BANK FOR INTERNATIONAL SETTLEMENTS

BIS Working Papers

No 817

Monetary policy hysteresis and the financial cycle

by Phurichai Rungcharoenkitkul, Claudio Borio and Piti Disyatat

Monetary and Economic Department

October 2019 (revised 17 May 2021)

JEL classification: E52, E58, E43.

Keywords: Monetary policy, financial cycle, money neutrality, hysteresis, natural rate of interest, intertemporal tradeoff.

This publication is available on the BIS website (www.bis.org).

ISSN 1020-0959 (print)
ISSN 1682-7678 (online)

Monetary Policy Hysteresis and the Financial Cycle*

Phurichai Rungcharoenkitkul[†] Claudio Borio[†] Piti Disyatat[†‡]

May 17, 2021

Check latest version here

Abstract

A long tradition of macroeconomic analysis accords monetary policy only a transient role in driving real outcomes. At the same time, a large body of evidence highlights the long-lasting impact of boom-bust cycles. We present a model where monetary policy, through its impact on and reaction to the financial cycle, influences long-term economic trajectories. The core setup is an overlapping generations model featuring bank financing – the creation of bank loans and inside money – which is critical for production and consumption. Monetary policy attains the first-best allocation by sustaining an efficient flow of financing. We then introduce coordination-failure frictions among lenders, which give rise to an endogenous boom-bust cycle in bank financing and an intertemporal policy tradeoff. A forward-looking policymaker optimally leans against excessive risk-taking during the boom, trading off short-term activity with longer-term stability. An inordinate focus on short-term outcomes can lead to 'monetary policy hysteresis', where low interest rates increase the vulnerability to financial busts over successive cycles. As a result, low rates can beget lower rates.

JEL Classification: E52, E58, E43
Keywords: Monetary policy, financial cycle, overlapping generations model, money neutrality, hysteresis, natural rate of interest, intertemporal tradeoff

*We would like to thank Iñaki Aldasoro, Frederic Boissay, Fiorella De Fiore, Benjamin Friedman, Martin Kliem, Vivien Lewis, Giovanni Lombardo, Cristina Manea, Emanuel Mönch, Cyril Monnet, Benoit Mojon, Stephane Moyen, Daniel Rees, Markus Roth, Kevin Sheedy, Hyun Song Shin, Nao Sudo, Kozo Ueda, Xuan Wang, Fabian Winkler, James Yetman, seminar participants at the BIS, the Bank of Japan, EEA Annual Congress, MMF Annual Conference, and the Swiss National Bank for their comments. All remaining errors are ours. The views expressed are those of the authors and do not necessarily represent those of the Bank for International Settlements or the Bank of Thailand.

[†]Bank for International Settlements

[‡]Bank of Thailand

On my view, there is no unique long-period position of equilibrium equally valid regardless of the character of the policy of the monetary authority. On the contrary there are a number of such positions corresponding to different policies.

J. M. Keynes

1 Introduction

What is the role of monetary policy in driving economic fluctuations? A long tradition of macroeconomic analysis accords monetary policy, and financial factors more broadly, only a transient effect on real outcomes. Under 'monetary neutrality', the equilibrium paths of all real variables are determined independently of the monetary and financial regime. This is reflected prominently in presumptions that real interest rates are tied to some 'natural rate of interest' whose variations are dictated purely by real saving-investment factors. At the same time, major economic dislocations, such as those that follow financial crises, are generally seen as being driven by exogenous shocks which are amplified by various propagation mechanisms. This is evident in much of the literature on macro-financial linkages spurred by the great financial crisis (GFC) as reviewed, for example, by Brunnermeier et al. (2013) and Claessens and Kose (2017).

We argue that this approach to modelling the role of monetary and financial factors may be overly restrictive. The influence of monetary policy, and particularly of the policy regime which governs the systematic component of monetary policy, may cast long shadows on real economic trajectories. A growing literature documents that monetary policy has a material influence on the financial cycle, not least through its impact on risk-taking, and that the financial cycle, in turn, has a long-lasting impact on output.[1] There are a number of potential mechanisms that could be at work. Some may be the familiar hysteresis effects operating through the impact of slumps on labour and capital markets (e.g. Reifschneider et al. (2015)). But others may include the much less recognised impact of financial booms and subsequent busts on resource misallocations (Cecchetti and Kharroubi (2015), Borio et al. (2016)). The potential for monetary policy to have long-lasting effects on the path of real variables through its impact on the financial cycle represents a new form of 'monetary hysteresis' that deserves greater attention.

At the same time, the focus on shocks and their propagation neglects the crucial role of underlying vulnerabilities that accumulate over time and influence the economy's proclivity, as well as response, to financial busts. Under this alternative view, major economic fluctuations arise from forces that are internal to the economy rather than exogenous shocks, resulting in recurrent periods of boom and bust (Beaudry et al. (2020)). This follows in the spirit of the long tradition of work on the financial cycle that highlight the critical interaction between financial fragility and the ebb and flow of investor sentiment (Kindleberger (2000), Minsky (1982, 1986)). To the extent that the build-up of vulnerabilities is endogenous to policy, omitting the channel understates the importance of policy frameworks for long-run financial and macroeconomic

[1] A number of studies document the fact that financial busts tend to generate long-lasting, if not permanent, output costs (e.g. Basel Committee on Banking Supervision (BCBS) (2016), Cerra and Saxena (2008), Blanchard et al. (2015)).

stability. Linking financial busts to the preceding booms is therefore crucial for a more complete assessment of the trade-offs policymakers face.

In this paper we study the role of monetary policy in a setting where, owing to financial frictions related to credit supply, lower interest rates boost risk-taking, thereby generating financial fragility. Financial fragility, in turn, generates sizeable and persistent output losses when stress materialises. As a result, monetary policy regimes have a pivotal role in determining the economy's vulnerability to boom-bust cycles and its long-run evolution. This form of monetary policy non-neutrality represents a key departure from typical frameworks, in which the long-run path is exogenous to monetary policy and the problem is one of cyclical stabilisation around that path.

For the central bank, this set-up generates path dependence – in the sense that future policy depends on constraints that are determined by policy actions in the past – which gives rise to an intertemporal policy trade-off. Easier policy today boosts output in the short run but at the expense of the build-up of financial imbalances and large output losses in the future. By changing its policy rule, the central bank can influence the economy's vulnerability to boom-bust cycles, and hence the long-run range of outcomes for output and real interest rates. While optimal policy involves both leaning and cleaning, it is the degree of leaning during the booms that generates substantive differences in long-run economic outcomes. We illustrate the possibility of a 'low interest rate trap', whereby monetary policy does not lean sufficiently against the build-up of financial imbalances resulting in a growing vulnerability to financial busts over successive cycles, which in turn leads to lower policy rates. In this sense, low rates beget lower rates.

The model features three key building blocks. The first generates long-run money non-neutrality, the second establishes a link between financial and real sectors, while the last gives rise to financial cycles.

First, we adopt an overlapping generations (OLG) setup which allows long-run equilibrium to be compatible with many interest rate levels. As in conventional OLG models with money, and in line with the well-known Tobin effect, the monetary policy rule in our model has long-lasting effects on real allocations and is non-neutral. The OLG setup breaks the link between consumption growth at the *individual* level, as governed by intertemporal substitution, and *aggregate* consumption growth, enabling many equilibria to be sustained at different real interest rates. By contrast, representative agent setups assume a uniform infinite planning horizon, which implies that aggregate demand growth is tied rigidly to individuals' intertemporal substitution decisions. In this case, there can only be one real interest rate level consistent with the goods market equilibrium.[2]

Second, finance is plays an integral role in the economy, acting to underpin economic activity. Financing is essential because firms need to pay for factors of production before output can be produced. Firms can do so only by borrowing from banks, which *generate* purchasing power or inside money. Banks here do not just intermediate saving or endowments, they create money (deposits) in tandem with loans that enable the production of goods. Bank financing

[2]In a similar spirit, much recent work has sought to address the unrealistic implications of representative agent intertemporal substitution, either through the introduction of heterogeneous agents (e.g. Kaplan et al. (2018)), departures from rational expectations (e.g. Gabaix (2019)), or finite planning horizons (e.g. Woodford (2018)).

thus plays a critical role, being a key transmission mechanism of monetary policy as well as a potential important driver of the business cycle.

Third, the banking system is inherently prone to instability, exhibiting phases of excessive risk-taking followed by periods of risk aversion. The corresponding financial cycle reflects two key frictions in the banking system. First, heterogeneous firms have limited liability and cannot be sorted from one another, so that even those with negative-value projects obtain loans. Second, short-lived bank managers focus on short-term profits and, in the presence of externalities, may inadvertently take excessive risk in competing for market share when interest rates are low. Depending on the level of bank capital, the loan market may be subject to multiple equilibria – a boom and a bust.

These features raise the possibility that the economy, including real interest rates, do not converge to a fixed steady state in the long run. Rather, the economy alternates between boom and bust regimes, in the spirit of Kindleberger (2000) and Minsky (1982, 1986). Because interest rates influence risk-taking, the nature of these fluctuations depends on the central bank's reaction function given that monetary policy sets banks' funding costs. The monetary regime is an integral part of the economy's equilibrium. The interaction between monetary policy and the leverage cycle has a long-term impact on the economy, which is of first-order importance in the evaluation of policy trade-offs. Of course, this is not to say that other trade-offs, not least between inflation and output, are less important. We are simply abstracting from these to focus on a particular mechanism that has received less consideration.

Our paper is related to several literature strands. The first is the burgeoning macro-finance literature spurred by the GFC. Much of this work has stressed the potential for financial frictions to amplify the impact of financial shocks. The vast majority of studies overlay those frictions on an otherwise standard real business-cycle framework. Gertler and Kiyotaki (2010) and Cúrdia and Woodford (2016) show how frictions on the lender side can amplify the effects of the traditional financial accelerator mechanism, which focuses on borrower-side frictions (e.g. Bernanke et al. (1999)). Gertler et al. (2017) and Brunnermeier and Sannikov (2014) introduce further amplification through non-linearities and feedback mechanisms. In all of these cases, financial factors simply increase the persistence of the effects of the shocks, rather than generating endogenous boom-bust cycles.[3]

The second literature strand is the growing body of empirical research on the financial cycle. In contrast to the theoretical work, this work strongly suggests that financial busts are linked to the booms that precede them. Many studies have found that strong credit and/or asset prices increases, beyond historical norms, are useful leading indicators of subsequent busts and financial crises (e.g. Borio and Lowe (2002), Borio and Drehmann (2009), Aldasoro et al. (2018), Reinhart and Rogoff (2009), Schularick and Taylor (2012)). It has also become increasingly evident that strong credit growth and/or easy financial conditions carry information about subsequent economic slowdowns (Mian and Sufi (2015), Mian et al. (2017), Claessens et al. (2012), Jordà et al. (2016), Drehmann et al. (2017)), large negative output gaps or possibly deeper recessions (Borio and Lowe (2004), Krishnamurthy and Muir (2017), Jordà et al. (2016),

[3]Relatively small shocks may however be sufficient to cause a boom to implode, given the amplification effect – see Boissay et al. (2016)

Borio et al. (2018)), or downside risks to output (Adrian et al. (2018)).[4]

A third literature strand is the one that has begun to explore the role of monetary policy in boom-bust cycles by focusing on endogenous vulnerabilities (see Adrian and Liang (2016) for a recent review). Building on the 'risk-taking channel' of monetary policy (Borio and Zhu (2012)), Adrian and Duarte (2017) present a New Keynesian model with financial vulnerabilities in the form of endogenous second moments. Within such a setup, it is optimal for monetary policy to lean against the build-up of financial vulnerabilities during times of easy financial conditions as they cause greater future downside risks to growth. Coimbra and Rey (2017) develop a model with time-varying endogenous macroeconomic risk that arises from the risk-shifting behaviour of heterogeneous financial intermediaries. They show that when interest rates are low, further monetary stimulus can increase systemic risk, resulting in a trade-off between growth and financial stability. Neither of these papers, however, deals directly with booms and busts, as we do here. Closer in spirit to ours is Filardo and Rungcharoenkitkul (2016), who examine optimal monetary policy in the context of an endogenous financial cycle and macroeconomic feedback.

A final literature strand includes studies that explore sources of long-run monetary policy non-neutrality. Benigno and Fornaro (2018) and Garga and Singh (2016) generate monetary policy hysteresis in New Keynesian models with endogenous growth. In both cases, deficient aggregate demand hampers aggregate supply by reducing innovation. Caballero and Simsek (2018) also focus on the possibility of a prolonged demand slump when interest rates are at the zero lower bound and the economy is caught in a negative feedback loop between pessimism and asset price declines. With respect to the long-run real interest rate, our work complements a recent set of papers that show how the natural rate of interest can be endogenous to policy in the presence of financial frictions (De Fiore and Tristani (2011), Benigno et al. (2014), Cúrdia and Woodford (2016), Sheedy (2018), and Vines and Wills (2018)). In contrast to these papers, the present one focuses on the non-neutrality that arises from the economy's vulnerability to boom-bust cycles in a setting where monetary policy, rather than saving-investment fundamentals, pins down the real interest rate path.

By exploring the role of monetary policy in influencing boom-bust cycles that have persistent impact on real outcomes, our paper combines these strands of research. It builds on previous empirical work that documents a significant and persistent impact of monetary policy on output through its impact on leverage and debt service burdens (see Juselius et al. (2017) and Juselius and Drehmann (2015)). It also complements the empirical findings of Borio et al. (2017), who document a robust link between monetary policy regimes and real interest rates over the long run. This paper can be seen as providing a theoretical counterpart to these empirical results.

The rest of the paper is organised as follows. The next section sets out the building blocks of a financing-based economy, deriving the absence of a natural rate of interest and policy nonneutrality results. Section 3 introduces lending externalities, and shows how they lead to a financial cycle and an intertemporal policy tradeoff. Section 4 investigates properties

[4]In particular, Adrian et al. (2018) find that the unconditional distribution of output is highly skewed to the left owing to the impact of financial conditions. Specifically, financial conditions boost growth in the near term but sap it in the longer term.

of the calibrated model, and characterises the optimal monetary policy and the associated dynamic equilibrium. Section 5 analyses the effects of different policy regimes on the economy's proclivity to boom-bust cycles, illustrating the possibility of 'monetary policy hysteresis'. The last section concludes.

