2013 | 21

# Working Paper

Norges Bank Research

# Monetary policy and financial stability in the long run

*Jin Cao and Lorán Chollete*

NORGES BANK

**Working papers fra Norges Bank, fra 1992/1 til 2009/2 kan bestilles over e-post:**
servicesenter@norges-bank.no

Fra 1999 og senere er publikasjonene tilgjengelige på www.norges-bank.no

Working papers inneholder forskningsarbeider og utredninger som vanligvis ikke har fått sin endelige form. Hensikten er blant annet at forfatteren kan motta kommentarer fra kolleger og andre interesserte. Synspunkter og konklusjoner i arbeidene står for forfatternes regning.

**Working papers from Norges Bank, from 1992/1 to 2009/2 can be ordered by e-mail:**
servicesenter@norges-bank.no

Working papers from 1999 onwards are available on www.norges-bank.no

Norges Bank's working papers present research projects and reports (not usually in their final form) and are intended inter alia to enable the author to benefit from the comments of colleagues and other interested parties. Views and conclusions expressed in working papers are the responsibility of the authors alone.

ISSN 1502-8143 (online)
ISBN 978-82-7553-774-2 (online)

# Monetary policy and financial stability in the long run ⋆

Jin Cao [a,c], Lorán Chollete [b,*]

[a] *Norges Bank, Norway*

[b] *UiS Business School, Norway*

[c] *CESifo, Germany*

**Abstract**

Most theoretical central bank models use short horizons and focus on a single tradeoff. However, in reality, central banks play complex, long-horizon games and face more than one tradeoff. We account for these issues in a simple infinite-horizon game with a novel tradeoff: higher rates deter financial imbalances, but lower rates reduce the likelihood of insolvency. We term these factors discipline and stability effects, respectively. The central bank's welfare decreases with dependence between real and financial shocks, so it may reduce costs with correlation-indexed securities. In our model, independent central banks cannot in general attain both low inflation and financial stability.

*JEL classification:* E500, G210, G280

*Key words:* Central Bank, Correlation-Indexed Security, Discipline Effect, Stability Effect

⋆ We are grateful for comments from Farooq Q. Akram, Gerhard Illing, Eric Schaanning, seminar participants at DIW Berlin, the University of Munich, and the CESifo Area Conference on Macro, Money & International Finance. The views expressed in the paper are those of the authors and should not be attributed to Norges Bank.

* Addresses: Jin Cao, Bankplassen 2, PB 1179 Sentrum, NO-0107 Oslo, Norway; Lorán Chollete, H142 Ellen og Axel Lunds Hus, NO-4036 Stavanger, Norway

*Email addresses:* jin.cao@norges-bank.no (Jin Cao), loran.g.chollete@uis.no (Lorán Chollete).

# 1 Introduction

What does a central bank do if it finds itself in a situation that is ex ante suboptimal for the banking system? There is a growing literature on optimal central bank policy towards banking intermediaries. Most of this literature uses three or four period examples of the tension between central bank and intermediary incentives, which often allows for attractive optimality results.[1] The relevant challenge which we tackle is that most theoretical models of monetary policy and stability consider short horizons, and only analyze a single tradeoff. This setting is problematic because real world central banks face multiple tradeoffs with long horizons. Such multiple tradeoffs are particularly onerous because many central banks have one main policy tool during normal times, the interbank interest rate. The question we address is 'What are the limits of monetary policy for a one-tool central bank which faces multiple policy tradeoffs that manifest over long horizons?'

Our paper's main contribution to the literature is that we characterize the tradeoffs between inflation, investment and fragility that plague monetary policy, in a simple game-theoretic framework which accounts for long horizons. Moreover, we extend research on central bank independence, by demonstrating the limits on a central bank that attempts to achieve both low inflation and financial stability.

We proceed by modelling excessive systemic risk-taking arising from a fire-sale externality, and then summarising the central bank's tradeoffs. We then model an economy with production and banking sectors that experience exogenous, correlated shocks, which feed back to the central bank. Subsequently, we endogenise feedback between the real economy and banking sector in a static and dynamic game, and demonstrate how a central bank can credibly support a cooperative equilibrium with minimal fragility. A more detailed breakdown of our approach is provided in the next subsection.

## 1.1 Overview of the paper

Since this paper covers a lot of ground, a brief overview to guide the reader is presented here. To establish a theoretical framework, in Section 2 we present a general form of the fragility-investment tradeoff, in which banks' investment decisions affect the likelihood of

[1] See Freixas et al. (2000); Chapter 7 of Allen and Gale (2007); and Allen et al. (2009).

systemic events. This negative externality is not internalised by individual banks, and the resulting excessive aggregate investment in the banking sector increases systemic risk.

We further develop this idea in Section 3 by modeling excessive systemic risk-taking arising from a fire-sale externality, building on Kashyap and Stein (2012). We proceed in two steps. First, we describe the basic setup, where banks are financing long term projects with a mixture of short and long term debt. While short term debt is cheaper, banks are exposed to liquidity risks when they need to roll over the debt. If creditors refuse to roll over the debts in the intermediate term, there will be bank failures and banks will have to engage in fire sales, thereby liquidating premature projects. This fire-sale cost increases with the total assets on sale in the market, but individual banks do not internalise such costs when they make their initial investment decisions, which leads to excessive investment. The market equilibrium thus features excessive financial fragility and inefficiently high social welfare costs from fire sales. Second, we summarize the central bank's tradeoff between financial fragility, output, and price stability. We formalize this scenario with a production sector and banking sector that experience exogenous output shocks and liquidity shocks, respectively, where these two shocks are potentially correlated. A social welfare maximizing central bank therefore faces dual tradeoffs between financial fragility and inflation, when setting short term rates to stabilize output. The covariance of real and financial shocks reflects a feedback effect that directly affects the central bank's expected loss. In Section 3, this feedback effect is taken as exogenous.

In Section 4 we present an extended model, where we explicitly model feedback between the real economy and banking sector, accounting for banks' strategic response to monetary policy. We proceed in three steps. First, we present the basic framework, with banks that invest in firms' capital via short- and long-term debt as before. In this framework, an intermediate output shock to real economy affects bank returns, making debt refinancing feasible. When the central bank sets short term interest rates to cushion an output shock, it affects both output and financial fragility. Second, the changing output level implies a change in banks' profit level, affecting their ability to refinance. Moreover, the short-term interest rate set by the central bank changes banks' refunding costs, which in turn changes the probability of bank failure. When there are more bank failures, more firms need to be liquidated, thus generating a loss in aggregate output. The key insight from Section 4 is that this feedback effect presents the central bank with a huge dilemma: when the policy rate needs to be raised due to an overheated economy, the central bank cannot raise the rate too high because that would increase the number of bank failures. Banks understand

the central bank's dilemma, and therefore from the start of the game engage in excessive investment, which increases equilibrium financial fragility. Third, at the end of Section 4, we analyze the central bank's problem in an infinitely repeated game. We show in the longer run how the central bank can restrict banks' aggregate investment, lower financial fragility and restore the first best solution by adhering to the socially optimal interest rate. The deviating banking sector will be punished by the central bank's best response, and we characterize the range of central bank discount rates that ensures this cooperative solution.

## 1.2 A dual policy tradeoff

Modern central banks have tended to focus on one policy tool during normal times, the interbank interest rate $r$. However, central banks face a variety of policy objectives. For the purposes of this paper, we consider three policy objectives–inflation, employment or investment, and financial stability. Since these objectives often conflict, the central bank faces a dual policy tradeoff.

The first tradeoff, between inflation and unemployment, is well documented and understood. This basic tradeoff is known as the Phillips curve, first documented by Phillips (1958), then placed in a micro-founded setting by Lucas (1972) and Woodford (2003). The Phillips curve's ramifications for central bank policy are examined by Barro and Gordon (1983). The second tradeoff, between investment and financial stability, has only been recently analyzed, see Cao and Illing (2012); Chollete and Jaffee (2012); and Kashyap and Stein (2012). The crux of this tradeoff is that an interest rate policy which encourages investment has an externality effect of excess credit supply, which in turn increases financial instability. In this paper we analyze both tradeoffs from the perspective of the central bank.

## 1.3 Contributions

Our paper contributes to the literature on optimal monetary policy by formalizing the investment fragility tradeoff faced by a central bank. We then characterize the dual tradeoff from inflation-investment-fragility that plagues monetary policy, in a simple static game. Finally, we extend our results to a dynamic game setting.

The remainder of the paper is organized in the following manner. Section 2 motivates the relationship between financial fragility and investment. Section 3 models the central

bank's tradeoff between fragility and investment. Section 4 studies an extended model that accounts for feedback between the real and financial sectors, and Section 5 is the conclusion.

# 2 Motivation for fragility-investment relation

In order to motivate the dual tradeoff approach to macroprudential policy, we need a functional relation for the second tradeoff[2], between fragility $f$ and aggregate investment liabilities $L$. That is, we require a simple micro-foundation for the function $f = f(L)$. A natural way to do this is to consider the well-documented *fire-sale externality* phenomena of banks: during normal times, banks overinvest. Since every bank does this, during extreme events the whole system is adversely affected.

There are two ways to represent such externalities: via the *likelihood* of extremes, or the *impact* of extreme events. The externality can be an excessively high likelihood of extreme events (Chollete (2012)), or an excessively high level of liabilities during extreme events (Allen and Gale (2007); Kashyap and Stein (2012)). The first method is somewhat general, so we present it in the Appendix. We focus on the liability-based approach below.

## 2.1 Leverage-based fragility measure

This approach models an externality from excessive investment, as in Fisher (1933); Keynes (1936); Allen and Gale (2007); and Kashyap and Stein (2012). We summarize the approach of Kashyap and Stein (2012). Consider a large number of banks that play an infinitely repeated game, in periods denoted by $t$, for $t = 0, 1, ...\infty$. In order to model revelation of information about systemic risk fragility, each period $t$ is further subdivided into three dates $t_0$, $t_1$, and $t_2$. Every bank $i$ holds exogenously determined equity $k^i$, whose value is uniformly distributed over the interval $[0, K]$.

Banks are in the business of investing, which can funded in two ways–short term or long term debt. Specifically, each period the representative bank has an investment opportunity, which, during *normal times,* transforms investment $I^i$ made at initial date $t_0$ into $\theta I^i$ two dates later at $t_2$, for $\theta > 1$. During extreme periods, investment yields no profit, and $\theta = 0$.

[2]For the first tradeoff, the Phillips curve, see Phillips (1958); Lucas (1972); and Woodford (2003).

In order to fund $I^i$, bank $i$ may issue an amount $mI^i$ of short-term debt and $(1-m)I^i$ of long-term debt. The gross interest rates on these two types of debts are $r_1$ and $r_1+\delta$, respectively. Therefore the bank's costs in normal times are $mI^i r_1 + (1-m)I^i(r_1+\delta) = I^i(r_1+\delta-m\delta)$, implying net profits of $I^i(\theta - r_1 - \delta + m\delta)$.

At date $t_1$ there is a public signal of the $t_2$ return from investment. With probability $1-s$ it will be a normal state in $t_2$, with investment return of $\theta I^i$, while with probability $s$ (for systemic) it will be in the crisis state, with investment return of $0$. In a normal state, debtors will roll over the banks' debt, while in a crisis state, debtors refuse to roll over debt. A bank goes bankrupt if its equity value $k^i$ is less than the value of short term debt $mI^i$ at date $t_1$. Since $k^i$ is uniformly distributed over $[0, K]$ the probability that bank $i$ will fail is simply $\frac{mI^i}{K}$. For simplicity, the likelihood of systemic risk is assumed to be linear in aggregate investment $I \equiv \sum_i I^i$. Similarly, the cost $C^i$ that each bankrupt firm imposes on society is also linear, $C^i = \gamma \cdot \sum_i I^i$, for $\gamma > 0$. Under the insolvency of bank $i$, its entire assets will be sold at a depressed price $p$, which applies to all other banks' assets and implies a cost for bank $i$ of $C^i = \gamma \cdot \sum_i I^i$. Thus bank $i$'s costs during extreme periods are given by $\frac{pmI^i\gamma\sum I^i}{K}$.

