyama, Kiminori, 2007, Credit Traps and Credit Cycles, *American Economic Review 97*, 503-516.

[37] Mei, Jianping and Anthony Saunders, 1997, Have U.S. Financial Institutions' Real Estate Investments Exhibited "Trend-Chasing" Behavior?, *Review of Economics and Statistics*, 79, 248–258.

[38] Monti, M., 1972, Deposit, credit and interest rate determination under alternative bank objective function. In: Shell, K., Szegö, G.P. (Eds.), *Mathematical methods in investment and finance*, North Holland, Amsterdam, 431-454.

[39] Myers, Stewart C., 1977, Determinants of Corporate Borrowing, *Journal of Financial Economics 5(2)*, 147-175.

[40] Myers, Stewart C., and Nicholas S. Majluf, 1984, Corporate financing and investment decisions when firms have information investors do not have, *Journal of Financial Economics 13,* 187-221.

[41] Myers, Stewart C. and Raghuram G. Rajan, 1998, The paradox of liquidity, *Quarterly Journal of Economics 113(3)*, 733-771.

[42] Naqvi, Hassan, 2007, Banking crises and the lender of last resort: How crucial is the role of information?, EFA 2007 Ljubljana meetings paper.

[43] OCC, 1988, Bank failure: An evaluation of the factors contributing to the failure of national banks, Washington D.C., Office of the Comptroller of the Currency.

[44] Pennacchi, George, 2006, Deposit insurance, bank regulation, and financial system risks, *Journal of Monetary Economics 53(1)*, 1–30.

[45] Prisman, Eliezer Z., Myron B. Slovin, and Marie E. Sushka, 1986, A general model of the banking firm under conditions of monopoly, uncertainty and recourse, *Journal of Monetary Economics 17(2)*, 293-304.

[46] Pulvino, Todd C., 1998, Do asset fire sales exist: An empirical investigation of commercial aircraft sale transactions, *Journal of Finance 53*, 939-978.

[47] Rajan, Raghuram G., 1994, Why Bank Credit Policies Fluctuate? A Theory and Some Evidence, *Quarterly Journal of Economics*, 109(2), 399–441.

[48] Rajan, Raghuram G., 2008, "Bankers' pay is deeply flawed", *Financial Times*, January 9.

[49] Reinhart, Carmen M., and Kenneth S. Rogoff, 2008, Is the 2007 US sub-prime financial crisis so different? An international historical comparison, *American Economic Review 98(2)*, 339-44.

[50] Ruckes, Martin, 2004, Bank competition and credit standards, *Review of Financial Studies 17*, 1073-1102.

[51] Shleifer, Andrei, and Robert W. Vishny, 1992, Liquidation values and debt capacity: A market equilibrium approach, *Journal of Finance 47*, 1343-1366.

[52] Stock, James H. and M. W. Watson, 2002, “Has the Business Cycle Changed and Why?", NBER Macroeconomics Annual, 159-218.

[53] Stromberg, Per, 2000, Conflicts of interest and market illiquidity in bankruptcy auctions: Theory and tests, *Journal of Finance 55*, 2641–2692.

[54] Thakor, Anjan, 2005, Do Loan Commitments Cause Overlending?, *Journal of Money, Credit and Banking*, 37-6, 1067-1100.