2 A Financing-based Economy

This section describes the basic 'financing-based economy', where an endogenous creation of credit and deposit by commercial banks facilitate production and trade among overlapping generations of agents (OLG). The need for financing arises from a production lag whereby firms require immediate labour input from households but the output is realised only a period later. Absent the ability to credibly commit the goods proceeds for repayment ex-post, firms are unable to obtain trade credit from households and no production takes place. The commitment problem is resolved by the existence of a universally accepted medium of exchange and store of value. This takes the form of bank deposits, which firms obtain by taking out loans from banks. Bank deposits also serve as a saving vehicle for households, enabling them to engage in intertemporal substitution. This inside money serves an analogous role to fiat money in the traditional OLG model of Samuelson (1958) – see below.

A key basic property of OLG models with monetary assets is that monetary policy is non-neutral, even in the long-run. This follows because equilibrium conditions in asset markets, which are influenced by monetary policy, generally matter for real allocations, a point emphasised by Tobin (1965). Our setup inherits this property. We show that the central bank, by regulating interest rates appropriately can attain the social planner's allocation. Alternative policies would result in different real interest rates and hence, sub-optimal outcomes. These properties mean there is no well-defined natural rate of interest. This setup lays the groundwork for the next section, where money non-neutrality carries over but the efficient allocation is no longer feasible due to frictions that give rise to boom-bust swings in bank financing. In that case, the optimal policy can only attain second-best outcomes, and largely depends on how the central bank contends with intertemporal policy tradeoffs.

2.1 The basic model

Technology and preferences

The economy consists of overlapping generations of households and firms who live for two periods, with a new cohort of equal size born every period. There is a single consumption good that firms can produce using the technology

$$F(L_t) = AL_t^{\alpha} \tag{2.1}$$

where L_t is the labour input that only young households can provide. The production takes one period to complete, and $F(L_t)$ is only available in period $t+1$. Each household chooses how much labour to supply in the first period of life and how to allocate consumption over their lifetime:

$$\max_{C_{1t},C_{2t+1},L_t}\left(\frac{C_{1t}^{1-\sigma}}{1-\sigma}+\rho\frac{C_{2t+1}^{1-\sigma}}{1-\sigma}-\frac{\Phi}{1+\chi}L_t^{1+\chi}\right) \tag{2.2}$$

where C_{1t} and C_{2t+1} are the consumption when young and old, respectively. The parameters $\sigma \in (0,1)$ and $\chi > 0$ determine the curvature of the utility of consumption and disutility of labour, respectively, while $\Phi > 0$ represents the relative weight of the two. $\rho \in (0,1)$ is the household's discount factor.

First-best allocation

The social planner's problem is to maximise households' utility 2.2 subject to the resource constraint

$$C_{1t+1}+C_{2t+1} \leq AL_t^{\alpha} \tag{2.3}$$

In the steady state, this yields the following result (where we suppress the time subscript).

Result 1. *The social planner solution is given by*

$$\Phi L^{\chi+1-\alpha(1-\sigma)} = \alpha(1+\rho^{\frac{1}{\sigma}})^{\sigma}A^{1-\sigma} \tag{2.4}$$

$$C_1 = \frac{1}{(1+\rho^{\frac{1}{\sigma}})}AL^{\alpha} \tag{2.5}$$

$$C_2 = \frac{\rho^{\frac{1}{\sigma}}}{(1+\rho^{\frac{1}{\sigma}})}AL^{\alpha} \tag{2.6}$$

The first-best equilibrium coincides with an Arrow-Debreu allocation where households can frictionlessly trade with firms by offering labour now in exchange for goods later, as well as among each other by borrowing and lending goods across generations.

Bank financing

In the decentralised case, no trade is feasible. As firms can costlessly walk away with goods once production is complete, households have no incentives to supply labour in the first place (a hold-up problem). The lack of trust and ability to pre-commit causes a breakdown in production and trade.[5]

Bank financing provides a solution. By extending new loans, banks create purchasing power in the form of deposits (inside money) which we assume function as a generally accepted means of payment.[6] Firms use deposits to acquire labour and start production; households use them to save and buy goods from firms; and banks accept them from firms as a means to repay

[5]OLG models usually focus on the inefficiency from no intergenerational trade. In our model, an 'autarky' outcome obtains if households possess the production technology but cannot borrow or lend ($C_{2t+1} \leq AL_t^{\alpha}, C_{1t} = 0$). This implies an inefficiently low but positive output $\Phi L^{\chi+1-\alpha(1-\sigma)} = \alpha A^{1-\sigma}$.

[6]Banks deposits are risk-free in our setup. In practice, deposit insurance and regulatory oversight helps ensure trust in bank deposits. Our emphasis on the primacy of financing builds on body of research which highlights the importance of the underlying frictions that give rise to money (see Wright (2018), Jakab and Kumhof (2018), Disyatat (2011), and Kiyotaki and Moore (2002).

loans. Loans and deposits are financial claims and liabilities created by banks, both being IOUs (not physical goods as assumed in some stylised models of banks). Banks have a monitoring technology that prevent firms from defaulting on loans. This gives firms incentives to accept deposits from households as payments for goods in the second period, as firms need deposits to repay loans. Loan repayment contracts firms' and banks' balance sheets and removes money from the economy, completing the financing process.

With bank financing, each household's budget constraint is given by

$$C_{2t+1} \leq R_t^d(W_t L_t - C_{1t}) + \pi_{t+1}^F \tag{2.7}$$

where W_t is the real wage and R_t^d is the gross real interest rate on bank deposits. We let households be born with ownership of the firms of the same cohort, and receive real dividends π_{t+1} (firms' profits) when old.[7] $W_t L_t - C_{1t}$ is the deposit balance that the household carries to the second period, and, together with the interest rate earned, represents its claim on goods. The part of wage that is immediately spent on goods when young, C_{1t}, earns no interest. Deposits are risk-free since they are always backed by claims on firms (see below).

Firms pay for the labour cost $W_t L_t$ upfront, financing it by one-period bank loans. Firms take real wage W_t and the real lending interest rate R_t as given, and choose the scale of production L_t to maximise profit

$$\pi_{t+1}^F = F(L_t) - R_t W_t L_t \tag{2.8}$$

The production cost $R_t W_t L_t$ is the value of goods that must be sold in order to obtain sufficient deposits to repay bank loans.

Banks are run by overlapping generations of 2-period-lived managers, who set the lending interest rate R_t in the first period. They take as given the deposit rate R_t^d, which is directly set by the central bank (we elaborate how shortly). Lending at time t earn banks a net interest income in the following period of

$$NII_{t+1} = (R_t - 1)W_t L_t - (R_t^d - 1)(W_t L_t - C_{1t}) \tag{2.9}$$

The first term on the right hand side is the interest income on the loan $W_t L_t$, while the second term is the interest payment to old depositors. The latter is calculated based on an outstanding deposit balance $W_t L_t - C_{1t}$ as only the deposit held to the second period is interest bearing. Young households spend C_{1t} immediately to buy goods from old firms, who in turn use the deposits to repay loans, extinguishing both.

We assume that banks engage in Bertrand competition and bid down the lending rate R_t to the effective funding cost, i.e. until $NII_{t+1} = 0$. The zero-profit condition pins down the equilibrium lending interest rate R_t for a given R_t^d. Any bank that sets a deposit rate lower than this market funding rate will experience an outflow of all of its deposits to other banks. At the chosen lending rate, banks *elastically* supply loans and automatically create deposits in the process.

[7] We assume a large number of firms and households with diversified portfolios, so that each household has no control rights, leading to the hold-up problem in the absence of banks.

For completeness, we spell out other assumptions that will be important later on. Households own the banks (transferring ownership to the next generation as with firms), and receive fixed but negligible dividends from banks. This means households care about the solvency of banks, but otherwise have no preferences over how the banking business is run. We assume banks have access to a goods storage technology, which they use to hold claims against the rest of the economy. This stock of goods occupies the asset side of banks' balance sheets, which reflect banks' retained earnings on the liability side. At present, zero bank profit implies that bank equity is time-invariant and irrelevant to the analysis. Later, when defaults are allowed, bank capital will be a key state variable.

Central bank

To close the model, we let the central bank set the deposit rate R_t^d as its policy rate, by acting as the intermediary between banks in the interbank market.[8] We assume that banks settle payments among themselves by transferring deposits at the central bank. Any bank that needs to transfer funds to another bank must therefore borrow from the central bank to do so. To facilitate this settlement, the central bank stands ready to lend to banks at an interest rate $R_t^d+\epsilon$ and remunerates balances held with them at an interest rate R^d (where $\epsilon > 0$ is small). Such a setup, which resembles 'corridor systems' used in practice by central banks to set overnight interest rates, ensures that no bank has an incentive to set deposit rates different from R_t^d.

To see this, suppose that all banks initially set deposit rates at R_t^d. If a bank sets a deposit rate lower than its peers, it will suffer a deposit outflow to other banks which it must fund by borrowing from the central bank at $R_t^d + \epsilon$. This higher funding cost means there is no incentive for any bank to deviate. Conversely, any bank that sets a higher deposit rate than R_t^d (say $R_t^d + \Delta$) will attract deposits from other banks and see correspondingly higher balances at the central bank remunerated at R_t^d. The negative spread $(-\Delta)$ rules out such a deviation. In the limit of $\epsilon \to 0$, the equilibrium deposit rate is exactly equal to the policy rate R^d, and the central bank neither requires nor earns real resources in running this operation.[9]

Figure 1 summarises the timeline of each period. First, the central bank sets the risk-free deposit interest rate, which banks take as given. Banks then issue new loans and creating new deposits for young firms, who transfer them to young households in exchange for labour in order to start production. The following period, production is completed and the goods market opens – both old and young households purchase goods using their deposits. Old firms repay banks with these deposits, settling their loans. Appendix A walks through the balance sheet mechanics underlying a financing round in more detail.[10]

[8]Specifying a monetary policy rule is necessary to pin down the return to the monetary asset, in our case bank deposit. In standard OLG models with fiat money, the policy rule takes the form of money supply growth.

[9]There is no other equilibrium. If all banks set deposit rates below R_t^d, each bank will have an incentive to deviate and set a slightly higher rate to attract deposits from other banks which they will receive in the form of balances at the central bank earning R_t^d, and hence a positive spread. Similarly, if all banks set rates above R_t^d, each bank will want to lower their deposit rate, incur outflows of funds which they more than make up by borrowing from the central bank at $R_t^d + \epsilon$.

[10]See also Jakab and Kumhof (2018) and McLeay et al. (2014) for detailed discussion on bank financing and its distinction from other modelling approaches.

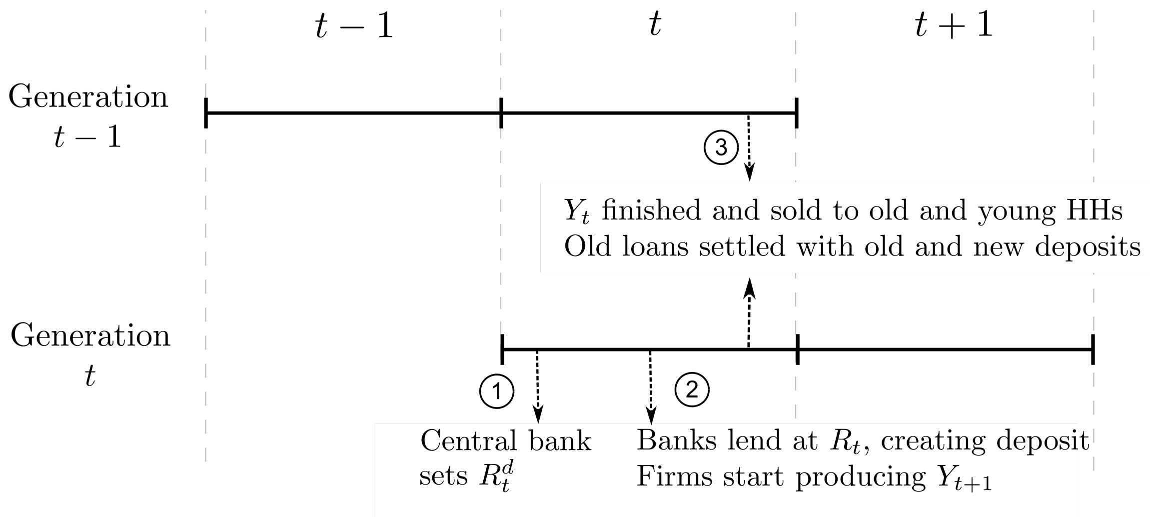


Figure 1: Model timeline

2.2 Financing equilibrium

We now proceed to characterise the equilibrium solutions. Solving the firms' problem, the profit-maximising amount of labour hired is

$$L_t = \left(\frac{\alpha A}{R_t W_t}\right)^{\frac{1}{1-\alpha}} \tag{2.10}$$

which implies firms' profit, to be distributed to households, of

$$\pi_{t+1}^F = \frac{(1-\alpha)}{\alpha} R_t W_t L_t \tag{2.11}$$

The first-order conditions of the household problem imply the Euler equation and the optimal work-leisure choice

$$C_{2t+1}^{\sigma} = \rho R_t^d C_{1t}^{\sigma} \tag{2.12}$$

$$\Phi L_t^{\chi} = \frac{W_t}{C_{1t}^{\sigma}} \tag{2.13}$$

Solving these together with the budget constraint 2.7 and firms' profit 2.11, one obtains the labour supply function

$$\Phi L_t^{\chi+\sigma} = W_t^{1-\sigma} \psi_{1t}^{\sigma} \tag{2.14}$$

where

$$\psi_{1t} \equiv \frac{1 + \rho^{\frac{1}{\sigma}} R_t^{d\,\frac{1}{\sigma}-1}}{1 + \frac{1-\alpha}{\alpha}\frac{R_t}{R_t^d}} \tag{2.15}$$

and R_t is the lending rate on bank loans.