Such a fire-sale externality implies that each bank does not internalise the cost it imposes on the other banks when it fails. The problem for bank $i$ at the beginning of period t, date $t_0$ is to maximize expected profit $\Pi^i$:

$$\max_{I^i} \Pi^i = I^i(\theta - r_1 - \delta + m\delta) - \frac{pmI^i\gamma\sum_i I^i}{K}. \tag{1}$$

The solution to (1) is given by

$$I^c = (\theta - r_1 - \delta + m\delta)\frac{K}{pm\gamma}, \tag{2}$$

where the $c$ denotes the competitive outcome, and we remove superscript $i$ for simplicity. The socially optimal value maximizes joint profits $\sum_i \Pi^i = \sum_i I^i(\theta - r_1 - \delta + m\delta) - \frac{pmI^i\gamma\sum_i I^i}{K}$. The corresponding optimal investment[3] is given by

$$I^p = (\theta - r_1 - \delta + m\delta)\frac{K}{2pm\gamma}, \tag{3}$$

[3]For further details, see p. 272 of Kashyap and Stein (2012).

where $p$ denotes a Pareto optimum. Similar to the results in the above subsection, the competitive level of investment $I^c$ exceeds the optimum $I^p$.

**Fragility Measure** $f^1$. In this framework, financial fragility $f$ is represented by the probability of bankruptcy due to excess leverage. With a continuum of competitive banks, this probability is the mass of banks with capital less than $mI^c$, that is, $\frac{mI^c}{K}$. We therefore define our main fragility measure $f^1$, as below. The alternative fragility measure $f^0$ is in Appendix A.1.

**Definition 1**. The **leverage-based fragility** measure is $f^1(I) = \frac{mI^c}{K}$.

We summarise the relevance of the fragility measures in Proposition 1, and the inefficiencies that motivate central bank intervention in Proposition 2, below.

**Proposition 1**: *In a competitive banking system, financial fragility $f$ can be represented as a function of equilibrium aggregate liabilities $L = \sum_i l^i$, or as a function of aggregate bank investment $I = \sum_i I^i$.*

**Proposition 2**: *In a competitive banking system with fire-sale externalities, the likelihood-based fragility $f^0$ and the leverage-based fragility $f^1$ are inefficiently high.*

Now that we have established the concept of financial fragility, we utilize it in the remainder of the paper in a game theoretic setup. For simplicity we use the symbol $f$ to denote fragility.

# 3 Modelling tradeoffs between fragility, investment and output

In this section we develop a simple model to establish a theoretical framework, and throughout we focus on describing the second, financial fragility tradeoff. We develop this model further in the succeeding section. Since the first tradeoff of unemployment and inflation is well understood, we utilise existing results from that literature directly. The setting is one of common knowledge, where the central bank knows the payoff functions of banks.[4] The central bank interacts with banks and entrepreneurs, and plays the game repeatedly. Such repetitions are cumbersome to represent, hence we simplify computation by focusing on a

[4]For earlier work, see Kydland and Prescott (1977), and Barro and Gordon (1983). For details on repeated games, see Fudenberg and Tirole (1991); Gibbons (1992); and Mas-Colell et al. (1995).

stage game, which is a one-shot version of the repeated game. If the full game is played $T$ times, then the payoff for each player is simply the discounted sum of the payoffs in each stage game.

**Notation**. This paper uses several types of notation in the next section. In order to ease the reader's understanding of the paper, we present below the most frequently used parameters and symbols.

- $a$: Cost to central bank of missing inflation or interest target
- $b$: Cost to central bank of missing investment target
- $\beta$: Phillips curve-based cost to central bank of missing interest target
- $r^*$: Target interest rate
- $I^E - I^*$: Gap between optimal and target investment
- $I^*$: Optimal investment
- $K$: Total equity in the banking system
- $\gamma$: Multiplier for systemic costs relative to $I$, which measures fire-sale costs
- $m$: Fraction of short-term debt in banking system, which measures system illiquidity
- $\varepsilon_y$: Output shock
- $G\left(E\left[\varepsilon_y^{i*}\right], \hat{r}\right)$: the number of banks that are expected to fail

**Setting for inflation-unemployment tradeoff**. The background here is well studied in the literature, so we comment only briefly. There are three main actors: employers, workers and the central bank. The central bank sets the interest rate, which determines the money supply and inflation rate. In order to capture anticipated effects, employees and workers set an imperfectly indexed wage. The central bank can set the interest rate lower than expected, which will yield reduced real wages and allow employment to increase, but also result in higher inflation.

**Setting for fragility tradeoff.** There are three actors in this part: a large number of firms, a large number of financial intermediaries ('banks') , and a central bank. The setting is a

sequential-move game where (time $t$) the entrepreneurs and banks negotiate on the level of risky loans to undertake, after which (time $t+1$) the central bank chooses the interbank interest rate and whether to bail out or provide liquidity to some banks that are distressed. Knowing the central bank's strategy, investors and banks will anticipate this in period $t$ when deciding on their optimal strategies. In particular, if the central bank plans to rescue some distressed banks and lower interest rates, this will induce both banks and investors to demand and supply excessive risky loans, thereby increasing systemic risk. At the same time, it is beneficial to lower interest rates and rescue some banks, since this encourages investor confidence, thereby providing a spur to innovative investment that might not otherwise occur.

Consequently, when deciding on bailouts, liquidity requirements and low interest rates, the policymaker faces a *fragility tradeoff*. This tradeoff features costs of *financial fragility* and systemic instability due to an excess of risky loans, and benefits of *enhanced investment*, to the extent that low rates and rescuing of some distressed banks cause result in more worthwhile innovative projects than anticipated being undertaken. The above interaction is repeated an infinite number of times, $T = \infty$. In order to focus on this long-run tradeoff, we analyse a reduced-form version of the repeated interaction, in the following stage game. We model financial fragility $f$ as an index that the central bank can affect directly, through its decisions on how many distressed banks to rescue, as well as the magnitude and duration of interest rate cuts.[5] For simplicity, $f$ can denote the number (or fraction) of banks that will be allowed to fail during an extreme event. In the stage game below, we combine the output-inflation tradeoff with this relatively recently studied fragility tradeoff.

## 3.1 A base model without feedback effects

First, banks and companies form expectations of financial fragility, $\hat{f}$, and inflation, $\hat{\pi}$, respectively. Second, the policymaker assesses expectations and chooses the actual fragility $f$ and inflation $\pi$. Banks and companies receive payoffs of $-(f-\hat{f})^2$ and $-(\pi-\hat{\pi})^2$, respectively. That is, banks (companies) desire to anticipate fragility (inflation) as accurately as possible, achieving their maximal payoff when $f=\hat{f}$ ($\pi=\hat{\pi}$). It is the policymaker's desire for fragility and inflation to be close to zero, but also for investment $I$ and employment $y$ to be at efficient[6] levels $I^*$ and $y^*$.

[5] More generally, $f$ will be a vector $[f, r]$ of the fraction of banks rescued and interest rates.

[6] Perfect financial market stability may not be efficient, if it means no investment is undertaken.

We let the real rate $\bar{r}$ be fixed, so that according to the Fisher relation $\pi = r_1 - \bar{r}$, it is sufficient to focus on the nominal short rate $r_1$ instead of inflation. The policymaker's loss function is therefore

$$\min_{r_1} L = (y - y^*)^2 + a\,(r_1 - r^*)^2 + b\,[f(I) - f^*]^2\,, \tag{4}$$

where $a > 0$ and $b > 0$ reflect costs to the central bank of missing its targets. All terms in (4) depend on the interest rate $r_1$, in a manner that we now make explicit via the inflation-output and fragility-investment tradeoffs, below. The inflation-output relationship is defined by the Phillips curve,

$$y = \alpha y^* + \beta\,(r_1 - \hat{r} + \varepsilon_y)\,, \tag{5}$$

where $\alpha > 0$ and $\beta > 0$, and $\varepsilon_y$ is a zero-mean shock to output, $\varepsilon_y \sim [0, \sigma_y^2]$. We model the relationship between financial investment and the short-term interest rate by building on the Kashyap and Stein (2012) specification, from equation (2) above:

$$I = (\theta - r_1 - \delta + m\delta + \varepsilon_f)\,\frac{K}{pm\gamma}.$$

Note that, unlike Kashyap and Stein (2012), we account for random financial shocks via the term $\varepsilon_f$, $E[\varepsilon_f] = 0$. Financial fragility is defined as the cost of a fire sale, namely $f\,(I) = \gamma\overline{I} = (\theta - r_1 - \delta + m\delta + \varepsilon_f)\,\frac{K}{pm}$. And the socially optimal financial fragility $f^*$ is $f^* = \gamma I^* = \gamma\,(\theta - r_1 - \delta + m\delta + \varepsilon_f)\,\frac{K}{2pm\gamma}$, or

$$f^* = (\theta - r - \delta + m\delta + \varepsilon_f)\,\frac{K}{2pm}. \tag{6}$$

Therefore, the central bank's decision problem in (4) can be rewritten in terms of $r_1$, using (5) and (6), as

$$\min_{r_1} L = [\alpha y^* + \beta\,(r_1 - \hat{r} + \varepsilon_y) - y^*]^2 + a\,(r - r^*)^2 + b\left[(\theta - r_1 - \delta + m\delta + \varepsilon_f)\,\frac{K}{2pm}\right]^2. \tag{7}$$

This objective is solved in Appendix A.2 to yield expressions for optimal interest rate $r^c$ and output $y$:

$$r^c = \frac{1}{B}\left[(1-\alpha)\beta y^* + ar^* + (\theta - \delta + m\delta)C\right] - \frac{1}{A}\left(\beta^2\varepsilon_y - C\varepsilon_f\right) \tag{8}$$

and

$$y = \alpha y^* + \frac{B}{A}\beta\varepsilon_y + \frac{C}{A}\beta\varepsilon_f = \alpha y^* + \frac{\beta}{A}\left(B\varepsilon_y + C\varepsilon_f\right), \tag{9}$$

where the coefficients $A, B$ and $C$ capture the relative weights attached to various forms of over- or undershooting: $A = a + \beta^2 + b\left(\frac{K}{2pm}\right)^2$, $B = a + b\left(\frac{K}{2pm}\right)^2$, and $C = b\left(\frac{K}{2pm}\right)^2$.

The expected welfare loss $E[L^n]$ is also computed in Appendix A.3, Equation (33), to be

$$\begin{aligned} E[L^n] &= (\alpha-1)^2 y^{*2} + a\left[\frac{(1-\alpha)\beta y^* + (\theta - \delta + m\delta - r^*)C}{B}\right]^2 \\ &\quad + C\left[\frac{(1-\alpha)\beta y^* + ar^* - (\theta - \delta + m\delta)a}{B}\right]^2 \\ &\quad + \frac{B^2 + \beta^2(a+C)}{A^2}\beta^2\sigma_y^2 + \frac{(a+\beta^2)^2(C^2+C)}{A^2}\sigma_f^2 + 2C\beta^2\frac{B+\beta^2}{A^2}\rho_{y,f}, \end{aligned}$$

where $\sigma_y^2$, $\sigma_f^2$ and $\rho_{y,f}$ are the real shock variance, financial shock variance, and correlation of the two shocks, respectively. In contrast, the cooperative, expected welfare loss $E[L^c]$ under perfect anchoring is computed in equation (34) of the Appendix to be

$$E\left[L^c\right] = E\left[\left[(\alpha-1)\,y^* + \beta\varepsilon_y\right]^2 + \left(\frac{K}{2pm}\varepsilon_f\right)^2\right] = (\alpha-1)^2\,y^{*2} + \beta^2\sigma_y^2 + \frac{C}{b}\sigma_f^2.$$

**Comments on the solution**. Two aspects of the above solutions appear cogent. First, the volatility of shocks to both the financial and real sectors matter for optimal loss of the central bank. Second, and perhaps more novel, the covariance of real and financial shocks, $\rho_{y,f}$, directly affects the central bank's expected loss. In particular, it is not enough for the central bank to focus on real or financial shocks: it must *also* identify the dependence

between real and financial shocks. We summarize this latter observation in a proposition, below.

**Proposition 3:** *In a dual tradeoff model, the central bank's expected loss increases monotonically with the dependence between financial and real shocks.*

## 3.2 Explanatory power and novel implications

Our theoretical framework provides a simple method to incorporate financial and real considerations, simultaneously, in a tractable form for central bank policy. It also offers a potential explanation of several interesting phenomena. In particular, a large, positive real economy shock $\varepsilon_y$ shifts the production function up and raises marginal productivity. Since marginal productivity determines marginal return on capital, banks increase their supply of credit. Consequently real shocks have financial effects, and our model explains the empirically observed procyclicality of credit.

Similarly, a large negative financial shock $\varepsilon_f$ reduces the amount of credit $I$ that banks supply. This, in turn, reduces firms' investment and therefore results in output contraction. Hence, our model explains the empirically observed occasional spillovers from financial crises to the real economy. These spillovers will be exacerbated, the larger the correlation between real and financial shocks.

**Implications**. Our model also has some novel implications. In particular, it suggests a dual tradeoff between enhancing economic output and maintaining financial fragility. The reason is that the central bank's policy instrument $r_1$, the short rate, appears positively in the real economy, equation (10) but negatively in the financial economy, equation (11). Central banks will experience the bite of this tradeoff, the larger the investment parameter $\tau$ and the larger the correlation between real and financial shocks.

**Potential policy tools**. In light of the above analysis, effective policy instruments in the face of this dual tradeoff must decrease $\tau$ or reduce real-financial correlations $\rho_{y,f}$. One possible approach is for central banks to market *correlation-indexed bonds*, i.e. securities that are indexed to the level of dependence between the real and financial sector.

# 4 A model with feedback between banks and firms

Thus far, we have modelled a simple link between banks and firms, via the correlation of shocks, $\rho_{f,y}$. More realistically, banks create credit $I$ that firms use to produce real output $y$. Hence the central bank's choice of interest rate $r$ affects bank credit, which affects real output. In turn, the real output is observed, which affects the central bank's choice of interest rate.

*Feedback from financial Sector to the real economy*. This framework allows us to account for the real effect of financial crises, since they reduce $I$ via fire sales.[7] In turn, the reduced $I$ decreases real output. Hence the Phillips curve is expanded to contain another element, namely $I$.

*Feedback from the Real Economy to the Financial Sector*. Our framework also allows us to account for feedback from the real economy, because the banking sector decide the amount of credit $I$ available depending on its expected marginal return. This marginal return, $\theta$ from the Kashyap and Stein (2012) specification, is equal in competitive markets to the marginal productivity of capital in the real economy.

We therefore revisit the Phillips curve and financial investment equations from the previous section. In light of the above observations, we propose to account for financial frictions in the real economy via an investment-augmented Phillips curve:

$$y = (I - I^*) + y^* + \beta \left( r_1 - \hat{r} + \varepsilon_y \right), \tag{10}$$

where $I^*$ represents investment under the natural rate of output $y^*$. We retain our previous specification of the relationship between financial investment and the short-term interest rate from Section 3 above:

$$I = \left( \theta(y) - r_1 - \delta + m\delta + \varepsilon_f \right) \frac{K}{pm\gamma}. \tag{11}$$

Hence the financial sector and real sector are linked in three ways. First, as discussed above, the two equations (10) and (11) are linked via the function $\theta(y)$, which represents the marginal return on bank capital in (11), and is determined by the marginal product of $y$ from equation (10). Second, they are linked by the available credit for investment $I$, which

[7] This approach is similar to the credit transmission channel of Bernanke and Gertler (1989).

is a component of real output. Finally, they are linked via the short rate $r_1$, which is set by the central bank in response to output shock $\epsilon_y$.

## 4.1 Formalisation of the model

Consider a large number of banks and firms that play an infinitely repeated game with a central bank, in periods denoted $t$, for $t = 0, 1, \ldots, +\infty$. Each period $t$ is further subdivided into three dates $t_0$, $t_1$, and $t_2$. The timing of the model is summarised in Figure 1.

Figure 1: Timeline of events in the model

| $t_0$ | $t_1$ | $t_2$ |
|---|---|---|
| Bank $i$ provides credit $I^i$ to firms, based on its expectation of central bank's policy rate $\hat{r}$;<br><br>The credit $I^i$ is financed by both short-term and long-term debt contracts, with interest rates $r_1$ and $r_1 + \delta$, respectively;<br><br>The firms use the credit to buy capital input, and start production. | The shock to the production sector $\varepsilon_y$ is revealed;<br><br>The central bank decides its policy rate $r^E$, based on the shock, inflation, and cost of bankruptcies in the financial sector;<br><br>The bank becomes insolvent if its equity becomes negative, and its assets are sold at depressed prices;<br><br>If the bank survives, it repays the existing short-term debt contracts, and rolls over part of the debt by issuing new short-term debt with interest rate $r^E$. | The surviving banks earn a return from their investment in firms, and repay both the short- and long-term debt contracts;<br><br>The failed banks are liquidated, and their debts cleared. |

**Stage** $t_0$: Each bank $i$ decides its loan or credit supply $I^i$ to firms, given its expectation on the central bank's short-term (one-period) policy rate $\hat{r}$ and the firms' expected output. To finance $I^i$, bank $i$ may issue an amount $mI^i$ of short-term debt and $(1-m)I^i$ of long-term debt, the later of which lives from $t_0$ to $t_2$. The short-term debt is rolled over at $t_1$ with expected rate $\hat{r}$. With similar notation to the previous section, the gross interest rates on these two types of debt are $r_1$ and $r_1 + \delta$, respectively, where compounding yields $r_1 = \hat{r}^2$.

The firms use the loan $I^i$ as an input. The return from the bank's investment in firms, $\theta$, is determined by the expected aggregate output $E[y]$, and $\theta(E[y])$ is an increasing function of $E[y]$. The bank's expected costs are therefore $C(m, r_1) = mI^i r_1 + (1-m)I^i(r_1 + \delta)$, implying expected net profits of $\theta(E[y])I^i - C(m, r_1)$, if it survives till $t_2$. Without much loss of generality, assume that $\theta(E[y]) = \omega E[y]$.

**Stage** $t_1$: At $t_1$ there is a publicly observed signal of the aggregate shock on date $t_2$'s return from firms. The shock is denoted by $\varepsilon_y \sim N\left(0, \sigma^2_{\varepsilon_y}\right)$. After the shock, the central bank sets its policy rate $r^E$, based on its loss function:

$$\min_{r^E} \quad L = (y - y^*)^2 + a\left(r^E - r^*\right)^2 - \gamma \sum_i I^i G\left(E\left[\varepsilon_y^{i*}\right], \hat{r}\right),$$

where the superscript $E$ denotes equilibrium. The term $G\left(E\left[\varepsilon_y^{i*}\right], \hat{r}\right)$ denotes the number of banks expected to fail, which depends on both real and monetary shocks, as explained later in this section. This policy rate will affect aggregate output at $t_2$. Actual output is determined by the input $I$, subject to price stickiness. We define this inflation-output relationship by an "augmented" Phillips curve

$$y\left(r^E, \varepsilon_y\right) = (I - I^*) + y^* + \beta\left(r^E - \hat{r} + \varepsilon_y\right).$$

The investment under natural output is denoted by $I^*$. Without loss of generality, the natural output is normalised as $y^* = I^*$.

In a state when the bank's return is higher than $r^E$, debtors will roll over banks' debt, while in other states, debtors refuse to roll over debt. The bank fails if its equity value $k^i$ (uniformly distributed on $[0, K]$) is less than the value of short term debt $mI^i$ net the expected value of its assets at date $t_2$. If the bank fails, all of its assets will liquidated in a fire sale, subject to the fire sale cost.

**Stage** $t_2$: If the bank survives, the firms will repay the loans, and its debtors withdraw.

The equilibrium of the model is featured by the following set of functions:

(1) Given $r^E$, the firms' aggregate output at $t_2$ is $y = (I - I^*) + y^* + \beta\left(r^E - \hat{r} + \varepsilon_y\right)$, implying that the gross return to the bank's investment is $\theta = \omega y$.

At $t_1$, the bank will fail if $mI^i\hat{r}r^E + (1-m)\, I^i\,(r_1 + \delta) - \omega y I^i > k^i$. Define $y^{i*}$ and $\varepsilon_y^{i*}$ such that $mI^i\hat{r}r^E + (1-m)\, I^i\,(r_1 + \delta) - \omega y^{i*} I^i = k^i$, and $y^{i*} = (I - I^*) + y^* + \beta\left(r^E - \hat{r} + \varepsilon_y^{i*}\right)$. Then the ex post probability that the bank will fail is $G\left(\varepsilon_y^{i*}\right)$.

(2) If the bank fails at $t_1$, its assets will be sold at the depressed price. The fire sale cost is $\gamma \sum_i I^i$.

At $t_0$ the bank's optimal decision on investment is determined by maximising its profit

$$\max_{I^i} L_B^i = \omega(E[y])I^i - C(m, r_1) - \gamma \sum_i I^i G\left(E\left[\varepsilon_y^{i*}\right], \hat{r}\right),$$

where the $B$ superscript denotes the bank, $G\left(E\left[\varepsilon_y^{i*}\right], \hat{r}\right)$ denotes the ex ante probability of insolvency, and rational expectations imply that $\hat{r} = E\left[r^E\right]$.

At stage $t_0$ the following relation holds: $\omega(E[y]) = \omega E\left[(I - I^*) + y^* + \beta\left(r^E - \hat{r} + \varepsilon_y\right)\right] = \omega I$, where $I$ is the aggregate credit supply, taken as given for each individual bank.

Given the bank's expectations for central bank policy, $\hat{r}$, the bank's expected probability of becoming insolvent[8] at $t_1$ is $G\left(E\left[\varepsilon_y^{i*}\right], \hat{r}\right) = \frac{mI^i\hat{r}^2 + (1-m)I^i(r_1+\delta) - \omega II^i}{K}$. Therefore, the bank's optimal decision problem at $t_0$ is

$$\max_{I^i} L_B^i = \omega II^i - mI^i r_1 - (1-m)I^i(r_1+\delta) - \gamma \sum_i I^i \frac{mI^i\hat{r}^2 + (1-m)I^i(r_1+\delta) - \omega II^i}{K}.$$

Note that since we are in a two period model, the short rate $\hat{r}$ and long rate $r_1$ are related by $r_1 = \hat{r}^2$.

**Bank's competitive equilibrium outcome** $I^E$. Under the above assumptions, the bank's first order condition yields

$$\frac{\partial L_B^i}{\partial I^i} = \omega I - mr_1 - (1-m)(r_1+\delta) - \gamma I \frac{m\hat{r}^2 + (1-m)(r_1+\delta) - \omega I}{K} = 0.$$

This condition is satisfied by the equilibrium investment $I^E$ such that

$$\omega I^E K - r_1 K - K\delta + Km\delta - \gamma I^E r_1 - \gamma I^E \delta + \gamma I^E m\delta + \gamma (I^E)^2 \omega = 0, \tag{12}$$

where the superscript $E$ denotes equilibrium. However, the bank does not take into account the fire-sale externality it imposes to the entire financial system, which leads to excessive investment.

[8] Our formulation of $G()$ therefore generalizes the Kashyap and Stein (2012) framework, since the short-rate enters $G()$. This formalises the notion that the central bank can affect the likelihood of default.

**Bank's social optimum** $I^p$. We now account for fire-sale externalities. If a planner forces the bank to consider fire-sale costs, the bank will solve the following decision problem:

$$\max_{I^p} L_B^p = \omega(E\left[y(I^p)\right])I^p - C\left(m, r_1\right) - \gamma \sum_i I^p G\left(\varepsilon_y^{i*}, \hat{r}\right)(i),$$

where the superscript $p$ denotes planner. This maximization problem can be rewritten

$$\max_{I^p} L_B^p = \omega(I^p)^2 - mI^p r_1 - (1-m)I^p\left(r_1+\delta\right) - \gamma I^p \frac{mI^p\hat{r}^2 + (1-m)\,I^p\left(r_1+\delta\right) - \omega(I^p)^2}{K}.$$

The first order condition yields

$$\frac{\partial L_B^p}{\partial I^p} = 2\omega I^p - mr_1 - (1\text{-}m)\left(r_1+\delta\right) - \frac{2\gamma I^p m\hat{r}^2 + 2\gamma I^p\left(1-m\right)\left(r_1+\delta\right) - 3\gamma\omega(I^p)^2}{K} = 0,$$

or the $I^p$ such that

$$2I^p\omega K - r_1 K - K\delta + Km\delta - 2I^p\gamma r_1 - 2I^p\gamma\delta + 2I^p\gamma m\delta + 3(I^p)^2\gamma\omega = 0. \qquad (13)$$

As summarised in Proposition 4 below, the bank's equilibrium investment is larger than the social optimum, the latter of which accounts for excess fragility.