Lastly from the banking side, the net interest margin can be written as

$$NII_{t+1} = (R_t - \psi_{2t})W_t L_t \tag{2.16}$$

where the unit cost of deposit is

$$\psi_{2t} \equiv R_t^d - \frac{R_t^d - 1}{\psi_{1t}} \tag{2.17}$$

For positive net interest rate $R^d > 1$, the unit cost of deposit in 2.17 is less than R^d because for each unit of deposits created in making loans, a part of it is extinguished and thus not liable for interest payments. Perfect price competition in the loan market then implies that the lending rate is set at the marginal financing cost

$$R_t = \psi_{2t} \tag{2.18}$$

so that banks make zero profit every period.

The equilibrium can be solved by combining the labour supply equation 2.14, the labour demand 2.10 and the lending rate 2.18. The resulting equilibrium employment (hence output) is a function of the policy interest rate

$$\Phi L_t^{1+\chi-\alpha(1-\sigma)} = (\alpha A)^{1-\sigma}\left(\frac{\psi_{1t}^{\sigma}}{\psi_{2t}^{1-\sigma}}\right) \tag{2.19}$$

where ψ_{1t} and ψ_{2t} can be solved in terms of only R_t^d, $\psi_{1t} = \alpha\left(1 + \rho^{\frac{1}{\sigma}} R_t^{d\,\frac{1}{\sigma}-1}\right) + (1-\alpha)\left(\frac{R_t^d - 1}{R_t^d}\right)$ and $\psi_{2t} = \alpha\left(1 + \rho^{\frac{1}{\sigma}} R_t^{d\,\frac{1}{\sigma}}\right)/\psi_{1t}$. Using this equation together with the budget constraint 2.7, the Euler equation 2.12 and firms' profit 2.11, household consumption is given by

$$C_{1t} = \frac{\alpha A L_t^{\alpha}}{\psi_{1t}\psi_{2t}} \tag{2.20}$$

$$C_{2t+1} = \frac{\alpha A L_t^{\alpha}}{\psi_{2t}}\left(R_t^d\left(1 - \frac{1}{\psi_{1t}}\right) + \left(\frac{1-\alpha}{\alpha}\right)\psi_{2t}\right) \tag{2.21}$$

The following result summarises the implications.

Result 2. *The bank financing equilibrium has the following properties:*

(i) There is no natural interest rate. Any $R_t^d > 0$ can be supported in the steady state.

(ii) Monetary policy determines real allocations, and is hence non-neutral.

(iii) Optimal policy is to set $R_t^d = 1$ for $\forall t$, which ensures the first-best allocation.

Proof. Provided the equilibrium is non-degenerate, equations 2.19-2.21 and the fact that ψ_{1t} and ψ_{2t} vary with R_t^d readily imply parts (i) and (ii). Non-degeneracy is guaranteed if $\psi_{1t}, \psi_{2t} > 0$, which is true as long as $R_t^d > 0$ as stipulated. For part (iii), note that $R_t^d = 1$ implies $\psi_{1t} = \alpha(1 + \rho^{\frac{1}{\sigma}})$ and $\psi_{2t} = 1$. This implies zero interest rate spread as both loans and deposits yield the same rate of return $R_t = R_t^d = 1$. Substituting these into equations 2.19-2.21 yield the same allocations as the social planner's solutions in equations 2.4-2.6. □

Discussion

The non-neutrality of monetary policy and the absence of a natural real interest rate echo properties of the classic Samuelson's OLG model with fiat money. In that setup, outside money growth affects real allocations by influencing the real rate of return to fiat money (the negative of inflation), and hence saving behaviour. This non-neutrality generalises even in the presence of competing store of value such as the physical capital under standard assumptions about the production function, a result known as the 'Tobin effect'.[11] Result 2 shows that non-neutrality also generalises to the financing economy based on inside money creation, and where the central bank sets the interest rate via open market operation. Appendix B shows that the non-neutrality result also obtains in the case where fiat money is used to finance production and monetary policy is characterised by a rule governing money growth.

Our model illustrates how the uniqueness of a natural interest rate and money neutrality in workhorse New Keynesian models may be sensitive to the representative agent with infinite-horizon construct. With a single representative household, aggregate consumption-saving decision boils down to a single Euler equation which tightly pins the steady-state growth rate of aggregate consumption to the discount factor ρ. In OLG models, multiple overlapping Euler equations mute the impact of any single cohort's saving plan on aggregate demand. In fact, in a steady state, one generation's saving is perfectly offset by dissaving of another. That is, consumption growth of *each cohort* can adjust in response to different interest rates without necessarily changing *aggregate* consumption growth. Multiple real interest rates could thus align aggregate demand with supply and clear the goods market, leaving the central bank to choose which interest rate would materialise. The next section exploits this property and explores a mechanism that links secular movements in real interest rates to the central bank reaction function and fluctuations in bank financing.

Optimal policy $R^d = 1$ is in effect a Friedman rule (except R^d is a real rather than a nominal interest rate). Since financing and grossing up bank balance sheet consumes no real resources, it is socially efficient to do so freely to equate the marginal product of labour to the marginal cost of supplying it. Optimal policy of zero net interest rate delivers the first-best outcome, with banks frictionlessly enabling trades between agents.

The model is cast in real terms, but nominal prices can in principle be incorporated. The extension would require an explicit nominal unit of account. One possibility is to let a fiat money play this role, as set out in Appendix B. Supply of this outside money, expressed in some nominal unit (e.g. dollars), controls its real rate of return, which defines inflation as cash pays zero nominal interest rate. As long as fiat money and bank deposits are imperfect substitutes, the two would earn different interest rates, and our key results would go through qualitatively. Economic activity and real allocations would be driven predominantly by the real deposit and lending rates, while fiat money supply determines inflation and all nominal interest rates.

[11] As Tobin (1965) wrote, *"Equilibrium capital intensity and interest rates are then determined by portfolio behavior and monetary factors as well as by saving behavior and technology."* See also more recent works by Lagos and Rocheteau (2008) and Altermatt and Wipf (2020).

3 The Economy with a Financial Cycle

We now introduce frictions that lead to boom-bust cycles in bank credits. On the production side, we add an unproductive sector that causes loan defaults which weaken banks' balance sheets. On the financial side, banks now operate in an imperfect competition environment and can price loans strategically to gain market share. The overall quality of the borrower pool deteriorates as lending interest rates decline, but which no single bank internalises. Lower lending rates, moreover, make unproductive firms less interest-sensitive when shopping for loans, a credit satiating effect. As a result, strategic complementarity arises among profit-maximising banks – both high and low lending rates (conservative and aggressive lending) can be supported in equilibrium.[12]

We introduce an agency friction between bank owners (households) and managers, which selects an equilibrium and generates a cycle. Bank managers are delegated to run the lending business, but have innate preferences for gaining market share via a low-rate strategy. Left to their own devices, managers tend to over-compete which deplete bank capital, possibly leading to bankruptcy. Owners prevent this by penalising a low-rate strategy increasingly as capital runs low. Managers incur a cost when changing strategy, and only coordinate on a 'regime switch' when capital is sufficiently low (owners' penalties dominate) or high (innate preferences dominate). For intermediate bank capital, history dictates equilibrium strategy due to the switching cost. Persistent spells of conservative and aggressive lending give rise to a financial cycle.[13]

3.1 Adding lending frictions

We start by describing the unproductive sector and the new bank profit-maximisation problem, before laying out the agency frictions between owners and managers and how the outcome gives rise to a cycle equilibrium.

Unproductive sector

The unproductive sector consists of a continuum of unproductive firms and asset owners who are endowed with an asset that unproductive firms require to produce. Banks have no means of differentiating unproductive firms from productive ones. Unproductive firms can utilise one unit of an asset to produce consumption goods, but only succeed with an infinitesimal probability. Due to limited liability, these firms would seek bank loans and produce as long as the best-case return exceeds the total cost of production. Each unproductive firm incurs a unit production cost $R_t q_t$, where q_t is the rental cost of the asset. Unproductive firms differ in their productivities $x > 0$, the density function of which is given by $\omega_d \bar{N}_d / x^{1+\omega_d}$, where $\bar{N}_d$ captures the total number of unproductive firms and ω_d regulates the shape of the distribution (a higher ω_d implies that the density is more concentrated at lower values of x). Integrating this density

[12] Our focus is on the frictions arising from the credit supply side. But a similar mechanism could well operate on the credit demand side. In reality, credit demand frictions are likely to be just as important, e.g. due to over-borrowing and balance sheet overhangs in the non-financial sector.

[13] The mechanism echoes the financial instability hypothesis of Minsky (1982, 1986), though in place of fluctuations in investor sentiment, here it is shifts in bank lending behaviour that drive boom-bust cycles.

function yields the aggregate asset demand

$$N_d(R_t q_t) = \frac{\bar{N}_d}{(R_t q_t)^{\omega_d}} \tag{3.1}$$

As a result, the number of unproductive firms entering the loan market, and hence asset demand, rises as the production cost $R_t q_t$ falls.

On the input side, asset owners are one-period-lived agents who expend effort to prepare the assets before renting them out. Once they receive the rent, they spend all bank deposits to consume immediately.[14] They solve

$$\max_{N_t} E_t \left(\frac{C_{At}^{1-\sigma}}{1-\sigma} - \frac{\Phi}{1+\chi} N_t^{1+\chi} \right) \tag{3.2}$$

subject to

$$C_{At} \leq q_t N_t \tag{3.3}$$

The first-order condition gives rise to the asset supply schedule

$$N_s(q_t) = \left[\frac{1}{\Phi}\right]^{\frac{1}{\chi+\sigma}} q_t^{\frac{1-\sigma}{\chi+\sigma}} \tag{3.4}$$

Equating asset demand and supply, equations 3.1 and 3.4, one obtains the equilibrium asset price q_t. This price is decreasing in R_t. The equilibrium amount of loans unproductive firms obtain also declines in R_t, so that

$$q_t N_t = \frac{\bar{N}}{R_t^{\omega}} \tag{3.5}$$

where

$$\bar{N} \equiv \bar{N}_d^{\frac{1+\chi}{1-\sigma+\omega_d(\chi+\sigma)}} \Phi^{\frac{1-\omega_d}{1-\sigma+\omega_d(\chi+\sigma)}} \tag{3.6}$$

$$\omega \equiv \frac{(1+\chi)\omega_d}{(1-\sigma+(\chi+\sigma)\omega_d)} \tag{3.7}$$

Note that ω is rising in ω_d (with $\omega = 1$ when $\omega_d = 1$), and the elasticity of asset demand translates into the elasticity of unproductive firms' demand for loans (i.e. the extensive margin's sensitivity to the production cost). We will calibrate ω directly below, on the understanding that it reflects the deep parameter ω_d.

Imperfectly competitive banks

Banks cannot differentiate firm types but know the aggregate quantity of loans demanded by unproductive firms at any given lending interest rate.[15] Thus aggregate bank profits depend

[14] It is also possible to assume, less restrictively, that asset owners live for two period and face a well-defined optimisation problem. Doing so does not affect the key results.

[15] If banks were able to specify both the loan amount and the interest rate, they could in principle exploit the heterogeneity of loan demand across types and screen for productive firms. We rule out this possibility by

on underlying loan quality, and are equal to net interest income from productive loans $(R_t - \psi_{2t})L_t W_t$ (equation 2.16) minus the loss on the loans to unproductive firms of $q_t N_t = \bar{N}/R_t^\omega$ (equation 3.5). To allow loan losses to matter for bank profits, we depart from the baseline case and assume that the loan market is imperfectly competitive. Individual banks may thus set different lending rates from their peers in an attempt to capture market share and influence loan composition. Specifically, we assume each small bank i takes the aggregate variables L_t, W_t, R_t^d, as well as the average lending rate set by its competitor banks R_t as given, and chooses its lending rate R_{it} to maximise its profit

$$\Pi(R_{it}, R_t) = (R_{it} - \psi_{2t})L_t W_t \left(\frac{R_t}{R_{it}}\right)^{\theta_1} - \frac{\bar{N}}{R_t^\omega}\left(\frac{R_t}{R_{it}}\right)^{\theta_2} \tag{3.8}$$

which is the sum of net interest income on the loans granted to productive firms and the loss on the loans granted to unproductive firms, adjusted by the corresponding market shares.[16] The market shares depend on how each bank prices its loans relative to competitors. $\theta_1 > 1$ and $\theta_2 \geq 0$ capture the interest sensitivities of loan demand for productive and unproductive firms, respectively, and represents banks' market power. Appendix C discusses the microfoundations for the market shares in equation 3.8.

Note the first source of externality – no single bank can on its own influence the overall quality of the pool of borrowers $\bar{N}/R_t^\omega$, which is determined by the aggregate lending rate R_t. Banks take R_t as given, failing to internalise the effect of their decision on it and on aggregate loan quality. This creates an incentive to over-lend, in line with the standard 'tragedy-of-the-commons' problem.

The second externality we introduce weakens banks' incentives to poach firms from their competitors when interest rates are high, and strengthen them when rates are low. When overall lending interest rates are high and credit relatively scarce, we assume that unproductive firms are more willing to switch to banks that offer cheaper loans – their loan demand is more price elastic at higher interest rates.[17] The assumption allows us to introduce strategic complementarity in a simple way. If a bank were to undercut competitors in this environment, it would disproportionately attract unproductive firms, leading to higher loan losses. If all banks lend conservatively and set a high lending rate R, then any individual bank would have little incentive to poach.

assuming that banks meet any loan demand at the chosen lending rate. More generally, what is central for the result is that the size of the loans extended to productive and unproductive activities are correlated. For instance, this would be the case if banks can only lend to a consortium of firms or production units.

[16]Each bank takes its funding cost ψ_{2t} as given when optimising, as ψ_{2t} depends only on R_t^d and the banking system's average lending rate R_t which an individual bank cannot influence. Note that there is no deposit funding cost for loans to unproductive firms because the corresponding deposits are paid out to asset owners who, in turn, spend them immediately, thereby extinguishing the deposits within the period.