**Proposition 4: Inefficiency of bank's competitive equilibrium investment.**
*Part A: In market equilibrium, the bank's investment is larger than the planner's solution, i.e., $I^E > I^p$. Moreover, $\sqrt{3}I^p < I^E < 2I^p$.*
*Part B: In market equilibrium, liquidity risk provides some market discipline, but investment is still excessive.*

Part A is straightforward, and provides an upper-bound on the magnitude of over-investment. Part B of the proposition says that a higher share of short-term debt (higher $m$) reduces banks' investment. Equivalently, liquidity risk restricts banks' risk-taking in the market equilibrium, thereby providing some market discipline. Nevertheless, due to the fire-sale externality, market discipline is not sufficient. Compared with the planner's solution, the competitive equilibrium still features excessive investment.

**Central bank's decision.** We now derive the central bank's optimal strategy for short rates. To establish the theoretical framework, we focus on the case where banks choose

the fragility-based excessive credit supply $I^E$ solved above.[9] The central bank's optimal decision at $t_1$, after observing $\varepsilon_y$, is to solve the following program:

$$\min_{r^E} L_C = (y - y^*)^2 + a\left(r^E - r^*\right)^2 + b\gamma \sum_i I^i G\left(\varepsilon_y^{i*}\right)(i),$$

where the subscript $C$ denotes the central bank. The actual output $y$, given the announced $r^E$, is $(I - I^*) + y^* + \beta\left(r^E - \hat{r} + \varepsilon_y\right)$, and the return of the banks' investment is therefore $\omega y$.

From the Appendix A.4, the central bank's optimal short rate $r^E$ is proved to satisfy

$$r^E = \frac{2aK}{2aK + b\gamma(I^E)^2 m} r^* - \frac{2\beta K\left(I^E - I^*\right)}{2aK + b\gamma(I^E)^2 m} + \frac{b\gamma\left(I^E\right)^2 \omega\beta}{2aK + b\gamma(I^E)^2 m} - \frac{\beta^2}{a + \beta^2}\varepsilon_y. \quad (14)$$

The solution for $r^E$ in (14) can be related to extant models. For example, as in Barro and Gordon (1983), $r^E$ should partially respond to demand shocks, which is intuitive. There are, however, some important differences between $r^E$ and the short rate obtained by a traditional central bank concerned with inflation targeting. We discuss these differences below.

## 4.2 Implications for macroprudential policy

In conventional inflation-targeting frameworks, the central bank's decision problem does not account for fire sales. Specifically, the central bank's problem is

$$\min_{r^0} L = \left[\left(I^E - I^*\right) + \beta\left(r^0 - \hat{r} + \varepsilon_y\right)\right]^2 + a\left(r^0 - r^*\right)^2,$$

in which $r^0$ is the policy rate set by a pure inflation targeting central bank. The first order condition yields

$$\frac{\partial L}{\partial r^0} = 2\beta\left[\left(I^E - I^*\right) + \beta\left(r^0 - \hat{r} + \varepsilon_y\right)\right] + 2a\left(r^0 - r^*\right) = 0,$$

[9] Given the symmetric structure of the problem, the central bank will solve the same program when banks supply the socially optimal credit $I^p$, except that everywhere in the program it would replace $I^E$ with $I^p$.

or

$$r^0 = \frac{ar^* + \beta^2\hat{r} - \beta^2\varepsilon_y - \beta\left(I^E - I^*\right)}{a + \beta^2}.$$

The consistency condition $E\left[r^0\right] = \hat{r}$ implies that $\hat{r} = E\left[r^0\right] = E\left[\frac{ar^* + \beta^2\hat{r} - \beta^2\varepsilon_y - \beta\left(I^E - I^*\right)}{a+\beta^2}\right] = \frac{ar^* + \beta^2\hat{r} - \beta\left(I^E - I^*\right)}{a+\beta^2}$, or $\hat{r} = r^* - \frac{\beta}{a}\left(I^E - I^*\right)$. By substituting this condition into the expression for $r^0$, we obtain the optimal short rate $r^0$ for an inflation-targeting central bank:

$$r^0 = r^* - \frac{\beta}{a}\left(I^E - I^*\right) - \frac{\beta^2}{a + \beta^2}\varepsilon_y. \tag{15}$$

Recall that for a central bank which takes financial stability into account, the optimal short rate $r^E$ from (14) satisfies

$$r^E = \frac{2aK}{2aK - b\gamma(I^E)^2 m}r^* - \frac{2\beta K\left(I^E - I^*\right)}{2aK - b\gamma(I^E)^2 m} + \frac{b\gamma\left(I^E\right)^2\omega\beta}{2aK + b\gamma(I^E)^2 m} - \frac{\beta^2}{a + \beta^2}\varepsilon_y.$$

**Observations on the solution.** By comparing the solution for $r^0$ in (15) to the one for $r^E$ above, we can see the role of financial stability in the central bank's decision making. In particular, we discern two effects.

The first is the *stability effect*: in the first term of $r^E$, $\frac{2aK}{2aK - b\gamma(I^E)^2 m} < 1$, which indicates that the central bank's "stability target" interest rate should be lower, once it cares about financial fragility. Thus, the central bank has the incentive to lower the policy rate in order to reduce the insolvency rate in the banking sector.

The second is a *discipline effect*: in the second term of $r^E$, $-\frac{2\beta K\left(I^E - I^*\right)}{2aK - b\gamma(I^E)^2 m} + \frac{b\gamma\left(I^E\right)^2\omega\beta}{2aK + b\gamma(I^E)^2 m} > -\frac{\beta}{a}\left(I^E - I^*\right)$, which indicates that the central bank's target rate should be higher, in order to restrict excessive investment. This effect is in line with suggestions by John Taylor and others, who argued in 2008 that a central bank which is concerned about financial stability should set higher policy rates (relative to the Taylor rule) in normal times, in order to deter build-up of financial imbalances. Here we do see such an incentive.

More broadly, these findings extend the literature on central bank independence (e.g. Rogoff (1985)). This literature argues that in order to achieve price stability, monetary policy should be delegated to a "conservative" central bank that only cares about inflation or, in the context of our model, $a \to +\infty$. In such a case, our model implies the following condition: $\lim_{a \to +\infty} r^0 = r^E = r^*$.

What does our solution imply about dynamics? There is obviously the potential for moral hazard. If banks have a systematically lower expectation $\hat{r}$ and higher $I^i$ at $t_0$, the central bank will be, ex post, forced to follow with a lower $r^E$ if it has a low discount factor. We discuss this issue further in the repeated game formulation below.

## 4.3 An impossibility result

Our framework may be valuable to clarify the limitations on central banking imposed by attempts to control both inflation and fragility. Intuitively, the only time a central bank's dual targets of inflation control and stability are compatible is when the policy rates $r^E$ and $r^O$ are equal. The central bank can use the same rate to control both inflation and targeting only if $r^E - r^O$ is identically zero. We solve this condition in the Appendix, which yields the following proposition:

**Proposition 5: General impossibility of both fragility control and inflation commitment**. *Even if a central bank is independent, it is generally unable to control simultaneously fragility and inflation. Such control is only possible in the knife-edge case of* $r^* = \frac{\beta}{a}(I^E - I^*) + \frac{\omega\beta}{m}$*, i.e. where the target rate is a specific linear function of the investment gap.*

## 4.4 Infinitely repeated game

Now consider the following infinitely repeated version of the game above. One solution above sets the vector $[I^E, r^E]$ such that, given banks' expectations, the marginal cost and benefits to the central bank from surprise investment offset each other. This outcome is subgame perfect, since the central bank is expected to allow a positive amount of fragility and indeed does so. However, the central bank would be better off if it could commit to a socially optimal scenario of $r^p$ (where $r^p \neq r^E$) that results in the optimal credit supply $I^p$. This is an inherent inconsistency problem, since the policymaker and banks have an

incentive to deviate from the socially optimal investment and short rate $[I^p, r^p]$. A simple way to assess the implications of such deviations is to examine an infinite horizon game.

**Infinite game formulation**. Let policymakers, firms and banks share a common discount factor $d$. To solve this game, we shall clarify the player payoffs and focus on pure strategies, then derive conditions under which $r = \hat{r} = r^p$ and $I = I^p$ every period, in a subgame perfect Nash equilibrium.

**Payoff functions for central bank and banks**. Denote the banks' choice of credit supply in equilibrium and under the social (planner) optimum as $I^E$ and $I^p$, respectively. Then the central bank's payoff as a function of short rates and the banks' credit supply $I^i \in \{I^E, I^p\}$, from the previous section, is

$$\begin{aligned} L_C(r^c, \hat{r}, I^i) &= \left[\left(I^i - I^*\right) + \beta\left(r^c - \hat{r} + \varepsilon_y\right)\right]^2 + a\left(r^c - r^*\right)^2 \\ &\quad + b\gamma I^i \frac{mI^i \hat{r} r^c + (1-m)\, I^i\,(r_1 + \delta) - \omega\left[I^i + \beta\left(r^c - \hat{r} + \varepsilon_y\right)\right] I^i}{K}, \end{aligned} \tag{16}$$

where $r^c \in \{r^E, r^p\}$. The bank's payoff function is

$$L_B(\hat{r}, I^i) = \omega I I^i - mI^i r_1 - (1-m) I^i\,(r_1 + \delta) - \gamma \sum_i I^i \frac{mI^i r_1 + (1-m)\, I^i\,(r_1 + \delta) - \omega I I^i}{K}. \tag{17}$$

We shall use the payoffs in (16) and (17) to compute optimal strategies and credible punishments for deviations.

**Strategies for the central bank and bank**. Given a bank's credit supply $I^i \in \{I^E, I^p\}$, the central bank's optimal interest rate strategy, from the previous section, is given by

$$r^c = \frac{2aK}{2aK + b\gamma (I^i)^2 m} r^* - \frac{2\beta K (I^i - I^*)}{2aK + b\gamma (I^i)^2 m} + \frac{b\gamma \left(I^E\right)^2 \omega\beta}{2aK + b\gamma (I^E)^2 m} - \frac{\beta^2}{a + \beta^2}\varepsilon_y. \tag{18}$$

Also from the previous section, the bank's equilibrium strategy for investment is $I^E$ such that the following relation holds:

$$\omega I^E K - K r_1 - K\delta + Km\delta - \gamma I^E r_1 - \gamma I^E \delta + \gamma I^E m\delta + \gamma (I^E)^2 \omega = 0. \tag{19}$$

By contrast, the socially optimal investment $I^p$ allows for fire-sale externalities, and $I^p$ is such that the following relation holds:

$$2\omega I^p K - K r_1 - K\delta + Km\delta - 2\gamma I^p r_1 - 2\gamma I^p \delta + 2\gamma I^p m\delta + 3\gamma (I^p)^2 \omega = 0. \qquad (20)$$

The central bank wishes to encourage socially optimal investment $I^p$ defined in (20), and stop intermediary banks from doing $I^E$ in (19). To accomplish this objective, the central bank can choose a trigger strategy, which sets $r^c = r^p$ computed from (18), as long as the intermediaries behave and choose credit supply equal to $I^p$. Otherwise, the central bank assumes banks are going to opt for their most profitable deviation of $I^E$ and sets $r$ equal to the best response $r^E$, computed from (18). In our model, the banks' expectations are correct on average, so $I^E$ and $I^p$ are in turn best responses to $r^E$ and $r^p$, respectively. Thus the central bank's strategy is supported in a subgame perfect Nash equilibrium.

The repeated game involves evaluating stage game payoffs to various strategies from the central bank and bank. These payoffs may be represented in the following matrix, Table 1. In Table 1, $L^p_C(\cdot)$ denotes the central bank's payoff in the socially optimal setting (where financial fragility is minimised), and $L^E_C(\cdot)$ denotes the central bank's payoff in the inefficient equilibrium, where financial fragility is a problem. The tilde denotes payoffs from a deviation. Thus $\tilde{L^E_B}$ is what the bank obtains by deviating to the equilibrium investment $I^E$.

Table 1: Stage Game Payoffs to Central Bank and Intermediary Bank

| | | **Bank** | |
|---|---|---|---|
| | | *Cooperate*: $r = r^p$ | *Deviate*: $r = r^E$ |
| **Central bank** | *Cooperate:* $I = I^p$ | $(L^p_C, L^p_B)$ | $(\tilde{L^E_C}, \tilde{L^E_B})$ |
| | *Deviate:* $I = I^E$ | $(\tilde{L^p_C}, \tilde{L^E_B})$ | $(L^E_C, L^E_B)$ |

More specifically, we consider two types of subgames that the central bank faces. The first type is one in which banks have chosen the socially optimal credit supply $I^p$ and the central bank sets the short rate to $r^c$ in (18), for the present and for all previous periods. The second

type is any one in which intermediary banks have deviated by choosing the individually rational but fire-sale provoking credit level $I^E$. The central bank and intermediary bank share a common discount factor $d$.

*Banks:* In the first period, banks hold the expectation $\hat{r} = r^p$, and supply the socially optimal credit $I^p$. In subsequent periods they expect $\hat{r} = r^p$, if and only if all prior credit supply was $I^p$ and short rates were actually $r^p$. Otherwise, bank expectations $\hat{r}$ are set equal to $r^E$, which we defined as optimal for the central bank if it wants to punish banks for excess credit supply in the stage game.

*Central Bank:* The central bank chooses $r = r^p$ if and only if current bank and firm expectations satisfy $\hat{r} = r^p$, all previous expectations have been $\hat{r} = r^p$, and all previous actual investment and short rates have been $I^p$ and $r^p$. Otherwise, the central bank assumes banks are building up excess credit $I^E$ and chooses $r^c = r^E$ from its best response (18).

**Infinite game equilibrium.** Let banks' first-period expectations be $\hat{r} = r^p$. Further, as in Table 1, let the central bank's loss function be $L_C^p(r^p, I^p)$. Now, given the bank's strategy, the central bank can focus on two possible paths of financial fragility. The first path involves $\hat{r} = r^p$, which results the next period in $\hat{r} = r^p$ and $I^p$. Consequently the central bank makes the same decision in the next period. The second alternative involves expectations $\hat{r} = r^E$ and credit supply $I = I^E$. This results in $(I^E, r^E)$ forever.

The payoffs from these two strategies are as follows: The first strategy ($r = r^p$ in this period), yields the central bank a payoff of $L_C^p(r^p, I^p)$ forever, for a total of $\frac{1}{1-d} L_C^p(r^p, I^p)$. The second strategy ($r = r^E$ this period) yields the central bank $\tilde{L_C^p}(r^p, I^p)$ this period, then $L_C^E(r^E, I^E)$ forever, for a total of $\tilde{L_C^p}(r^p, I^p) + \frac{d}{1-d} L_C^E(r^E, I^E)$. Therefore the central bank's strategy is a best response to the banks' strategy if the following condition holds: $\frac{1}{1-d} L_C^p(r^p, I^p) \geq \tilde{L_C^p}(r^p, I^p) + \frac{d}{1-d} L_C^E(r^E, I^E)$. This condition can be rewritten as

$$d \geq \frac{\tilde{L_C^p}(r^p, I^p) - L_C^p(r^p, I^p)}{\tilde{L_C^p}(r^p, I^p) - L_C^E(r^E, I^E)}, \tag{21}$$

which has a ready economic implication. Economically speaking, the central bank finds it optimal not to deviate from its commitment to minimise future financial fragility, if and only if its discount rate (i.e. its willingness to sacrifice present pain for future gain from enforcing low fragility) is large enough.

### 4.5 Discount rate, investment and policy surfaces

In order to illustrate our results visually, we depict the threshold discount rate $d$, aggregate investment $I$, and the optimal policy rate in traditional and fragility-based settings, i.e. equations (19) and (20), respectively. This communicates the difference between our approach and previous research in an intuitive manner.

Details of our calibration are in the Appendix.[10] The surface for the threshold discount rate is displayed in Figure 2. Intuitively, $d$ is more likely to be below 1 when the weight of financial stability gets higher in the central bank's loss function. We show results of our investment calibration in Figure 3. Evidently investment is too low and unresponsive when the central bank only targets inflation. Figure 4 displays the optimal short rate for a central bank that cares about fragility or only inflation. The yellow surface shows that the optimal short rate is unresponsive to fragility concerns, which is straightforward. The dark blue surface corresponds to a relatively large $m$ (illiquidity). It is lower than the case (light blue) of small $m$, indicating attenuation of optimal short rates in the case of a highly illiquid financial sector. In sum, the calibration results are quite reasonable. Nevertheless, we do not place much emphasis on this aspect of the paper, since our focus is to illustrate, in a tractable theoretical framework, the forces at work when a central bank faces the dual tradeoffs of inflation and fragility.

## 5 Conclusions

We develop a framework that characterises central banks' dual tradeoff between financial stability and inflation control. We analyse two models: a base model where there is no feedback, and an extended model with feedback effects between the real and financial sectors. The base model offers a potential explanation of several interesting phenomena. In particular, a large, positive real economy shock shifts the production function upward and raises marginal productivity. Since marginal productivity determines the marginal return on capital, banks increase their supply of credit. Consequently real shocks have financial effects, and our model explains the empirically observed procyclicality of credit. Similarly, a large negative financial shock reduces the credit banks supply. This, in turn, reduces firms' investment and therefore results in output contraction. Hence, our model explains the

[10] Solution of the equations was performed in Maple.

empirically observed occasional spillovers from financial crises to the real economy. These spillovers will be more prominent, the larger the correlation between real and financial shocks.

In the extended model, we account for both intermediaries' and central banks' incentives to deviate from appropriate bailout, liquidity, and interest rate targets. Our model introduces a novel tradeoff between enhancing economic output and maintaining financial stability. Central banks will experience the bite of this tradeoff, the larger the correlation between real and financial shocks. Specifically, higher interest rates deter buildup of financial imbalances, but lower rates reduce the likelihood of insolvency. We term these factors the discipline effect and stability effect, respectively. We show that the central bank's welfare loss increases with dependence between the real and financial shocks. Thus, a central bank may be able to reduce tradeoff costs by marketing correlation-indexed securities.

Our findings extend the literature on central bank independence (e.g. Rogoff (1985)). This literature demonstrates that in order to achieve price stability, monetary policy should be delegated to a conservative central bank whose sole concern is inflation. Our framework suggest a further impossibility result: in our model, even with an independent central bank, the attainment of both low inflation and financial stability is generally not possible.

# References


Allen, F., Carletti, E., Gale, D., 2009. Interbank market liquidity and central bank intervention. Journal of Monetary Economics 56 (5), 639–652.

Allen, F., Gale, D., 2007. Understanding Financial Crises. Clarendon Lectures in Finance. Oxford University Press.

Barro, R., Gordon, D., 1983. Rules, discretion, and reputation in a model of monetary policy. Journal of Monetary Economics 12, 101–121.

Bernanke, B., Gertler, M., 1989. Agency costs, net worth, and business fluctuations. American Economic Review 79 (1), 14–31.

Cao, J., Illing, G., 2012. "interest rate trap", or: Why does the central bank keep the policy rate too low for too long time? Tech. rep., Norwegian Central Bank.

Chollete, L., 2012. A model of endogenous extreme events. Tech. rep., UiS Business School.

Chollete, L., Jaffee, D., 2012. Financial implications of extreme and rare events. Tech. rep., University of Stavanger.

Cornes, R., Sandler, T., 1996. The Theory of Externalities, Public Goods and Club Goods: Second Edition. Cambridge University Press, New York.

Fisher, I., 1933. The debt-deflation theory of great depressions. Econometrica 1 (4), 337–357.

Freixas, X., Parigi, B., Rochet, J., 2000. Systemic risk, interbank relations, and liquidity provision by the central bank. Journal of Money, Credit and Banking 32 (3, part 2), 611–638.

Fudenberg, D., Tirole, J., 1991. Game Theory. MIT Press, Cambridge, MA.

Gibbons, R., 1992. Game Theory for Applied Economists. Princeton University Press, New Jersey.

Kashyap, A., Stein, J., 2012. Optimal conduct of monetary policy with interest on reserves. American Economic Journal: Macroeconomics 4, January (1), 266–282.

Keynes, J., 1936. The General Theory of Employment, Interest and Money. Harcourt, Brace, and Company, New York.

Kydland, F., Prescott, E., 1977. Rules rather than discretion: The inconsistency of optimal plans. Journal of Political Economy 85, 473–491.

Lucas, R., 1972. Expectations and the neutrality of money. Journal of Economic Theory 4, 103–124.

Mas-Colell, A., Whinston, M., Green, J., 1995. Microeconomic Theory. Oxford Press.

Phillips, A., 1958. The relation between unemployment and the rate of change of money wage rates in the united kingdom, 1861-1957. Economica XXV, 283–299.

Rogoff, K., 1985. The optimal degree of commitment to an intermediate monetary target. Quarterly Journal of Economics 100, 1169–1189.

Shin, H., 2010. Risk and Liquidity. Oxford University Press, New York.

Woodford, M., 2003. Interest Rates and Prices: Foundations of a Theory of Monetary Policy. Princeton University Press, Princeton, New Jersey.

# A Appendix

## A.1 Alternative fragility measure $f^0$, and proofs of propositions 1 and 2

This approach observes that individual banks may not consider the full impact of their individual risk-taking on the aggregate likelihood of extreme events. The mechanism can be elucidated with a simple externality model (Cornes and Sandler (1996)), as in Chollete (2012). Consider an economy populated with a large number $N$ of small banks, indexed by $i = 1, ...N$. All banks maximize the same expected yield function $Y$. The banks choose their assets $a^i$ and liabilities $l^i$, and aggregate liabilities $L$ are simply the sum of individual liabilities, $L \equiv \sum_i l^i$. Since banks are small, they cannot control other banks' behavior, and take the total of other banks' liabilities, $L^{-i} \equiv \sum_{j \neq i} l^j$, as given. The 'price' of holding one unit of assets is the foregone interest $r$, while the price of holding liabilities is the bankruptcy cost $b$ in event of a default.

**Welfare effects of competitive banking**. We first examine the competitive solution. The typical bank's problem is to choose a combination of assets $a^i$ and liabilities $l^i$ to maximize expected yield, $Y(a^i, l^i, s(L))$. The function $Y(\cdot)$ is concave and continuously differentiable, and the last term $s$ represents the likelihood of *systemic risk*. As the total amount of liabilities $L$ rises, it increases the likelihood of systemic risk. This effect occurs for a number of reasons (liquidity-mismatch, changing investor confidence, etc.), which we summarize with a continuously differentiable function $s = s(L)$. Partial derivatives are denoted with a subscript, for example, $Y_a \equiv \partial Y / \partial a$. In order to capture the fact that systemic risk increases with total liabilities, let $s_L \equiv s'(L) > 0$.[11] Since banks are averse to systemic risk, we let $Y_s \equiv \partial Y / \partial s < 0$. Now the representative bank solves the following problem:

$$\max_{\{a^i, l^i\}} Y[a^i, l^i, s(l^i + L^{-i})] \tag{22}$$

subject to

$$r \cdot a^i + b \cdot l^i \leq k^i, \tag{23}$$

where $k^i$ is the initial capital of bank $i$. In a competitive industry, excess profits will be eroded such that (23) becomes an equality. Under the given assumptions, the bank's problem has a solution, with the following necessary and sufficient conditions: $Y_a^i - \lambda \cdot r = 0$ and $Y_l^i + Y_s^i s_L(L) - \lambda \cdot b = 0$. These conditions combine to yield $\frac{Y_l^i + Y_s^i s_L(L)}{Y_a^i} = \frac{b}{r}$, which can be simplified in terms of the likelihood of systemic risk as

$$s_L^c(L^c) = \frac{Y_a^i \cdot \frac{b}{r} - Y_l^i}{Y_s^i}, \tag{24}$$

where the superscript $c$ denotes the competitive solution. The above relation characterizes sensitivity of the banking industry to increased likelihood of systemic risk, for a unit increase in total liabilities $L$, at the competitive optimum. Thus, it captures a measure of financial fragility.

[11] A similar approach is used in Shin (2010).

In order to assess efficiency of the competitive solution above, we compute a Pareto optimum. A social planner calculates the socially optimal level of investment by solving the following optimization program:

$$\max_{\{a^i,l^i\}} \sum_{i=1}^{N} Y^i[a^i, l^i, s(L)] \tag{25}$$

subject to

$$r\sum_i a^i + b\sum_i l^i \leq \sum_i k^i.$$

The solution involves the following first order conditions: $Y_a^i - \lambda \cdot r = 0$, all $i$, and $Y_l^i + \left(\sum_j Y_s^j\right) \cdot s_L(L) - \lambda \cdot b = 0$, which combine to yield $\frac{Y_l^i + \left(\sum_j Y_s^j\right) \cdot s_L(L)}{Y_a^i} = \frac{b}{r}$. Further simplification results in the Pareto optimum condition for aggregate liabilities:

$$s_L^*(L^*) = \frac{Y_a^i \cdot \frac{b}{r} - Y_l^i}{\left(\sum_j Y_s^j\right)}, \tag{26}$$

where the superscript $^*$ denotes a Pareto optimum. Evidently, (26) differs from the competitive solution in (24) by dint of the larger denominator. Thus, $s_L^*(L^*) < s_L^c(L^c)$.