[17]One interpretation is that firms face a trade-off between their preferences for bank diversity (the source of banks' market power) and obtaining the cheapest possible loans (see Appendix C). We assume that the shape of preferences is such that the second consideration tends to dominate when overall financial conditions are tight. This is consistent, for example, with empirical evidence of a higher price elasticity at higher interest rates, leading to a 'kinked' loan demand function – see Karlan and Zinman (2008).

Specifically, we let $\theta_2(R_t)$ be an increasing step-function of R_t

$$\theta_2 = \begin{cases} \theta_H & \text{under high-}R_t \text{ equilibrium} \\ \theta_L & \text{under low-}R_t \text{ equilibrium} \end{cases} \tag{3.9}$$

where $\theta_H > \theta_L$. The first-order condition from equation 3.8 implies bank i's best-response function

$$R_{it} = \frac{\theta_1 \psi_{2t}}{\theta_1 - 1} + \frac{\theta_2 \bar{N}}{(\theta_1 - 1) L_t W_t R_t^{\omega}} \left(\frac{R_t}{R_{it}} \right)^{\theta_2 - \theta_1} \tag{3.10}$$

In the symmetric Nash equilibrium where $R_{it} = R_t$, the aggregate loan supply function dictates that the lending rate must solve

$$R_t = \frac{\theta_1 \psi_{2t}}{\theta_1 - 1} + \frac{\theta_2 \bar{N}}{(\theta_1 - 1) L_t W_t R_t^{\omega}} \tag{3.11}$$

The first term shows that banks exercise market power by setting a mark-up over the funding cost ψ_{2t} when lending to productive firms. As individual banks' incentive to steal market share increases (θ_1 is higher), the Nash equilibrium mark-up declines towards 1. The second term reflects incentives to contain losses from lending to unproductive firms, internalised up to the parameter θ_2. Strategic complementarity arises because θ_2 is increasing in R_t. At higher interest rates, the loan demand by unproductive firms is more interest rate sensitive, reducing individual banks' incentive to undercut competitors and hence limiting loan losses. Thus when average interest rates are high, each bank wants to maintain a high rate. There are two Nash equilibria: a 'bust' with high R_t, where banks are conservative and careful about undercutting each other ($\theta_2 = \theta_H$), and a 'boom', with low R_t, where banks compete more aggressively for customers ($\theta_2 = \theta_L$).

Despite profit maximisation, a bank may well make a loss in equilibrium, because each has no control over the aggregate share of unproductive firms. It must take the system-wide profit or loss as given. Raising its own lending rate R_{it} may deter unproductive firms, but only at the cost of turning away productive borrowers. Doing so may in fact lower profit even further if unproductive firms are relatively price-insensitive, as is the case in a boom. As a result, a bank cannot guarantee itself a non-negative profit. We assume R_{it} is the only choice variable of a bank, and there is no option to shut down the business temporarily to avoid a bad equilibrium (as Charles Prince famously remarked, "As long as the music is playing, you've got to get up and dance").

Agency frictions and equilibrium selection

We now describe agency frictions between ownership and management. Households own the banks and delegate the running of them to two-period-lived managers, who receive a (small) fraction γ of profit $\Pi_{t+1}(R_{it}, R_t)$ as remuneration. This incentivises managers to maximise bank profit and follow the best-response rule derived in equation 3.10. The equilibrium lending rate R_t is thus one of the two solutions to equation 3.11, with lower and higher values denoted by R_L and R_H respectively (dropping time subscript to ease notation, with an understanding that these depend on R_t^d). As it is common knowledge that all managers pursue to same objective, the strategic interaction between managers is reduced to a 2-by-2 coordination game, with action space $\{R_L, R_H\}$.

We assume that bank managers have an innate predisposition to expand market share, which can be achieved by setting low interest rates. This is captured by assuming that managers derive a positive utility D from lending at rate R_L. Bank owners have an incentive to counter this bias to guard against bankruptcy that may arise from over-competition. They do so by imposing a penalty $H(K_t)$ on managers for each period that they choose R_L, more so at lower bank capital levels K_t – formally $H'(K) < 0$ and $\lim_{K \to 0} H(K) = \infty$ so that bankruptcy never occurs.[18] [19] Lastly managers incur a switching cost C each time they shift between R_L and R_H (e.g. the cost of having to explain to the board and customers the reasons for such change). A bank manager's payoff is thus given by

$$\Pi^M(R_{it}, R_{it-1}|R_t) = \gamma\Pi(R_{it}, R_t) + H(K)I(R_{it} = R_H) \\ +DI(R_{it} = R_L) - CI(R_{it} \neq R_{it-1}) \tag{3.12}$$

where $I()$ is an indicator function. We assume that $\gamma, C, D, H(K)$ are all small and can be ignored when computing bank profits or social welfare. γ is also small relative to $C, D, H(K)$ (see below).

For a bank manager, switching to a different lending strategy yields a payoff that depends not only on her private cost $H(K)$ and C, but also what she expects other managers to do. Determining the switching point in equilibrium is thus an equilibrium selection problem. We adopt a dominant strategy criterion, according to which the equilibrium switches, say from R_L to R_H, when each manager finds it optimal to do so regardless of what others choose.[20] That is

$$\Pi^M(R_H, R_L|\tilde{R}) \geq \Pi^M(R_L, R_L|\tilde{R}) \tag{3.13}$$

for $\tilde{R} \in \{R_L, R_H\}$. Combining this with equation 3.12 and exploiting the assumption that γ is small (relative to C, D, and $H(K)$), the switching condition from R_L to R_H boils down to

$$H(K) - D \geq C \tag{3.14}$$

Conversely, the switching from the R_H equilibrium to the R_L counterpart occurs when

$$H(K) - D \leq -C \tag{3.15}$$

These conditions define the threshold levels of bank capital at which equilibrium switches take

[18]Consistent with this agency cost assumption is some recent empirical evidence that higher bank capital leads to greater risk-taking. See Dell'Ariccia et al. (2017).

[19]The fact that $H(K)$ is state contingent could be justified by owners having to incur monitoring and due diligence costs each time they decide to assess banks' insolvency risk and managers' strategy. $H(K)$ could thus represent the expected penalty, decreasing in K partly because owners are more likely to begin due diligence process when bank capital runs low.

[20]When γ is small, as we assume, it turns out that the Harsanyi-Selten *risk dominance* equilibrium selection concept delivers exactly the same outcome.

place, namely

$$K_c \equiv H^{-1}(C + D) \tag{3.16}$$

$$K_r \equiv H^{-1}(D - C) \tag{3.17}$$

where $K_c < K_r$ (subscripts standing for 'crisis' and 'recovery') because $H'(K) < 0$.

Intuitively, when bank capital falls below K_c, the owner penalty dominates and pushes all bank managers to coordinate on R_H. While this conservative lending equilibrium prevails, all banks benefit from lower loan losses, earn positive profits and build up capital. Higher capital level lowers the threat of penalty relative to managers' preferences for competition ($H(K) - D$ declines). But due to the switching cost, it takes time before managers switch back to aggressive lending at $H(K) - D \leq -C$. For intermediate bank capital $K_t \in (K_c, K_r)$, the prevailing equilibrium regime thus persists into the next period, a form of hysteresis. See Bebchuk and Goldstein (2011) and Rajan (1994) for a related approach in dealing with multiplicity.

We allow C to be a random variable, and hence equilibrium switches stochastic. This is the only source of uncertainty in the model. We assume that the stochastic properties of C, through equations 3.16-3.17, induce a logistic probability distribution of regime switches over bank capital K_t. Parameters K_c an K_r are the centres of these distribution, and the transition probabilities between the two equilibria or regimes are

$$\begin{aligned} P(s_t = 2|s_{t-1} = 1, K_t) &= G(a_c(K_t - K_c)) \\ P(s_t = 1|s_{t-1} = 2, K_t) &= G(a_r(K_t - K_r)) \end{aligned} \tag{3.18}$$

where $G(x) = e^x/(1 + e^x)$ is the logistic function, $a_c < 0, a_r > 0$. The regime $s_t = 1$ (a boom) corresponds to the equilibrium where banks compete actively with $R_t = R_L$. The regime $s_t = 2$ (a bust) is associated with the conservative lending equilibrium with $R_t = R_H$. This formulation nests the deterministic case when a_c and a_r are large, though we will focus on the more interesting case where both are finite.[21]

Bank capital is a key state variable in the economy, which in a symmetric Nash equilibrium evolves according to

$$K_{t+1} = K_t + \Pi(R_t, R_t) \tag{3.19}$$

Bank capital reflects the various stages of the financial cycle. In the initial stages of a boom, bank capital is ample and banks can withstand an extended period of losses. In the early stages of a bust, bank capital is low and takes time to build up. Despite the persistence of bank risk-taking, regime switches occasionally reverse the bank capital dynamics, ensuring that in equilibrium banks never become bankrupt nor accumulate net worth indefinitely.

The absence of a natural interest rate applies also in this richer model with lending frictions. Note that banks hold goods on the asset side of their balance sheets, which mirror bank equity capital. When banks make profits, they are acquiring claims on goods in the economy. Firms transfer these goods to banks as part of their loan settlement, when their loan obligations

[21] The deterministic case implies predictable crises, and the central bank can always stabilise bank capital just above K_c to avoid them. There is no financial cycle and policy tradeoff in such case.

exceed the deposits they acquire through sales of goods. Similarly, when banks make losses, they run down their goods inventory to meet the excess of deposit withdrawals over loan repayments. This serves to make deposits risk-free. The evolution of bank capital then reflects banks' profit and loss – banks accumulate capital when making profit, and run down capital when making losses.[22] Bank inventories of goods serve the same function as inventories in national income accounting (as part of business investment). In our model, bank financing ensures that these adjust to absorb any demand excesses or shortages, thus coordinating aggregate production with aggregate demand. Goods market clearing again obtains without the need for the interest rate or other prices to adjust.

3.2 Equilibrium conditions and the financial cycle

The equilibrium conditions consist of the labour supply in equation 2.14, labour/credit demand in equation 2.10, and the regime-specific lending interest rate in equation 3.11:

$$L_t = \left(\frac{W_t^{1-\sigma} \psi_{1t}^{\sigma}}{\Phi} \right)^{\frac{1}{\chi+\sigma}} \tag{3.20}$$

$$L_t = \left(\frac{\alpha A}{R_t W_t} \right)^{\frac{1}{1-\alpha}} \tag{3.21}$$

$$R_t = \begin{cases} \dfrac{\theta_1 \psi_{2t}}{\theta_1 - 1} + \dfrac{\theta_L \bar{N}}{(\theta_1 - 1) L_t W_t R_t^{\omega}}, & \text{for } s_t = 1 \\ \dfrac{\theta_1 \psi_{2t}}{\theta_1 - 1} + \dfrac{\theta_H \bar{N}}{(\theta_1 - 1) L_t W_t R_t^{\omega}}, & \text{for } s_t = 2 \end{cases} \tag{3.22}$$

For a given regime, these three conditions pin down a unique equilibrium in L_t, W_t and R_t period by period, at *any* level of policy interest rate R_t^d. The solutions imply all other variables, including consumption, output as well as firms' and banks' profits. Finally, the laws of motion for bank capital in equation 3.19 and the regime-switching process in equation 3.18 govern the economy's evolution.

The relationship between output and interest rates is subject to two opposing forces. In equation 3.20, labour supply depends positively on R_t^d due to an income effect – an increase in interest income makes households more willing to work. This dampens the usual labour demand effect, which lowers employment as rates rise. For some parameter configurations, the income effect could dominate the labour demand effect, so that higher rates would go hand-in-hand with higher employment and output. This is more likely in a bust because the lending rate, and hence labour demand, is less responsive to policy. We rule out this possibility of an upper 'reversal interest rate', where a sufficiently high interest rate could be expansionary, by adding a technical assumption that the labour supply is perfectly elastic at W_{min}, calibrated to always be binding in a bust. As a result, changes in employment during a bust stem entirely from labour demand shifts.

A financial boom-bust cycle exists when bank profits are persistently negative during the boom $s_t = 1$, and positive during the bust $s_t = 2$. Under such condition, a boom would deplete bank capital and inevitably trigger a bust, while a bust would restore bank capital and

[22] In Appendix A, we explain the balance sheet mechanics in more detail via explicit numerical examples.

eventually usher in a boom. The economy never reaches a resting steady state as a result, but goes through phases of declining and rising bank capital, aggressive and conservative lending as well as high and low economic activity. A financial cycle is more likely to emerge the stronger the lending frictions, which under a special case can be expressed as simple parametric restrictions.