**General fragility measure** $f^0$. The discussion above shows that the sensitivity $s_L(L)$ of the economy to systemic risk is suboptimally large in a competitive banking system. Since $s_L^c(L^c)$ measures sensitivity of the banking sector to systemic risk, we use it to define our general fragility measure $f^0$, as below.

**Definition 2**. The **likelihood-based fragility** measure is $f^0(L) = s_L^c(L^c)$.

Proofs of Propositions 1 and 2:

**Proposition 1**: *In a competitive banking system, financial fragility $f$ can be represented as a function of equilibrium aggregate liabilities $L = \sum_i l^i$, or as a function of aggregate bank investment $I = \sum_i I^i$.*

**Proof of Proposition 1**. By Definition 1 in Section 2 and Definition 2 above. ■

**Proposition 2**: *In a competitive banking system with fire-sale externalities, the likelihood-based fragility $f^0$ and the leverage-based fragility $f^1$ are inefficiently high.*

**Proof of Proposition 2**. Inefficiency of $f^0$ follows from comparing equations (24) and (26) above. Inefficiency of $f^1$ follows from comparing equations (2) and (3) in Section 2. ■

## A.2 Derivation of optimal policy response $r^c$ of section 3.1

Recall from equation (7) that the the central bank's decision problem is

$$\min_r L = \left[\alpha y^* + \beta\left(r - \hat{r} + \varepsilon_y\right) - y^*\right]^2 + a\left(r - r^*\right)^2 + b\left[\left(\theta - r - \delta + m\delta + \varepsilon_f\right)\frac{K}{2pm}\right]^2.$$

The first-order condition is

$$\frac{\partial L}{\partial r} = 0 = 2\beta\left[\alpha y^* + \beta\left(r - \hat{r} + \varepsilon_y\right) - y^*\right] + 2a\left(r - r^*\right) + 2b\left[\left(\theta - r - \delta + m\delta + \varepsilon_f\right)\frac{K}{2pm}\right]\left(-\frac{K}{2pm}\right),$$

which can be solved to obtain the central bank's best response, denoted $r^c$:

$$r^c = -\frac{1}{a + \beta^2 + b\left(\frac{K}{2pm}\right)^2}\left[\alpha\beta y^* - \beta^2\hat{r} + \beta^2\varepsilon_y - \beta y^* - ar^* - b\left(\frac{K}{2pm}\right)^2(\theta + \varepsilon_f) + (1-m)\,\delta b\left(\frac{K}{2pm}\right)^2\right]. \tag{27}$$

To compute the above expression, we require the interest rate expectations $\hat{r}$, to which we now turn. If expectations are on average correct in an equilibrium, then $\hat{r} = E\left[r^c\right]$, whence we can derive the following relation:

$$\hat{r} = -\frac{1}{a + \beta^2 + b\left(\frac{K}{2pm}\right)^2}\left[\alpha\beta y^* - \beta^2\hat{r} - \beta y^* - ar^* - b\left(\frac{K}{2pm}\right)^2\theta + (1-m)\,\delta b\left(\frac{K}{2pm}\right)^2\right].$$

This expression can be simplified to

$$\hat{r} = -\frac{1}{a + b\left(\frac{K}{2pm}\right)^2}\left[\alpha\beta y^* - \beta y^* - ar^* - b\left(\frac{K}{2pm}\right)^2\theta + (1-m)\,\delta b\left(\frac{K}{2pm}\right)^2\right],$$

or, equivalently,

$$\hat{r} = \frac{1}{a + b\left(\frac{K}{2pm}\right)^2}\left[(1-\alpha)\,\beta y^* + ar^* + (\theta - \delta + m\delta)\left(\frac{K}{2pm}\right)^2 b\right]. \tag{28}$$

By substituting the $\hat{r}$ term from (28) into (27) we obtain the equilibrium policy rate, namely,

$$r^c = -\frac{1}{a+\beta^2+b\left(\frac{K}{2pm}\right)^2}\left\{\frac{\beta^2}{a+b\left(\frac{K}{2pm}\right)^2}\left[\alpha\beta y^* - \beta y^* - ar^* - b\left(\frac{K}{2pm}\right)^2\theta + (1-m)\,\delta b\left(\frac{K}{2pm}\right)^2\right] + \beta^2\varepsilon_y\right\}$$

$$-\frac{1}{a+\beta^2+b\left(\frac{K}{2pm}\right)^2}\left[\alpha\beta y^* - \beta y^* - ar^* - b\left(\frac{K}{2pm}\right)^2(\theta+\varepsilon_f) + (1-m)\,\delta b\left(\frac{K}{2pm}\right)^2\right].$$

Now we just simplify the above expression. First, note that upon collecting terms, it becomes

$$r^c = -\frac{1}{a+\beta^2+b\left(\frac{K}{2pm}\right)^2}\left\{\frac{a+\beta^2+b\left(\frac{K}{2pm}\right)^2}{a+b\left(\frac{K}{2pm}\right)^2}\left[\alpha\beta y^* - \beta y^* - ar^* - b\left(\frac{K}{2pm}\right)^2\theta + (1-m)\,\delta b\left(\frac{K}{2pm}\right)^2\right]\right\}$$

$$-\frac{1}{a+\beta^2+b\left(\frac{K}{2pm}\right)^2}\left\{\beta^2\varepsilon_y - b\left(\frac{K}{2pm}\right)^2\varepsilon_f\right\}.$$

Now define $A = a+\beta^2+b\left(\frac{K}{2pm}\right)^2$, $B = a+b\left(\frac{K}{2pm}\right)^2$, and $C = b\left(\frac{K}{2pm}\right)^2$. Then the above expression implies

$$r^c = \frac{1}{B}\left[-\alpha\beta y^* + \beta y^* + ar^* + b\left(\frac{K}{2pm}\right)^2\theta - (1-m)\,\delta b\left(\frac{K}{2pm}\right)^2\right] - \frac{1}{A}\left[\beta^2\varepsilon_y - b\left(\frac{K}{2pm}\right)^2\varepsilon_f\right]$$

or

$$r^c = \frac{1}{B}\left[(1-\alpha)\,\beta y^* + ar^* + (\theta - \delta + m\delta)\,C\right] - \frac{1}{A}\left[\beta^2\varepsilon_y - C\varepsilon_f\right]. \tag{29}$$

Similarly, we can simplify the expected interest rate $\hat{r}$ from equation (28) to be

$$\hat{r} = \frac{1}{B}\left[(1-\alpha)\,\beta y^* + ar^* + (\theta - \delta + m\delta)\,C\right] - \frac{1}{A}\left(\beta^2\varepsilon_y - C\varepsilon_f\right) \tag{30}$$

Thus $r^c$ and $\hat{r}$ have the following relation, $r^c = \hat{r} - \frac{1}{A}(\beta^2\varepsilon_y - C\varepsilon_f)$.

We also need to obtain the term for output $y$. From the Phillips curve relation (5), actual output is

$$y = \alpha y^* + \beta\left(r^c - \hat{r} + \varepsilon_y\right),$$

which upon substitution yields

$$y = \alpha y^* + \frac{a + b\left(\frac{K}{2pm}\right)^2}{a + \beta^2 + b\left(\frac{K}{2pm}\right)^2}\beta\varepsilon_y + \frac{b\left(\frac{K}{2pm}\right)^2}{a + \beta^2 + b\left(\frac{K}{2pm}\right)^2}\beta\varepsilon_f.$$

and finally,

$$y = \alpha y^* + \frac{B}{A}\beta\varepsilon_y + \frac{C}{A}\beta\varepsilon_f = \alpha y^* + \frac{\beta}{A}\left(B\varepsilon_y + C\varepsilon_f\right). \tag{31}$$

## A.3 Derivation of expected welfare loss $E[L^n]$ in section 3.1.

To compute this moment, we first calculate the welfare loss expression $L^n$, by substituting (29) and (31) into the objective (7):

$$\begin{aligned}
L &= \left[\alpha y^* + \frac{B}{A}\beta\varepsilon_y + \frac{C}{A}\beta\varepsilon_f - y^*\right]^2 \\
&\quad + a\left[\frac{1}{B}\left[(1-\alpha)\beta y^* + ar^* + (\theta - \delta + m\delta)C\right] - \frac{1}{A}(\beta^2\varepsilon_y - C\varepsilon_f) - r^*\right]^2 \\
&\quad + b\left[\left(\theta - \left(\frac{1}{B}\left[(1-\alpha)\beta y^* + ar^* + (\theta - \delta + m\delta)C\right] - \frac{1}{A}(\beta^2\varepsilon_y - C\varepsilon_f)\right) - \delta + m\delta + \varepsilon_f\right)\frac{K}{2pm}\right]^2 \\
&= \left[(\alpha-1)y^* + \frac{B}{A}\beta\varepsilon_y + \frac{C}{A}\beta\varepsilon_f\right]^2 + a\left[\frac{(1-\alpha)\beta y^* + (a-B)r^* + (\theta - \delta + m\delta)C}{B} - \frac{\beta^2\varepsilon_y - C\varepsilon_f}{A}\right]^2 \\
&\quad + b\left[\left(\frac{B(\theta - \delta + m\delta) - \left[(1-\alpha)\beta y^* + ar^* + (\theta - \delta + m\delta)C\right]}{B} + \frac{\beta^2\varepsilon_y + (A-C)\varepsilon_f}{A}\right)\frac{K}{2pm}\right]^2.
\end{aligned}$$

To simplify, note that $(a-B)r^* = -Cr^*$, in the third line; and $(B-C)(\theta-\delta+m\delta) = a(\theta-\delta+m\delta)$, and $(A-C)\varepsilon_f = (a+\beta^2)\varepsilon_f$, for the fourth line. Thus we can rewrite the above expression as

$$\begin{aligned}
L &= \left[(\alpha-1)y^* + \frac{B}{A}\beta\varepsilon_y + \frac{C}{A}\beta\varepsilon_f\right]^2 + a\left[\frac{(1-\alpha)\beta y^* + (\theta - \delta + m\delta - r^*)C}{B} - \frac{\beta^2\varepsilon_y - C\varepsilon_f}{A}\right]^2 \\
&\quad + b\left(\frac{K}{2pm}\right)^2\left[\left(\frac{-(1-\alpha)\beta y^* - ar^* + (\theta - \delta + m\delta)a}{B} + \frac{\beta^2\varepsilon_y + (a+\beta^2)\varepsilon_f}{A}\right)\right]^2 \\
&= \left[(\alpha-1)y^* + \frac{B}{A}\beta\varepsilon_y + \frac{C}{A}\beta\varepsilon_f\right]^2 + a\left[\frac{(1-\alpha)\beta y^* + (\theta - \delta + m\delta - r^*)C}{B} - \frac{\beta^2\varepsilon_y - C\varepsilon_f}{A}\right]^2 \\
&\quad + C\left[\left(-\frac{(1-\alpha)\beta y^* + ar^* - (\theta - \delta + m\delta)a}{B} + \frac{\beta^2\varepsilon_y + (a+\beta^2)\varepsilon_f}{A}\right)\right]^2.
\end{aligned} \tag{32}$$

Now that we have the expression for welfare loss $L$, we can compute its moments, which rely on joint stochastic properties of the real and financial shocks $\varepsilon_y$ and $\varepsilon_f$. Recall that the distribution of the shocks is the following. $\varepsilon \equiv [\varepsilon_y, \varepsilon_f]'$, then $\varepsilon \sim [0, V]$, where the covariance matrix $V$ is of the form $V = \begin{bmatrix} \sigma_y^2 & \rho_{y,f} \\ \rho_{y,f} & \sigma_f^2 \end{bmatrix}$.