Result 3. *When $\theta_L = 0$ and $R_t^d = 1$, sufficient conditions for a financial cycle to exist are*

$$\theta_H > \theta_1 - 1 \tag{3.23}$$

$$\kappa > 0 \tag{3.24}$$

$$\theta_1 > \theta_1^* \tag{3.25}$$

where $\kappa \equiv \omega - \frac{1+\chi}{\chi+\sigma+(1-\alpha)(1-\sigma)}$ and θ_1^* solves

$$\frac{(\theta_1^* - 1)^{\kappa+1}}{\theta_1^{*\kappa}} = \frac{1}{\bar{N}} \left[(\alpha A)^{1+\chi} \left(\frac{\psi_1^\sigma}{\Phi} \right)^\alpha \right]^{\frac{1}{\chi+\sigma+(1-\alpha)(1-\sigma)}}$$

Proof. See Appendix D. □

In other words, a financial cycle tends to arise when θ_H is high relative to θ_1, and θ_1 is high relative to θ_L (assumed zero here). Intuitively, high θ_H and low θ_L implies that unproductive firms can have very different effects on the lending market equilibrium and bank profitability depending on which regime prevails. In the next section, we calibrate the parameters to satisfy these sufficient conditions as a basic requirement. However, in the case of optimal time-varying R_t^d, the conditions for financial cycle to exist become more complicated and cannot be pinned down in closed-form.[23]

4 Simulation and the Optimal Policy

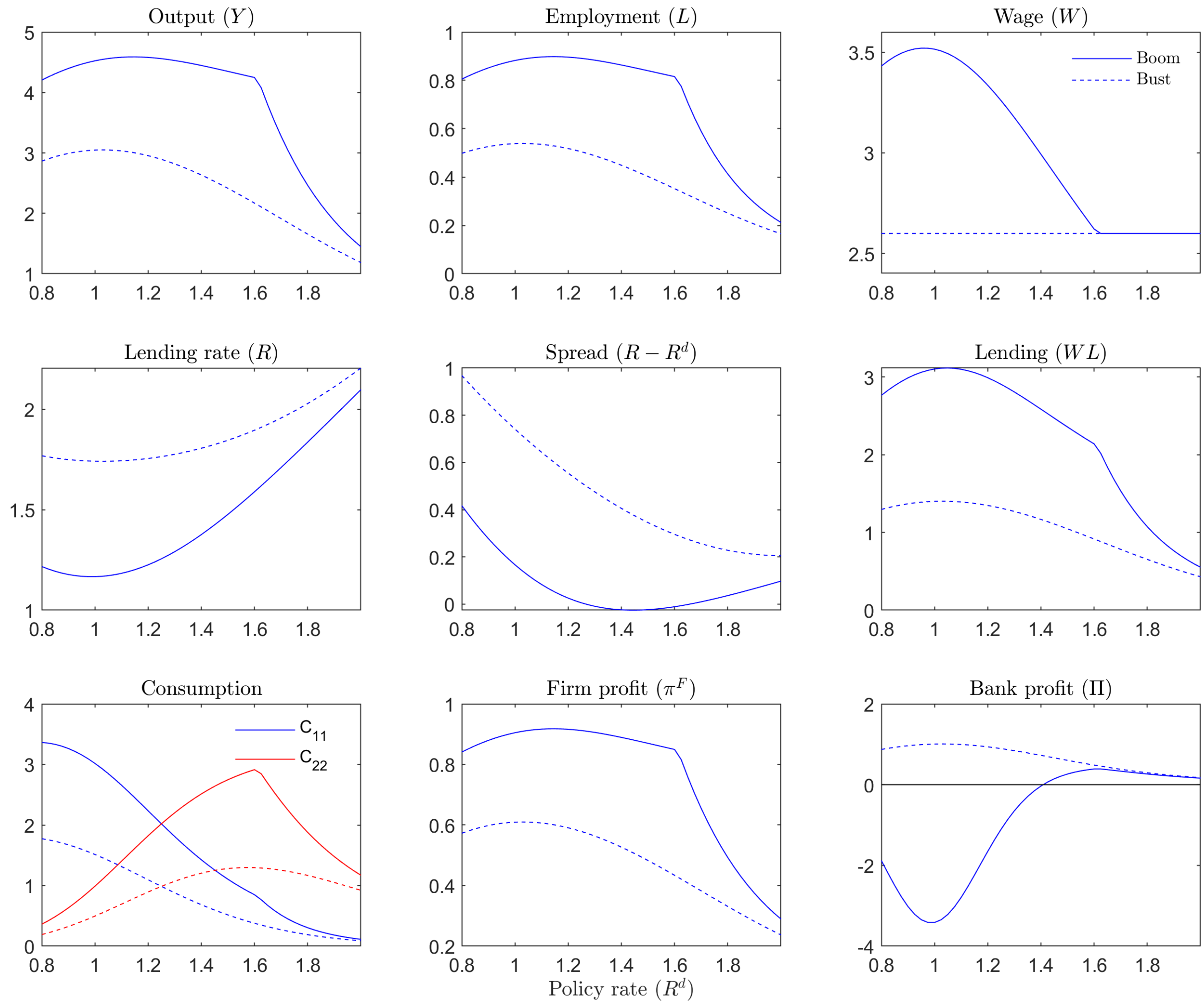
4.1 Period equilibrium

We numerically solve for the equilibrium values of endogenous variables conditional on the level of R^d, using the set of parameters in Table 1. Those parameters pertaining to households' intertemporal problem are standard. Others are calibrated such that a financial cycle does arise in equilibrium, to illustrate the full implications of the model with banking sector externalities.[24] Figure 2 plots the equilibrium values of the endogenous variables in an arbitrary period (vertical axis), as a function of the contemporaneous policy interest rate R^d (horizontal axis). Each panel displays the equilibrium functions for both boom and bust regimes (solid and dash lines, respectively). A number of features are worth highlighting.

First, monetary policy influences the real economy not only through aggregate demand, but also through supply. A lower R^d pushes down the lending rate R in equilibrium. This raises

[23] The challenge is the mutual dependence between optimal policy and the financial cycle. As the optimal policy problem considered here is highly nonlinear and not available in closed-form, it is also impossible to characterise the financial cycle analytically. We rely on numerical methods in the following.

[24] The model clearly nests the case where a financial cycle does not arise, e.g. when $\bar{N} = 0$ so that there are no unproductive firms.



Note: This figure plots the equilibrium values of endogenous variables in each period (vertical axis), for any given level of policy interest rate R^d (horizontal axis). Solid and dash lines correspond to boom and bust regimes, respectively.

Figure 2: Period equilibrium and policy interest rate

Table 1: Calibration assumption

Parameters	Values	Parameters	Values
ρ	0.8	ω	12
Φ	3	$\bar{N}$	25
α	0.8	K_c	4
χ	0.5	K_r	16
σ	0.2	a_c	-1
A	5	a_r	1
θ_1	7	β	0.9
θ_L	0	W_{min}	2.6
θ_H	150		

Note: This table reports the baseline parameters used to calibrate the model and compute equilibrium as well as optimal policy.

productive firms' demand for labour, which in turn increases wages, employment and output. And as R^d falls, households bring forward consumption – a standard Euler equation result. In the figure, this is reflected in the rise in consumption of the young (C_{11}) and decline in those of the old (C_{22}). Alongside the expansion of aggregate consumption demand, productive firms' production increases resulting in a boost to output, as firms take advantage of lower funding costs to borrow more and scale up production. This is consistent with the 'cost channel' of monetary transmission, which emphasises the real effects of monetary policy working through the supply side, here in the form of variations in the marginal cost of production linked to changes in interest expenses (e.g. Ravenna and Walsh (2006) and Gaiotti and Secchi (2006)).

Second, while lower interest rates expand firm profits, they generally *reduce* bank profits (bottom right panel of Figure 2). This is due to the higher participation of unproductive firms, which depresses the overall quality of the borrower pool. In this calibration, there exists a cut-off level for R^d below which bank profits become negative. This creates a key tension: a more accommodative monetary policy boosts economic activity today, but saps bank profitability and drains banks' net worth. Banks take on too much risk amidst intense competition.[25]

Third, in the bust regime (dashed lines), banks are more conservative and set a higher R for any given R^d. In other words, the interest rate spread $R - R^d$ jumps as the economy moves from a boom to a bust, reflecting banks' increased conservatism.[26] Consequently, output, employment, consumption and firms' profits are lower for any given level of R^d. Because of the fall in labour demand, the minimum wage becomes binding at every level of R^d. The binding minimum wage ensures that a lower policy rate is still expansionary for the economy during

[25]Negative bank profits during the boom may seem contrary to the observation that banks usually report strong earnings during the upswing. But it is precisely during the upswing that key fragilities build up, obscured because either asset prices are inflated or accounting valuations are backward-looking. Thus, one can interpret our results as simply reflecting a more accurate measurement of the true underlying health of banks, incorporating also the vulnerabilities usually hidden from view. Put differently, our model indicates that *even if there is no bias* in the measurement of risk and valuations, boom-bust dynamics can arise. For recent empirical evidence, see Haubrich (2015) who documents a countercyclical behaviour of bank capital.

[26]The spread is in general non-monotonic in the policy rate R^d. On the one hand, a higher interest rate improves the loan pool quality, lowering the required compensation for loan defaults and the spread. On the other hand, a higher interest rate induces households to save more, increases the proportion of deposits that are held to maturity, and raises banks' funding cost ψ_{2t}, putting an upward pressure on the spread.

the bust, even though the policy transmission to the lending rate R is weaker. However, bank profits are higher in this regime, as banks seek to contain bad loans. A bust is painful while it lasts, but allows banks to re-build their capital and repair their balance sheets even at a low risk-free rate.

Finally, note that monetary policy is subject to an effective lower bound. Figure 2 shows that for sufficiently low R^d (approximately less than 1), the lending rate R increases as monetary policy is eased. This second instance of a reversal rate is directly related to bank intermediation (similar to Brunnermeier and Koby (2019)), and thus we let it define the effective lower bound for monetary policy. To see why it exists, note that the effective deposit cost for the bank ψ_{2t} (equation 2.17) is increasing in ψ_{1t} when $R_t^d > 1$, and decreasing when $R_t^d < 1$. Intuitively, below this inflection point, banks are earning rather than paying money to depositors (because of the negative *net* interest rates) so that a reduction in the policy rate leads to lost income on deposits rather than lower costs. Banks try to offset this by increasing lending rates, leading to perverse output effects.

4.2 Optimal monetary policy

We next endogenise the setting of R_t^d in each period, and let the central bank set the policy rate R_t^d to maximise the social welfare function. This is defined in each period to be an equally-weighted sum of the welfare of households and asset owners – the only agents that consume in the economy – discounted over the infinite future. The policy problem at each time t is

$$\max_{R_t^d} E_t \sum_{\tau=t}^{\infty} \beta^{\tau} \left(W_{H,\tau} + W_{A,\tau} \right) \tag{4.1}$$

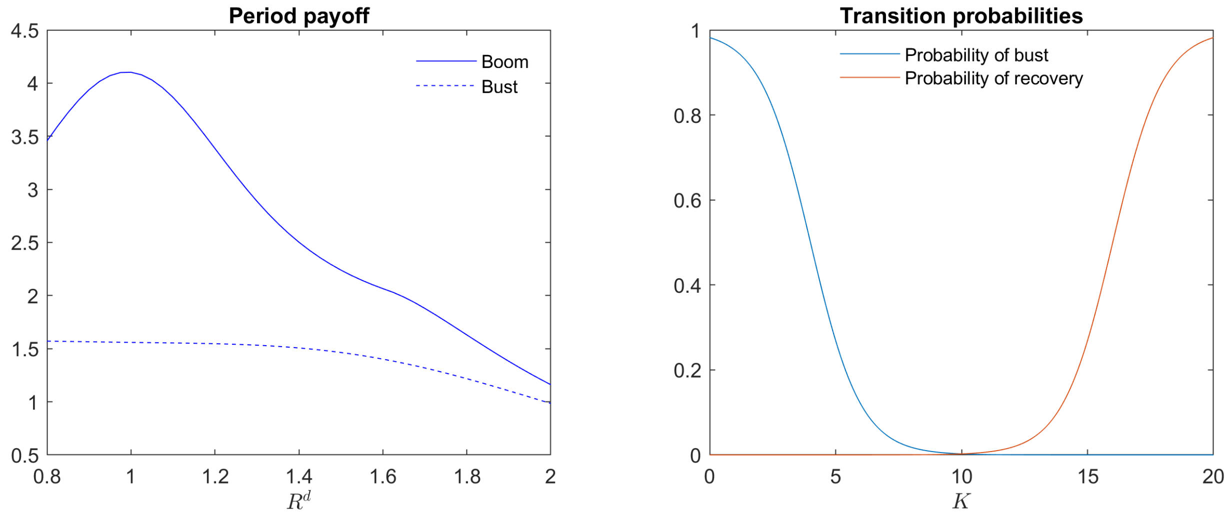
where

$$W_{H,t} = E_t \left(\frac{C_{1t}^{1-\sigma}}{1-\sigma} + \rho \frac{C_{2t+1}^{1-\sigma}}{1-\sigma} - \frac{\Phi}{1+\chi} L_t^{1+\chi} \right) \tag{4.2}$$

$$W_{A,t} = E_t \left(\frac{C_{A,t}^{1-\sigma}}{1-\sigma} - \frac{\Phi}{1+\chi} N_t^{1+\chi} \right) \tag{4.3}$$

are the present value lifetime utilities of households and asset owners in period t, respectively. The discount factor β describes how the central bank compares the payoffs and hence consumption of different generations. It is a normative parameter, and need not equal households' rate of time preference ρ.

The setup gives rise to an intertemporal trade-off for monetary policy. As described above, monetary policy does not just affect output and consumption, but also the degree of risk-taking during booms and hence bank profits and the evolution of bank capital. A higher policy rate today restrains the boom, at the cost of more subdued growth in the short run. At the same time, this 'leaning' policy promotes a more robust financial sector and reduces the risk of a bust with severe output consequences down the road. The optimal policy problem involves managing this trade-off.



Note: The left panel shows the period payoff, the contemporaneous aggregate utility of the agents in the economy as a function of the policy interest rate R^d (horizontal axis), corresponding to boom and bust regimes (solid and dashed lines, respectively). The right panel shows the transition probabilities between boom and bust regimes, depending on the initial regime (blue for boom, orange for bust) and the level of bank capital K (horizontal axis).

Figure 3: Period welfare and regime switching probabilities

The left panel of Figure 3 shows the short-term benefit in terms of the period payoff

$$W(R_t^d, s_t, K_t) \equiv W_H(R_t^d, s_t) + W_A(R_t^d, s_t) \tag{4.4}$$

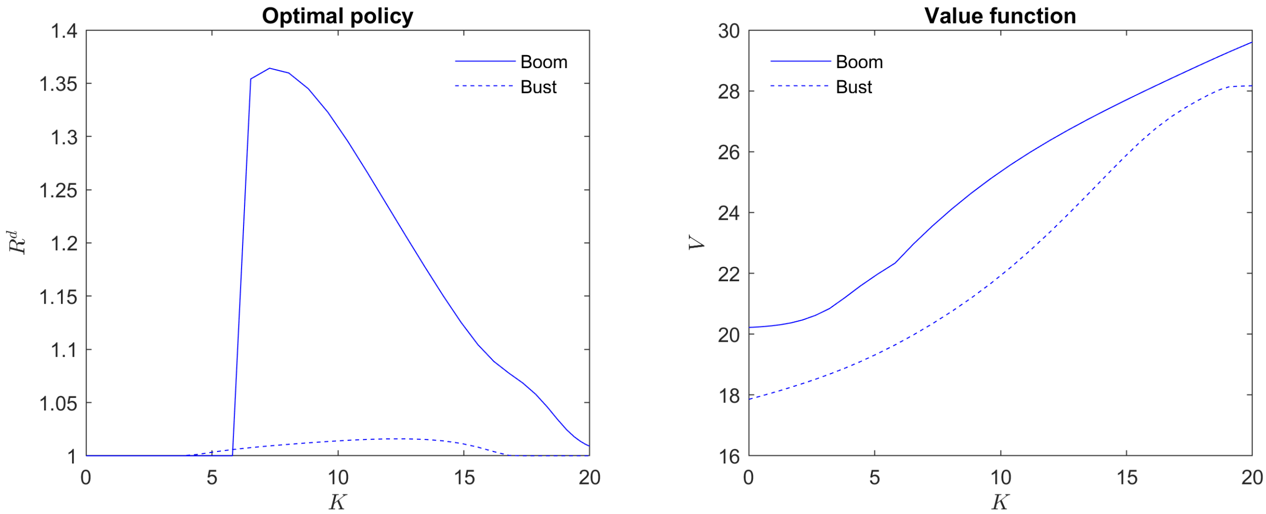
corresponding to each regime s_t. As can be seen, the period payoff in the boom dominates that in the bust for any policy interest rate R^d. And in both regimes, there is an incentive to set a low policy interest rate and stimulate the economy in the short-term.

Implementing a low R^d, however, feeds excessive competition and risk-taking, which adversely affects bank earnings. At a low enough R^d, bank profits turn negative, eroding banks' net worth and pushing the economy closer to a bust. The probability of this regime switch rises as K declines (right panel of Figure 3, blue line). In addition, the probability of recovering from a bust is positive only after K has reached a sufficiently high level (same panel, red line). Because the recovery threshold is higher than the bust threshold ($K_r > K_c$), both boom and bust regimes are persistent states. Once in a bust, it takes time to escape from it. And during the boom, policy has a window of opportunity to avoid a bust.

The resulting policy problem, which is captured in equation 4.1, can be formally written in recursive form as

$$V(K_t, s_t) = \max_{R_t^d} \left[W(R_t^d, s_t, K_t) + \beta \left(\sum_{s_{t+1} \in \{1,2\}} P(s_{t+1} | s_t, K_t) V(K_{t+1}(R_t^d), s_{t+1}) \right) \right] \tag{4.5}$$

subject to the dynamic evolution of the regime and bank capital in equations 3.18 and 3.19,



Note: The left panel shows the optimal level of policy interest rate R^d, as a function of the prevailing regime (solid for boom and dashed for bust) and the level of bank capital K (horizontal axis). The right panel shows the corresponding value functions under optimal policy.

Figure 4: Optimal lean-against-the-wind policy

reproduced here.

$$P(2|1, K_t) = G(a_c(K_t - K_c)) \tag{4.6}$$

$$P(1|2, K_t) = G(a_r(K_t - K_r)) \tag{4.7}$$

$$K_{t+1} = K_t + \Pi(R_t(R_t^d), R_t(R_t^d)) \tag{4.8}$$

What does optimal monetary policy look like? The answer is that it 'leans against the wind'. The solution to this dynamic programming problem is shown in Figure 4. The left-panel presents the optimal choice of R^d as a function of bank capital K and regime s. In a boom, the optimal policy generally involves giving up some short-term gains by setting R^d at a higher level than that which maximises the period payoff (as we saw in the left hand side of Figure 3, optimal short-term policy would set R^d close to 1). This 'leaning-against-the-wind' policy confers two key benefits. First, it promotes a stronger banking system and lowers the probability of entering a costly bust, effectively internalising the effects of excessive bank competition. Second, by moderating the boom and maintaining higher bank capital on average, it strengthens the resilience of the banking system to face a bust, should it occur. This, in turn, shortens the expected duration of a bust, which boosts welfare.

The optimal policy has a non-linear profile. The central bank leans moderately when bank capital is ample, and steps up the degree of leaning as bank capital declines. At its peak, leaning almost completely stabilises bank capital and fully offsets the bankers' incentives to over-compete (we consider the case where bank capital is indeed fully stabilised below). For a sufficiently low K, when a bust becomes imminent, optimal policy calls for a sharp reduction in the leaning intensity. In this region, the cost of leaning no longer justifies the benefit, and the central bank switches to implementing a very low interest rate to stimulate short-term activity.

In the bust regime, the central bank has significantly less control over the financial system's evolution, as banks enter a high risk-aversion and balance-sheet-repair mode – equivalent to pushing on a string. Here bank profits are much less sensitive to changes in R^d (see bottom right panel of Figure 2). In addition, bank profits are now positive at any level of the policy rate, which implies that there is no financial stability benefit of setting a high R^d_t. In fact, a lower interest rate, by stimulating economic activity at a time when banks are conservative, provides additional support to bank profits and helps the economy recover faster.

Result 4. *During booms, optimal policy leans against the wind. In busts, optimal policy eases to stimulate economic activity to the extent possible.*

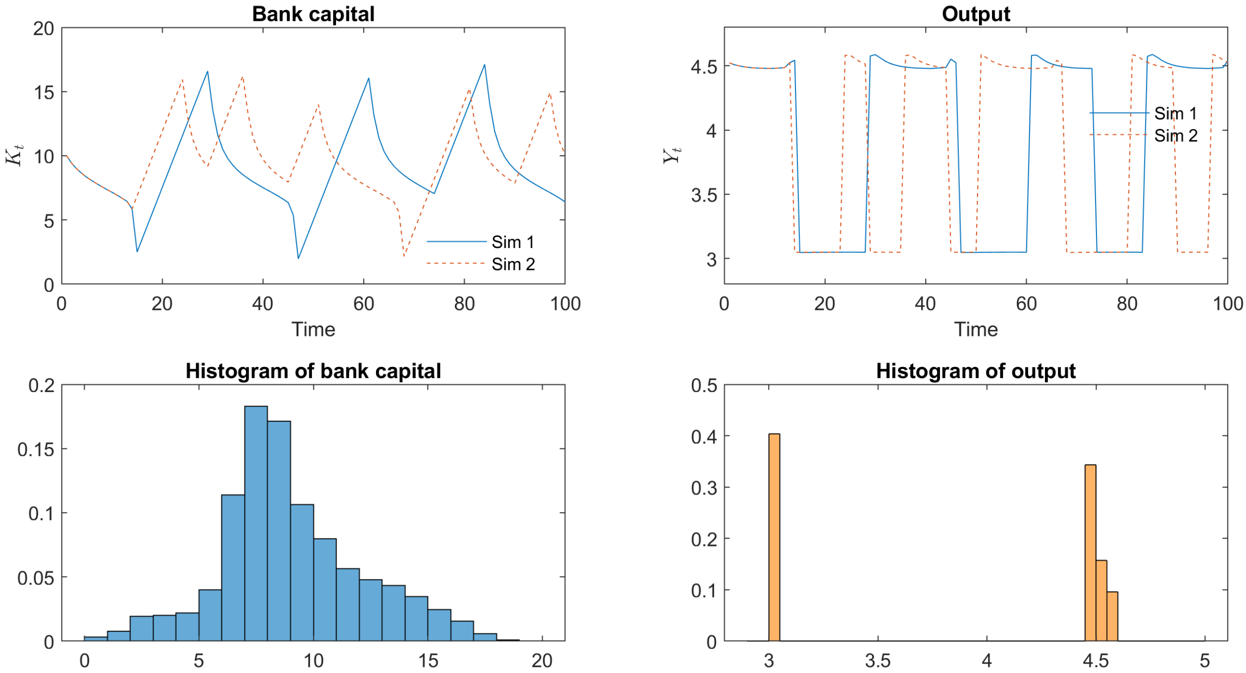
Thus, the results support both 'leaning' and 'cleaning' policies. During a boom, it is optimal to lean against the build-up of financial imbalances (deterioration in bank capital), especially as these grow. And once the economy is in a bust, it is optimal to clean, quite aggressively, to put the economy back on its feet as quickly as possible. In practice, the challenge is to determine when the bust phase is over and to move from cleaning to leaning in a timely manner. Moving too early would prolong the slump. Moving too late would jeopardise the sustainability of the recovery and make it more likely that the economy ends up in a bust again.

The welfare consequences are substantial. The right panel of Figure 4 shows the value functions $V(K, s)$ – the discounted sum of period payoffs under the optimal policy, corresponding to each regime. Welfare in the boom phase strictly dominates that in the bust phase, reflecting the fact that regimes are persistent – once a bust ensues, not only is current consumption low, it is also expected to be low for a protracted period. As bank capital increases and a boom phase edges closer, the discounted sum of the payoffs rises, so that the value function is upward sloping. The virtue of leaning is that it shores up bank capital and thus the continuation value, even as it sacrifices some short-term payoff.

4.3 Cycle dynamics and the equilibrium distribution

A key feature of the model is that it accommodates the possibility of recurrent financial cycles. To illustrate, we conduct dynamic simulations for our baseline calibration over 100 periods and under different random draws that determine the exact timing of regime switches. The top panels in Figure 5 show the evolution of bank capital K_t (top left panel) and output Y_t (top right panel) from two different simulations. Both K_t and Y_t cycle endogenously over time, as banks switch between spells of aggressive and conservative lending equilibria. The economy starts in a boom, with bank capital equal to 10. Aggressive lending leads to loan losses, initially reducing K_t. The fall in bank capital accelerates at lower values of K_t, as the central bank optimally stops leaning at this point, and the economy eventually enters a bust. In this phase, bank capital is rebuilt until K_t rises sufficiently and there is a recovery into a boom phase. The cycle then continues. There is a similar periodic swing in the level of output (top right panel), with high output in booms and low output in busts.

Due to stochastic regime switching, individual cycles vary, hence the difference between the two simulations. In the long run, the distribution of equilibrium variables converges to a stationary one –the model's analogue of a 'steady state'. This distribution can be approximated numerically (thanks to the ergodic theorem) by simulating the model and obtaining the empirical distribution. The histograms for output and bank capital are shown in the two lower panels of



Note: Dynamic simulation under optimal monetary policy, when $\beta = 0.9$. Histograms based on 100 rounds of simulations, each lasting 100 periods. The economy is initialised with $K = 10$ and a boom regime.

Figure 5: Dynamic simulation and equilibrium distribution

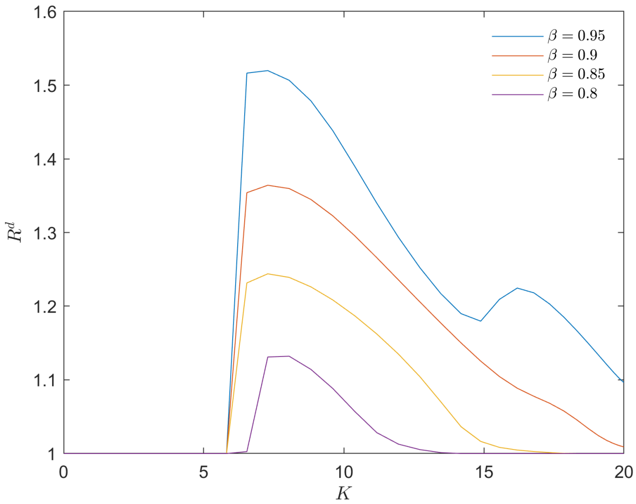
Figure 5, computed from 100 rounds of simulations each for 100 periods.

Because of the recurrent financial cycle, the stationary distribution for output is bi-modal. In a bust, the distribution of output is concentrated in a very narrow range, as it is hardly affected by policy. In a boom, the distribution is skewed, with the mode being lower than the mean. This reflects the leaning policy, as the central bank gives up some upside for output to curtail the financial boom. The equilibrium distribution for K_t is unimodal and shows a similar skew for the same reason – policy tends to favour higher capital and stabilise it during the boom.

A distinctive feature of the economy's evolution is path dependence (Figure 5). Past decisions influence the trade-offs policy faces today. Bygones are not bygones. This dependence reflects the path of bank capital and the sequence of boom-bust phases – the state variables in our model that describe the evolution of financial imbalances. The structural link between the booms and the subsequent busts generates a powerful and rather neglected monetary policy transmission channel. We next explore its implications.

5 The role of monetary policy frameworks

In our setup, the monetary policy framework acts as the anchor for the economy. Absent a unique 'natural interest rate' to which the economy gravitates, the long-run evolution of interest rates depends on the monetary regime. Moreover, such a regime determines the economy's proclivity to boom-bust cycles – a key property that influences the output path in the long-run.



Note: This figure shows the optimal level of policy interest rate R^d corresponding to various values of the central bank's discount factor β. All other parameters are kept unchanged.

Figure 6: Optimal policy under 3 central banks

What are the implications of different policy frameworks? We next derive two main results. First, the extent to which the central bank discounts future welfare has a bearing on how much it is willing to lean against the boom, which in turn has different long-term consequences for the real economy. Second, the degree of policy inertia has non-trivial implications, but these depend critically on whether the central bank takes future inertia into account when setting policy today.