Then we can take expectations of (32). In the following, we first express the constant terms, then the stochastic ones, i.e. those that depend on the moments of $\varepsilon_y$ and $\varepsilon_f$. Specifically,

$$\begin{aligned}
E[L] &= (\alpha-1)^2 y^{*2} + a\left[\frac{(1-\alpha)\beta y^* + (\theta-\delta+m\delta-r^*)C}{B}\right]^2 \\
&\quad + C\left[\frac{(1-\alpha)\beta y^* + ar^* - (\theta-\delta+m\delta)a}{B}\right]^2 \\
&\quad + \left(\left[\frac{B}{A}\right]^2\beta^2 + a\left[\frac{\beta^2}{A}\right]^2 + C\left[\frac{\beta^2}{A}\right]^2\right)\sigma_y^2 + \left(\left[\frac{C}{A}\right]^2\beta^2 + a\left[\frac{C}{A}\right]^2 + C\left[\frac{(a+\beta^2)}{A}\right]^2\right)\sigma_f^2 \\
&\quad + 2\left(\frac{B}{A}\frac{C}{A}\beta^2 - a\frac{\beta^2}{A}\frac{C}{A} + C\frac{\beta^2}{A}\frac{(a+\beta^2)}{A}\right)\rho_{y,f} \\
&= (\alpha-1)^2 y^{*2} + a\left[\frac{(1-\alpha)\beta y^* + (\theta-\delta+m\delta-r^*)C}{B}\right]^2 \\
&\quad + C\left[\frac{(1-\alpha)\beta y^* + ar^* - (\theta-\delta+m\delta)a}{B}\right]^2 \\
&\quad + \frac{B^2 + a\beta^2 + C\beta^2}{A^2}\beta^2\sigma_y^2 + \frac{C^2(\beta^2+a) + C(a+\beta^2)^2}{A^2}\sigma_f^2 + 2C\beta^2\frac{B-a+a+\beta^2}{A^2}\rho_{y,f},
\end{aligned}$$

or

$$\begin{aligned}
E[L^n] &= (\alpha-1)^2 y^{*2} + a\left[\frac{(1-\alpha)\beta y^* + (\theta-\delta+m\delta-r^*)C}{B}\right]^2 \\
&\quad + C\left[\frac{(1-\alpha)\beta y^* + ar^* - (\theta-\delta+m\delta)a}{B}\right]^2 \\
&\quad + \frac{B^2 + \beta^2(a+C)}{A^2}\beta^2\sigma_y^2 + \frac{(a+\beta^2)^2(C^2+C)}{A^2}\sigma_f^2 + 2C\beta^2\frac{B+\beta^2}{A^2}\rho_{y,f} \qquad (33)
\end{aligned}$$

In contrast, the cooperative solution features two conditions: $\hat{r} = r^c = r^*$; and $r^* = \theta - (1-m)\delta$. Substituting these conditions into the objective function (7) and taking expectations yields the expected loss under perfect anchoring, $E[L^c]$:

$$\begin{aligned}
E[L^c] &= E\left[[(\alpha-1)y^* + \beta\varepsilon_y]^2 + \left(\frac{K}{2pm}\varepsilon_f\right)^2\right] \\
&= (\alpha-1)^2 y^{*2} + \beta^2\sigma_y^2 + \frac{C}{b}\sigma_f^2, \qquad (34)
\end{aligned}$$

where the second line uses the fact that $C = b\left(\frac{K}{2pm}\right)^2$. The difference in expected loss $\Delta = E[L] - E[L^c]$ is calculated from (33) and (34) to be

$$\begin{aligned}\Delta &= a\left[\frac{(1-\alpha)\beta y^* + (\theta - \delta + m\delta - r^*)C}{B}\right]^2 + C\left[\frac{(1-\alpha)\beta y^* + ar^* - (\theta - \delta + m\delta)a}{B}\right]^2 \\ &+ \frac{B^2 + \beta^2(a+C) - A^2}{A^2}\beta^2\sigma_y^2 + \frac{b(a+\beta^2)^2(C^2+C) - A^2C}{bA^2}\sigma_f^2 + 2C\beta^2\frac{B+\beta^2}{A^2}\rho_{y,f}\end{aligned} \quad (35)$$

**Proposition 3:** *In a dual tradeoff model, the central bank's expected loss increases monotonically with the dependence between financial and real shocks.*

**Proof:** We have to show that $\Delta = E[L] - E[L^c]$ increases with dependence. Dependence is measured by $\rho_{y,f}$, and the expected loss differential is presented in equation (35). From examination of the last term, it is evident that the coefficient on $\rho_{y,f}$ is unambiguously positive. Therefore $E[L] - E[L^c]$ increases with $\rho_{y,f}$, as was to be shown. ■

**Proposition 4: Inefficiency of bank equilibrium investment.** *Part A: In market equilibrium, the bank's investment is larger than the planner's solution, i.e., $I^E > I^p$. Moreover, $\sqrt{3}I^p < I^E < 2I^p$. Part B: In market equilibrium, liquidity risk provides some market discipline, but investment is still excessive.*

*Proof of Part A:* Subtracting (13) from (12), one can get

$$\gamma\omega\left(3(I^p)^2 - (I^E)^2\right) + \left(2I^p - I^E\right)\left(\omega\hat{r}K - \gamma r_1 - \gamma\delta + \gamma m\delta\right) \equiv 0.$$

The banks borrow short-term debt only if long-term debt is too costly, or $\omega < \omega I^p \leq r_1 + \delta$. Therefore

$$\omega\hat{r}K - \gamma\left(r_1 + \delta\right) + \gamma m\delta \geq \omega\left(\hat{r}K - \gamma\right) + \gamma m\delta > 0$$

as long as $\hat{r}K - \gamma > 0$. This latter condition is fairly weak: since $\hat{r} > 1 > \gamma$, the inequality strictly holds if $I^p > K > 1$.

Suppose that $I^E < I^E$, then $\gamma\omega\left(3(I^p)^2 - (I^E)^2\right) + \left(2I^p - I^E\right)\left(\omega\hat{r}K - \gamma r_1 - \gamma\delta + \gamma m\delta\right) > 0$. A contradiction.

Suppose that $I^p \leq I^E \leq \sqrt{3}I^p$, then $\gamma\omega\left(3(I^p)^2 - (I^E)^2\right) + \left(2I^p - I^i\right)\left(\omega\hat{r}K - \gamma r_1 - \gamma\delta + \gamma m\delta\right) > 0$. A contradiction.

Suppose that $I^E \geq 2I^p$, then $\gamma\omega\left(3(I^p)^2 - (I^E)^2\right) + \left(2I^p - I^i\right)\left(\omega\hat{r}K - \gamma r_1 - \gamma\delta + \gamma m\delta\right) < 0$. A contradiction.

Therefore, $\sqrt{3}I^p < I^E < 2I^p$.

*Part B*. This part of the proof involves evaluating comparative statics of the liquidity effect. Using the implicit function theorem, we can take the derivative of (17) with respect to $m$ to get

$$\frac{\partial I^E}{\partial m} = -\frac{K\delta + \gamma I\delta}{\omega K - \gamma r_1 - \gamma\delta + \gamma m\delta + 2I^E\gamma\omega}.$$

The denominator can be rewritten as

$$\omega K + \gamma(\omega I^E - r_1) + \gamma(\omega I^E - \delta) + \gamma m\delta.$$

Since $\omega I^E$ is the return on investment, $\omega I^E > r_1 > \delta$ by assumption. Therefore the denominator is positive, and $\frac{\partial I^E}{\partial m} < 0$. ■

## A.4 Derivation of optimal short rate $r^E$ in model of section 4.2

Recall that the central bank's program is

$$\min_{r^E} L_C = (y - y^*)^2 + a\left(r^E - r^*\right)^2 + b\gamma\sum_i I^i G\left(\varepsilon_y^{i*}\right)(i),$$

For simplicity, we have assumed that bankruptcy formally occurs in the end of the period, $t_2$. Thus, the share of bankrupted banks can be written as

$$\frac{mI^E\hat{r}r^E + (1-m)I^E(r_1+\delta) - \omega\left[I^E + \beta\left(r^E - \hat{r} + \varepsilon_y\right)\right]I^E}{K}.$$

The central bank's decision problem can now be rewritten as

$$\begin{aligned}\min_{r^E} L_C \quad = \quad & \left[\left(I^E - I^*\right) + \beta\left(r^E - \hat{r} + \varepsilon_y\right)\right]^2 + a\left(r^E - r^*\right)^2 \\ & + b\gamma I^E\frac{mI^E\hat{r}r^E + (1-m)I^E(r_1+\delta) - \omega\left[I^E + \beta\left(r^E - \hat{r} + \varepsilon_y\right)\right]I^E}{K}.\end{aligned}$$

The first order condition is linear in $r^E$ so that we can solve $\hat{r}$ by taking expectations. The first order condition yields

$$\frac{\partial L_C}{\partial r^E} = 2\beta\left[\left(I^E - I^*\right) + \beta\left(r^E - \hat{r} + \varepsilon_y\right)\right] + 2a\left(r^E - r^*\right) + \frac{b\gamma I^E}{K}\left(mI^E\hat{r} - \omega I^E\beta\right) = 0.$$

Since in a rational expectations equilibrium $E\left[r^E\right] = \hat{r}$, we take expectations on the first order condition, in order to obtain

$$2\beta\left(I^E - I^*\right) + 2a\left(\hat{r} - r^*\right) + \frac{b\gamma I^E}{K}\left(mI^E\hat{r} - \omega I^E\beta\right) = 0,$$

which implies that interest rate expectations satisfy

$$\hat{r} = \frac{2aKr^* - 2\beta K\left(I^E - I^*\right) + b\gamma(I^E)^2\omega\beta}{2aK + b\gamma(I^E)^2 m}.$$

We now insert $\hat{r}$ into the first order condition, and solve for the central bank's optimal short rate to obtain $r^E = \frac{2aKr^* - 2\beta K\left(I^E - I^*\right) + b\gamma(I^E)^2\omega\beta}{2aK - b\gamma(I^E)^2 m} - \frac{\beta^2}{a+\beta^2}\varepsilon_y$,

or, equivalently,

$$r^E = \frac{2aK}{2aK + b\gamma(I^E)^2 m}r^* - \frac{2\beta K\left(I^E - I^*\right)}{2aK + b\gamma(I^E)^2 m} + \frac{b\gamma\left(I^E\right)^2\omega\beta}{2aK + b\gamma(I^E)^2 m} - \frac{\beta^2}{a+\beta^2}\varepsilon_y,$$

which is presented as equation (14) of the text.

**Proof of the impossibility result, Proposition 5**

We have to show that when the central bank's rates are equal, $r^E = r^O$, then the target rate is a linear function of the investment gap: $r^* = \frac{\beta}{a}(I^E - I^*) + \frac{\omega\beta}{m}$.

The central bank can use the same rate to control both inflation and targeting only if $r^E - r^O$ is identically zero. We solve this condition by taking the difference $r^E - r^0$ :

$$\begin{aligned}
r^E - r^0 &= \frac{2aK}{2aK + b\gamma(I^E)^2 m}r^* - \frac{2\beta K(I^E - I^*)}{2aK + b\gamma(I^E)^2 m} + \frac{b\gamma(I^E)^2\omega\beta}{2aK + b\gamma(I^E)^2 m} - \frac{\beta^2}{a+\beta^2}\varepsilon_y \\
&\quad - \left[r^* - \frac{\beta}{a}(I^E - I^*) - \frac{\beta^2}{a+\beta^2}\varepsilon_y\right] \\
&= \left[\frac{2aK}{2aK + b\gamma(I^E)^2 m} - 1\right]r^* + \beta[I^E - I^*]\left[\frac{1}{a} - \frac{2K}{2aK + b\gamma(I^E)^2 m}\right] + \frac{b\gamma(I^E)^2\omega\beta}{2aK + b\gamma(I^E)^2 m} \\
&= \left[\frac{2aK - 2aK - b\gamma(I^E)^2 m}{2aK + b\gamma(I^E)^2 m}\right]r^* + \beta[I^E - I^*]\left[\frac{2aK + b\gamma(I^E)^2 m - 2aK}{a[2aK + b\gamma(I^E)^2 m]}\right] + \frac{b\gamma(I^E)^2\omega\beta}{2aK + b\gamma(I^E)^2 m} \\
&= \left[\frac{-b\gamma(I^E)^2 m}{2aK + b\gamma(I^E)^2 m}\right]r^* + \beta[I^E - I^*]\left[\frac{b\gamma(I^E)^2 m}{a[2aK + b\gamma(I^E)^2 m]}\right] + \frac{b\gamma(I^E)^2\omega\beta}{2aK + b\gamma(I^E)^2 m} \\
&= \left[\frac{-b\gamma(I^E)^2 m}{2aK + b\gamma(I^E)^2 m}\right]r^* + \left[\frac{b\beta\gamma(I^E)^2 m}{a[2aK + b\gamma(I^E)^2 m]}\right](I^E - I^*) + \frac{b\gamma(I^E)^2\omega\beta}{2aK + b\gamma(I^E)^2 m},
\end{aligned} \tag{36}$$

or

$$r^E - r^0 = \left[\frac{b\gamma(I^E)^2}{2aK + b\gamma(I^E)^2 m}\right]\left[m\left(\frac{\beta}{a}(I^E - I^*) - r^*\right) + \omega\beta\right]. \tag{37}$$

The central bank can use the same rate to control both inflation and targeting only if the expression in (37) is identically zero, i.e. at least one of the terms on the RHS is zero. Let us examine the first term. This term is non-zero because the central bank cares about fragility ($b > 0$), fire sales occur ($\gamma > 0$), investment is positive ($I^E > 0$), and there is always some short term debt in the economy ($m > 0$). Therefore it suffices to examine only the second term. This second term is zero if the target rate is a specific linear function the investment gap: $r^* = \frac{\beta}{a}(I^E - I^*) + \frac{\omega\beta}{m}$, as was to be shown. ■

# B Calibrations for optimal investment, short-rates, and threshold discount rate

We calibrate the key variables from Section 4, in order to illustrate them graphically.