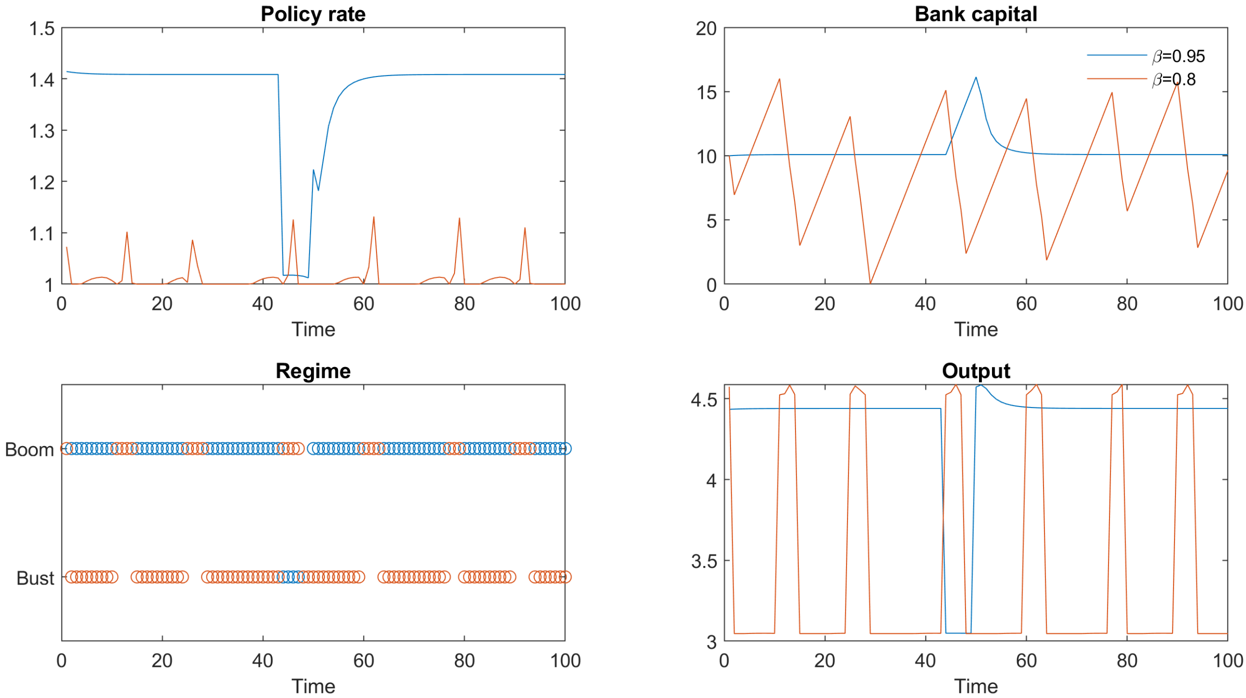
5.1 Central banks' discount factors

Given the intertemporal trade-off, the key element of the monetary policy framework is the relative weight assigned to the welfare of agents of different generations. We therefore compute optimal policy for central banks with various discount factors β. The higher the β, the more forward-looking – and hence more 'egalitarian' – the central bank is in assessing welfare. Conversely, as β falls, the greater is the relative weight placed on short-term outcomes.

In the baseline calibration above, $\beta = 0.9$. Here we explore the system's behaviour for $\beta = 0.95$, $\beta = 0.85$, and $\beta = 0.8$. Figure 6 plots the optimal policy corresponding to the four policymakers during the boom regime.[27] A lower discount factor leads to a weaker leaning policy for every K. The optimal timing is also different. Short-termist central banks start leaning late, preferring to wait until bank fragility becomes more evident (K is lower). They also give up leaning more quickly as fragility worsens, preferring to stimulate the economy even when the probability of a bust remains relatively low.

The implications of monetary policy frameworks for the evolution of the economy can be significant. Consider two simulation draws over 100 periods corresponding to central banks with $\beta = 0.95$ and $\beta = 0.8$, shown in Figure 7. Clearly, outcomes under the more myopic central bank with $\beta = 0.8$ are more volatile, with output fluctuating much more as the economy

[27] In a bust, all central banks set the interest rate at a similarly low level to maximise short-term payoff.



Note: Dynamic simulation under optimal monetary policy, when the central bank's discount rate is $\beta = 0.95$ and $\beta = 0.8$. Both simulations last for 100 rounds, and are based on the same sequence of random numbers used to determine regime switches (i.e. identical shocks). The economy is initialised with $K = 10$ and a boom regime.

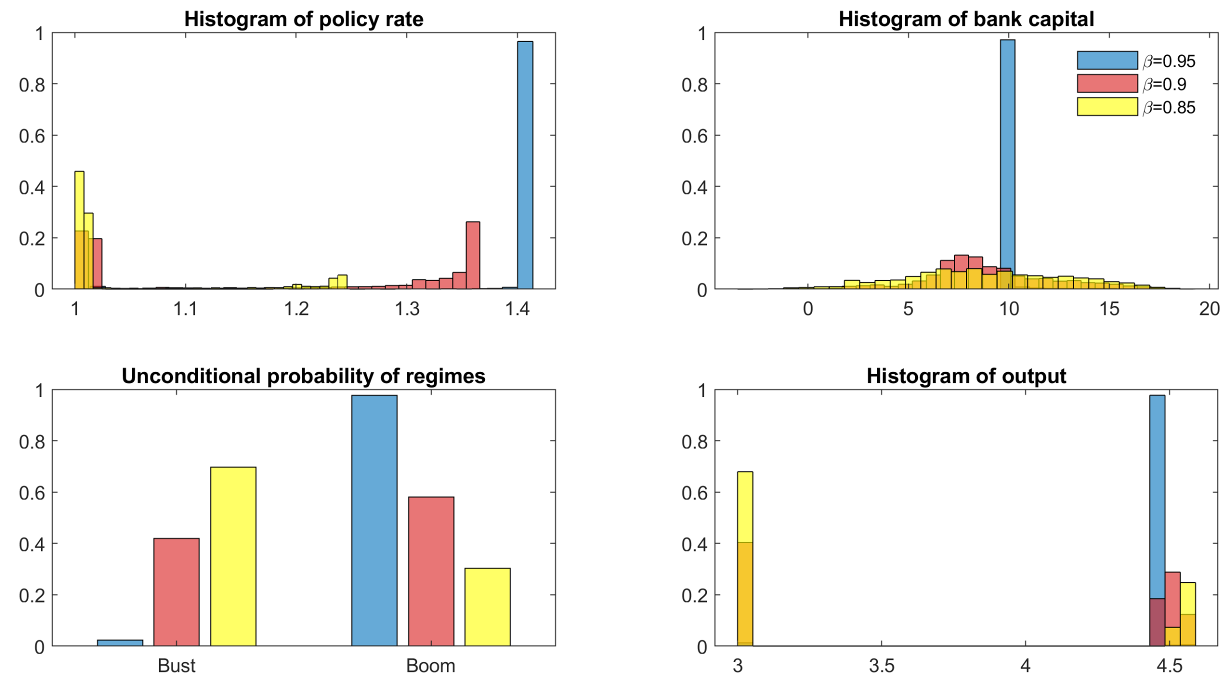
Figure 7: Simulated time-series under two monetary policy frameworks

alternates frequently between booms and busts. Policy leans very little, with interest rates set at low levels most of the time to maximise short-run payoffs. The more rapid decline in bank capital then triggers the bust, forcing the central bank to lower the policy rate again. Here, the interest rate is the same on average over successive cycles. In this sense, low interest rates beget low interest rates. We show later a case where it actually declines.

In contrast, a central bank with $\beta = 0.95$ places a greater weight on future outcomes and implements a higher interest rate for any given level of bank capital. In this case, monetary policy is actually able to completely stabilise bank capital and the economy remains in a boom for an extended period. As a result, the interest rate is higher on average, since busts are few and far between. In this particular simulation, it takes a large negative shock near the middle of the sample to trigger a bust. The bottom right-hand panel of Figure 7 confirms that the relative time spent in a bust is much shorter.

Another way to illustrate the importance of monetary policy frameworks is to examine long-term equilibrium outcomes. As before, we conduct 100 simulations, each lasting 100 periods, for the economy under the three different policy regimes. Figure 8 shows histograms serving as numerical proxies for the ergodic distribution.

The more forward-looking the central bank, the higher is the interest rate, on average (top-left panel). This results in a more stable financial system with bank capital generally higher and less volatile (top-right panel). The economy therefore spends less time in a bust (bottom-left panel) and output is generally high and stable, though slightly below the highest



Note: Histograms based on 100 rounds of simulations, each lasting 100 periods. The economy is initialised with $K = 10$ and a boom regime.

Figure 8: Equilibrium distribution under different monetary policy frameworks

level achieved in the other regimes given the short-run trade-off (bottom-right panel).

Result 5. *More forward-looking (egalitarian) central banks with a higher discount factor β lean more and earlier against the booms. As a consequence, they deliver on average more stable financial systems with more prolonged booms and higher real interest rates.*

5.2 Policy inertia

The analysis so far assumes that optimal monetary policy adjusts freely over time in either direction. As a result, policy rate adjustments are generally very sharp. In practice, central banks typically adjust policy gradually and in small steps. Coibion and Gorodnichenko (2012), for example, provide evidence of strong policy inertia and discuss possible rationales. And such interest rate smoothing is a standard feature in Taylor rule estimates, (e.g. Carlstrom and Fuerst (2014)). A key question, then, is whether policy inertia contributes to policy being 'behind the curve' and hence to raising the economy's proclivity to experiencing boom-bust cycles.

To examine this issue, we explore the implications of assuming that changing the policy rate involves adjustment costs. We conduct two exercises, which lead to very different outcomes. In one, the central bank attempts to track the first-best optimal policy but is forced to adjust policy only gradually in small steps. In this case, the boom-bust cycle becomes more virulent over time as policy is hindered in leaning against the upswing. As a result, interest rates drift downwards and the economy spends more and more time in the bust phase, despite no asymmetry in the gradualism constraint.

In the second exercise, the central bank is subject to quadratic adjustment costs but fully anticipates and internalises these costs, current and prospective, in setting policy today. In this case, the central bank leans even more strongly against the upswing than under the baseline. In doing so, it avoids busts and the associated policy adjustment costs down the road.

5.2.1 Constrained policy adjustments

Consider the first case. Here, the central bank cannot adjust the policy rate by more than a fixed amount δ each period. Let $R_t^{d^*}$ denote the optimal policy. Specifically, the central bank sets the interest rate according to

$$R_t^d = \begin{cases} R_t^{d^*}, & \text{if } R_t^{d^*} \in [R_{t-1}^d - \delta, R_{t-1}^d + \delta] \\ R_{t-1}^d + \delta, & \text{if } R_t^{d^*} > R_{t-1}^d + \delta \\ R_{t-1}^d - \delta, & \text{if } R_t^{d^*} < R_{t-1}^d - \delta \end{cases} \tag{5.1}$$

In other words, the central bank attempts to mimic the baseline optimal policy as much as possible, subject to changing its policy rate by at most δ per period.

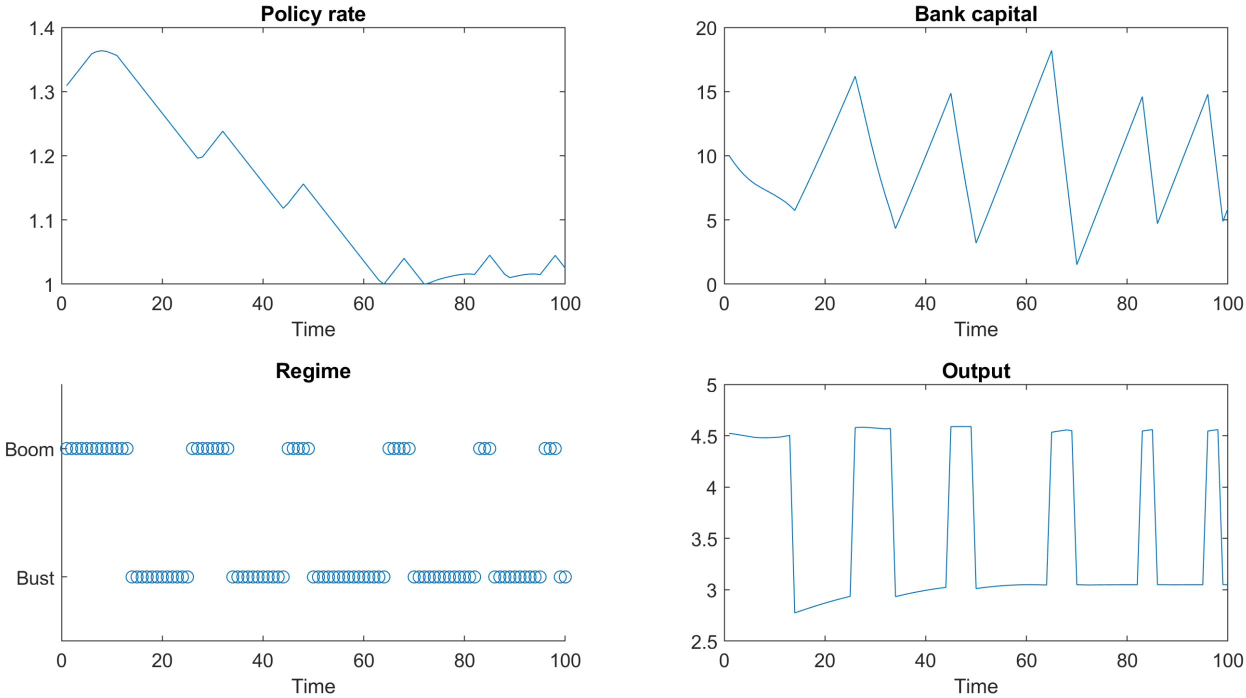
Figure 9 depicts the simulated time series of key variables. The economy starts with R_t^d close to the baseline optimal policy. Initially, the constrained central bank is able to lean against the financial upswing not very differently from the unconstrained one (top left panel). Once a bust ensues, the central bank lowers the interest rate, albeit more gradually and by less than in the baseline. As the economy recovers and enters a boom again, the central bank finds itself with a lower starting interest rate and a binding adjustment constraint which curtails the degree of leaning against the boom. Indeed, before the central bank can raise the policy rate enough to curb the boom or normalise the policy rate level, another bust occurs, which forces another round of rate cuts.

This ratcheting process continues to push the interest rate down over successive cycles, with busts becoming not only more frequent but also more drawn out. The constrained central bank in effect falls into a 'low interest rate trap', with an increasingly potent financial cycle and a downward trend in interest rate reinforcing each other, rendering an escape impossible. Here low rates beget *lower* rates.