## B.1 Calibration of investment and short-rates

In this part we calibrate variables from Section 4.2 and 4.3. Aggregate investment is represented in equations (12) and (13), while short rates under fragility and traditional central banking are in equations (15) and (14). The main parameters are described below.

Table 2: Description of Variables in Investment and Short Rates

| **Parameter** | **Interpretation** |
|---|---|
| 1. $a$ | Cost to central bank of missing inflation or interest target |
| 2. $b$ | Cost to central bank of missing investment target |
| 3. $\beta$ | Phillips curve-based cost to central bank of missing interest target |
| 4. $r^*$ | Target interest rate |
| 5. $I^E - I^*$ | Gap between optimal and target investment |
| 6. $I^*$ | Optimal Investment |
| 7. $K$ | Total equity in banking system |
| 8. $\gamma$ | Multiplier for systemic costs relative to $I$, which measures fire-sale costs |
| 9. $m$ | Fraction of short-term debt in banking system, which measures system illiquidity |
| 10. $\varepsilon_y$ | Output shock |

We approach determination of the parameter values through a combination of natural constraints, theoretical considerations, and banking practice. Notice that:

- Typical inflation targeting central banks put much higher weight on inflation loss than output loss, i.e., $a >> 1$;

- The ratio $\frac{I^E}{k}$ reflects the bank's leverage ratio. The typical leverage in the banking industry is about $12.5$, rarely below $5$; therefore, $\frac{I^E}{K} = 5$ (the lowest banking leverage) may be a good guess;
- The value $\gamma \Sigma_i I^i$ is the fire sale loss of one unit liquidated asset, therefore $\gamma I^E < 1$ — even in the worse case when all the bank assets are liquidated, the banks can still recover some value ($1 - \gamma I^E$ per unit);
- The reliance on short-term debt varies quite a lot across banks, and depends mainly on how much investment banking business one bank gets involved. However, post-crisis banks tend to focus more on the mainstream banking, and $m$ is mostly between $20\%$ and $40\%$ in Europe and US;
- For regularity, we require $2aK - b\gamma \left(I^E\right)^2 m > 0$.

## B.2 Calibrating the threshold discount rate

We now calibrate the threshold discount rate $d$ for the infinitely repeated game from Section 4.4. The discount rate is given by expression (21), reproduced here:

$$d \geq \frac{\tilde{L}_C^p - L_C^p}{\tilde{L}_C^p - L_C^E} \equiv \bar{d}. \tag{38}$$

The expressions for $\tilde{L}_C^p$, $L_C^p$ and $L_C^E$ depend on similar parameters to those for the optimal policy rates in section 1 above. In particular, equation (21) gives the central bank's loss function as

$$\begin{aligned} L_C(r^c, \hat{r}, I^i) &= [(I^i - I^*) + \beta(r^c - \hat{r} + \varepsilon_y)]^2 + a(r^c - r^*)^2 \\ &\quad + b\gamma I^i \frac{mI^i \hat{r} r^c + (1-m) I^i (r_1 + \delta) - \omega[I^i + \beta(r^c - \hat{r} + \varepsilon_y)] I^i}{K}, \end{aligned}$$

where $r^c \in \{r^E, r^p\}$ and $I^i \in \{I^E, I^p\}$.

Preliminary results for computing threshold.

Before we compute the threshold $\bar{d}$, recall that $r^E$ and $r^p$ are calculated from equation (23) in the text, as

$$r^p = \frac{2aK}{2aK + b\gamma (I^p)^2 m} r^* - \frac{2\beta K (I^p - I^*)}{2aK + b\gamma (I^p)^2 m} + \frac{b\gamma (I^p)^2 \omega \beta}{2aK + b\gamma (I^p)^2 m} - \frac{\beta^2}{a + \beta^2} \varepsilon_y$$

and

$$r^E = \frac{2aK}{2aK + b\gamma (I^E)^2 m} r^* - \frac{2\beta K (I^E - I^*)}{2aK + b\gamma (I^E)^2 m} + \frac{b\gamma (I^E)^2 \omega \beta}{2aK + b\gamma (I^E)^2 m} - \frac{\beta^2}{a + \beta^2} \varepsilon_y.$$

Further, $I^E$ and $I^p$ satisfy the following conditions from equations (17) and (18) in the text:

$$\omega I^E K - r_1 K - K\delta + Km\delta - \gamma I^E r_1 - \gamma I^E \delta + \gamma I^E m\delta + \gamma (I^E)^2 \omega = 0,$$

and

$$2\omega I^p K - r_1 K - K\delta + Km\delta - 2\gamma I^p r_1 - 2\gamma I^p \delta + 2\gamma I^p m\delta - 3\gamma (I^p)^2 \omega = 0,$$

respectively.

Computing the threshold $\bar{d}$.

Now we can compute the three different elements of the threshold discount rate in (38)–namely $L_C^p$, $L_C^E$, and $\tilde{L}_C^p$. The first two are straightforward since they involve the bank and central bank both playing cooperate or deviate:

$$\begin{aligned} L_C^p(r^p, \hat{r}, I^p) &= [(I^p - I^*) + \beta(r^p - \hat{r} + \varepsilon_y)]^2 + a(r^p - r^*)^2 \\ &\quad + b\gamma I^p \frac{mI^p \hat{r} r^p + (1-m)I^p(r_1 + \delta) - \omega[I^p + \beta(r^p - \hat{r} + \varepsilon_y)]I^p}{K}; \quad (39) \end{aligned}$$

and

$$\begin{aligned} L_C^E(r^E, \hat{r}, I^E) &= [(I^E - I^*) + \beta(r^E - \hat{r} + \varepsilon_y)]^2 + a(r^E - r^*)^2 \\ &\quad + b\gamma I^E \frac{mI^E \hat{r} r^E + (1-m)I^E(r_1 + \delta) - \omega[I^E + \beta(r^E - \hat{r} + \varepsilon_y)]I^E}{K} \quad (40) \end{aligned}$$

The third payoff, $\tilde{L}_C^p$, is different because it reflects the central bank's payoff (by choosing $r = r^p$) if the bank deviates ($I = I^E$) for one period:

$$\begin{aligned} \tilde{L}_C^p(r^p, \hat{r}, I^E) &= [(I^E - I^*) + \beta(r^p - \hat{r} + \varepsilon_y)]^2 + a(r^p - r^*)^2 \\ &\quad + b\gamma I^E \frac{mI^E \hat{r} r^p + (1-m)I^E(r_1 + \delta) - \omega[I^E + \beta(r^p - \hat{r} + \varepsilon_y)]I^E}{K}. \quad (41) \end{aligned}$$

Thus the threshold discount factor $\bar{d}$ in (38) can be computed from equations (39) to (41), and the short rate and investment expressions from Section 2.1.

Figure 2: Threshold Discount Rate for Infinite-Horizon Game

The figure shows the threshold discount rate $d$ from equation (21). The calibration is as in Appendix B. The parameters $a$ and $b$ denote the central bank's weight on inflation and fragility, respectively.

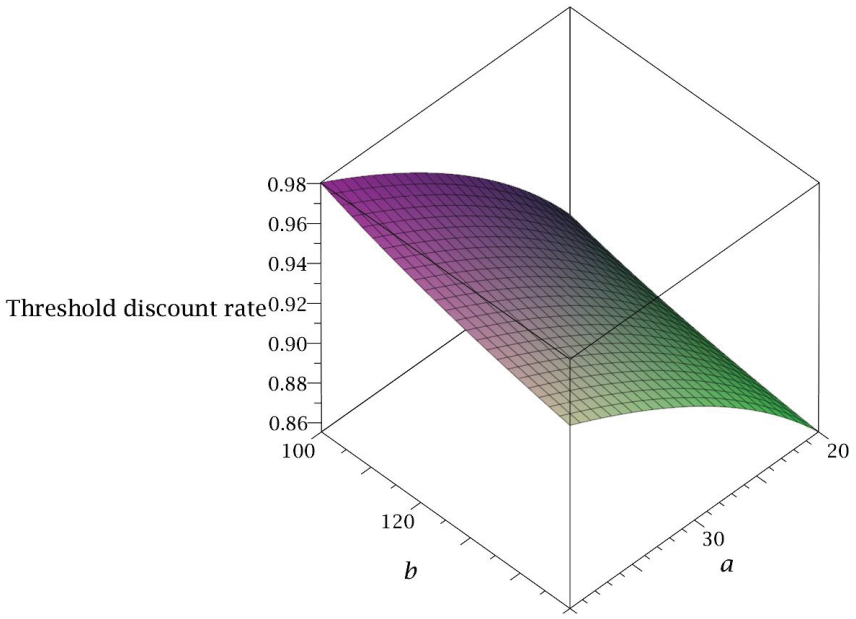

Figure 3: Aggregate Investment with Fragility-Conscious Central Bank

The figure shows the average investment surface from equations (12) and (13). The calibration is as in Appendix B. The parameters $a$ and $b$ denote the central bank's weight on inflation and fragility, respectively. The blue surfaces correspond to equation (13), and show the optimal investment when the central bank targets both inflation and fragility– the light blue corresponds to $m = 0.4$ while the dark blue corresponds to $m = 0.6$. The yellow surface corresponds to equation (12), and shows the optimal investment when the central bank only targets inflation.

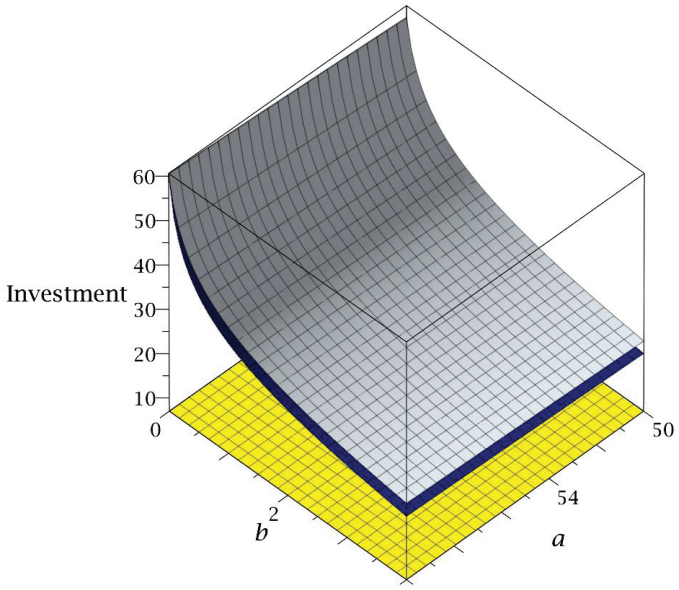

Figure 4: Optimal Policy Rates under Traditional Targeting and Fragility Concerns

The figure shows the optimal policy rates $r^0$ and $r^E$ from equations (15) and (14). The calibration is as in Appendix B. The parameters $a$ and $b$ denote the central bank's weight on inflation and fragility, respectively. The blue surfaces correspond to equation (14), and depict the optimal policy rate when the central bank targets both inflation and fragility– the light blue corresponds to $m = 0.4$, while the dark blue corresponds to $m = 0.6$. The yellow surface corresponds to equation (15), and shows the optimal short rate when the central bank only targets inflation.