Result 6. *Failure to address the booms over successive cycles can lead to more frequent busts and, as a result, a secular decline in the real interest rate.*

5.2.2 Optimal policy under adjustment costs

A qualitatively different result emerges in the second scenario, in which the central bank fully internalises the effects of adjustment costs associated with its future decisions when formulating policy today. It is convenient to consider this problem in the context of quadratic adjustment costs, which simplifies the computation of interior optimal policy. Specifically, the central bank's



Note: Dynamic simulation when the central bank attempts to implement an interest rate as close as possible to the baseline optimal policy, subject to a constraint that it cannot change the policy rate by more than $\delta = 0.01$ per period. The simulation lasts for 100 rounds, and are based on a sequence of random numbers used to determine regime switches. The economy is initialised with $K_0 = 10$, a boom regime and $R_0^d = 1.3$.

Figure 9: Simulation with constrained policy adjustments

period payoff is assumed to take the form

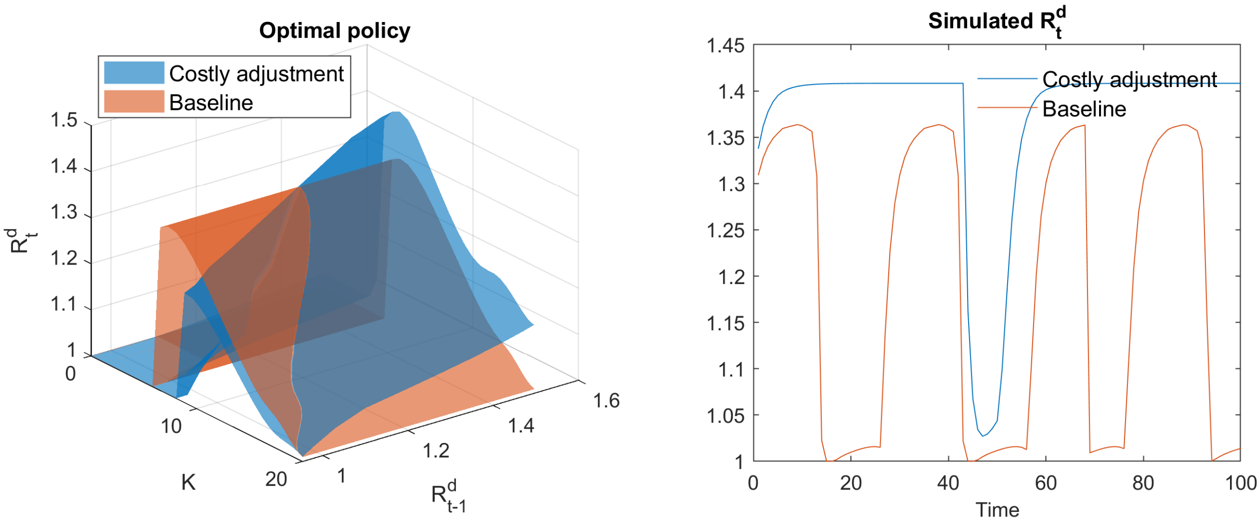
$$W^*(R_t^d, R_{t-1}^d, s_t, K_t) = W(R_t^d, s_t, K_t) - c_{I(R_t^d > R_{t-1}^d)}(R_t^d - R_{t-1}^d)^2 \tag{5.2}$$

where the adjustment cost $c_{I(R_t^d > R_{t-1}^d)}$ is equal to c_1 when the interest rate is increased and c_0 when it is cut. This formulation allows for an asymmetry in the adjustment cost if $c_0 \neq c_1$. For illustrative purposes, we assume that it is costlier to increase the interest rate than to cut it by an equal amount $(c_1 > c_0)$.[28]

The dynamic programming problem now involves an additional state variable, R_{t-1}^d, reflecting an additional form of history-dependence. The left panel of Figure 10 shows the optimal policy surface under adjustment costs as a function of bank capital K_t and the previous period interest rate level R_{t-1}^d. For comparison, the panel shows also the optimal policy from the baseline case without adjustment costs, which by construction does not vary with R_{t-1}^d. Qualitatively, the result appears quite intuitive. The optimal policy varies positively with the initial level of the policy rate, implying that monetary policy now involves some inertia.

Strikingly, this case strengthens the case for leaning against the wind. The reason is that acting pre-emptively helps the central bank avoid large policy adjustments associated with booms and busts. Optimal policy actually involves *more* leaning at the outset. This is illustrated in the right panel of Figure 10, which shows two simulations of R_t^d corresponding

[28]The case with symmetric adjustment costs yields qualitatively similar results.



Note: The left panel shows the optimal policy surfaces as a function of K_t and R^d_{t-1} during a boom, corresponding to costly policy adjustment case and the baseline. For costly adjustment, it is assumed that $c_1 = 10$ and $c_0 = 2$. The right panel shows two simulations under an identical draw of shock and initial conditions ($K_0 = 10$, a boom regime and $R^d_0 = 1.3$).

Figure 10: Optimal policy under costly policy adjustments

to the baseline and the adjustment cost case, both with identical initial conditions, parameters and shock sequence. Despite a penalty on interest rate increases, the central bank sets interest rate to a higher level right at the outset relative to the baseline. The intuition is that leaning earlier, and incurring adjustment costs immediately, reduces the risk of a future bust and the associated need for large and costly interest rate adjustments. Taken together, the two cases suggest that policy inertia *per se* is not sufficient for policymakers to fall behind the curve with respect to the financial cycle. Indeed, were policymakers to internalise adjustment costs, they would lean even more today to avoid situations where large interest adjustments are needed down the road.

Policy inertia is more problematic when it constrains large required policy actions in real time in ways not anticipated beforehand. This is arguably the more realistic case, as policymakers can hardly foresee how their hands may be tied in the future (e.g. how uncertainty may force them to act cautiously). The analysis points to systematic deviations from optimal policy that are not internalised as one cause of policy falling behind the curve, with the effects cumulating over time. The predictions of the model in this case, with interest rate normalisation being repeatedly stunted by financial dislocations, appear broadly consistent with developments over the last two decades.

5.3 Policy discussion

Our framework underlines the path-dependence of policy which lends a new perspective to several important policy debates. With financial fragility acting as a state variable that links developments today to outcomes in the distant future, policy options available today depend on the long sequence of previous policy decisions. This sequence of decisions does not just have a transient impact on the economy, but leaves a long-lasting imprint.

A case in point is the long-run evolution of real interest rates. Prevailing discussion typically attributes their evolution to shifts in the natural rate of interest, driven by real saving-investment factors such as preferences, productivity growth and demographics. Our framework points to the monetary policy regime as a possible factor, through its influence on the economy's proclivity to boom-bust cycles. In our setting, a policy rule that places a smaller weight on future outcomes, and hence leans less against the financial cycle, not only prescribes a lower interest rate in normal times, but also leads to a higher incidence of busts and hence lower interest rates in the future. As weakening balance sheets weigh on the economy over successive cycles, it becomes increasingly hard to raise rates without damaging the economy in the short term. This monetary hysteresis property echoes earlier works that emphasise the long-lasting financial strains left by crises such as those in the early 1990s and the GFC (e.g. Drehmann et al. (2012), Borio (2014b)).

The possibility of monetary hysteresis also highlights the nature of policy trade-off associated with leaning against the wind. If financial stresses arise only as a result of exogenous shocks, then leaning may be necessary only as a last resort once risks have become imminent and other policy tools exhausted. But if financial imbalances accumulate incrementally, continuously shaped by monetary policy in the process, then a failure to lean entails a cost that grows as time passes. The intertemporal nature of this policy trade-off becomes critical, and neglecting it leads to inferior macroeconomic outcomes.

Could macroprudential tools solve the problem set out in our model? While a precise answer depends on explicit modelling assumptions, which we leave for future research, we note a couple of general observations. First, given that there are no sectoral imbalances in our framework, a key benefit of macroprudential policy in dealing with specific segments of the economy would be absent. Second, conventional macropruential tools could not surgically remove the underlying friction at play here – that banks do not internalise how their interest rate settings collectively affect the aggregate quality of borrowers. Standard measures such as minimum capital requirements would effectively force banks to set higher interest rates, resulting in lower output. As such, macroprudential measures would face the same trade-off as monetary policy. That said, macroprudential tools would add an extra degree of freedom, enabling the policymaker to regulate both risk-taking (the lending rate) and household intertemporal consumption allocation (the deposit rate). But this would be second-order relative to the welfare consequences of boom-bust cycles. So long as the intertemporal trade-off could not be eliminated by any policy tool, the first-order welfare gain from leaning is likely to justify a combination of (mutually reinforcing) policies.

6 Conclusion

The debate on the appropriate role of monetary policy with respect to financial stability has been an active one, spurred once more by the GFC. Views necessarily differ depending on one's underlying model of the economy. We have argued that by assuming long-run money neutrality, prevailing frameworks neglect a key, and potentially important, transmission channel. By relaxing this assumption, the monetary policy regime becomes a significant factor in influencing longer-run economic outcomes, including the real interest rate. We are by no means arguing that it is the only factor – central banks set policy in response to economic developments and are hence influenced by a confluence of factors related to the structure of the economy. Rather, we simply offer a complementary explanation to that based on traditional saving-investment factors, whose empirical performance in explaining real interest rates has been rather weak (e.g. Borio et al. (2017)).

In drawing this conclusion, it is also important to distinguish between the specific model developed in this paper and broader considerations. The model is purely illustrative. It focuses on a specific source of instability – incentive problems in the supply of credit. In practice, there are other sources of instability, such as limitations in the measurement of risk owing to imperfect information and problems on the demand side, which can result in over-indebtedness and debt overhangs (Borio et al. (2001)). Moreover, as empirical evidence indicates, financial cycles may have persistent first-order effects on the misallocation of resources and hence productivity. Such considerations, ignored in our analysis, would strengthen our conclusions concerning the long-term impact of monetary policy.

Our goal in this paper was simply to shed light on a set of theoretical conditions that could justify the implementation of a more financial stability-oriented monetary policy. Of course, reaching a proper conclusion requires consideration of a broader set of issues, many of which are empirical and extend to the analysis of alternative instruments, such as prudential policies (e.g. Borio (2014a) and Borio et al. (2018)). The paper provides just one more piece of the puzzle.

References

Adrian, T. and F. Duarte (2017): "Financial Vulnerability and Monetary Policy," Staff Reports 804, Federal Reserve Bank of New York.

Adrian, T., F. Duarte, F. Grinberg, and T. Mancini-Griffoli (2018): "Monetary Policy and Financial Conditions: A Cross-Country Study," CEPR Discussion Papers 12681.

Adrian, T. and N. Liang (2016): "Monetary Policy, Financial Conditions, and Financial Stability," CEPR Discussion Papers 11394.

Aldasoro, I., C. Borio, and M. Drehmann (2018): "Early Warning Indicators of Banking Crises: Expanding the Family," *BIS Quarterly Review*, March, 29–45.

Altermatt, L. and C. Wipf (2020): "Liquidity, the Mundell-Tobin Effect and the Friedman Rule," Discussion papers, University of Bern.

Basel Committee on Banking Supervision (BCBS) (2016): "An Assessment of the Long-term Economic Impact of Stronger Capital and Liquidity Requirements," July.

Beaudry, P., D. Galizia, and F. Portier (2020): "Putting the Cycle Back into Business Cycle Analysis," *American Economic Review*, 110, 1–47.

Bebchuk, L. A. and I. Goldstein (2011): "Self-fulfilling Credit Market Freezes," *Review of Financial Studies*, 24, 3519–3555.

Benigno, G. and L. Fornaro (2018): "Stagnation Traps," *Review of Economic Studies*, 85(3), 1425–1470.

Benigno, P., G. Eggertsson, and F. Romei (2014): "Dynamic Debt Deleveraging and Optimal Monetary Policy," NBER Working Papers 20556.

Bernanke, B., M. Gertler, and S. Gilchrist (1999): "The Financial Accelerator in a Quantitative Business Cycle Framework," in *Handbook of Macroeconomics*, ed. by J. B. Taylor and M. Woodford, Elsevier, vol. 1C, chap. 21, 1341–1393.

Blanchard, O., E. Cerutti, and L. Summers (2015): "Inflation and Activity – Two Explorations and their Monetary Policy Implications," NBER Working Papers 21726.

Boissay, F., F. Collard, and F. Smets (2016): "Booms and Banking Crises," *Journal of Political Economy*, 124, 489–538.

Borio, C. (2014a): "Macroprudential Frameworks: (Too) Great Expectations?" *Central Banking Journal*, 29–45, 25th Anniversary issue.

——— (2014b): "Monetary Policy and Financial Stability: What Role in Prevention and Recovery?" BIS Working Papers 440.

Borio, C., P. Disyatat, J. Juselius, and P. Rungcharoenkitkul (2017): "Why so Low for so Long? A Long-Term View of Real Interest Rates," BIS Working Papers 685.

Borio, C., P. Disyatat, M. Juselius, and P. Rungcharoenkitkul (2018): "Monetary Policy in the Grip of a Pincer Movement," BIS Working Papers 706.

Borio, C. and M. Drehmann (2009): "Assessing the Risk of Banking Crises - Revisited," *BIS Quarterly Review*, March, 29–46.

Borio, C., C. Furfine, and P. Lowe (2001): "Procyclicality of the Financial System and Financial Stability: Issues and Policy Options," in *Marrying the Macro- and Micro-prudential Dimensions of Financial Stability*, Bank for International Settlements, vol. 1 of *BIS Papers chapters*, 1–57.

Borio, C., E. Kharroubi, C. Upper, and F. Zampolli (2016): "Labour Reallocation and Productivity Dynamics: Financial Causes, Real Consequences," BIS Working Papers 534.

Borio, C. and P. Lowe (2002): "Asset Prices, Financial and Monetary Stability: Exploring the Nexus," BIS Working Papers 114.

——— (2004): "Securing Sustainable Price Stability: Should Credit Come Back from the Wilderness?" BIS Working Papers 157.

Borio, C. and H. Zhu (2012): "Capital Regulation, Risk-Taking and Monetary Policy: A Missing Link in the Transmission Mechanism?" *Journal of Financial Stability*, 8, 236–251.

Brunnermeier, M. and Y. Sannikov (2014): "A Macroeconomic Model with a Financial Sector," *American Economic Review*, 104, 379–421.

Brunnermeier, M. K., T. M. Eisenbach, and Y. Sannikov (2013): *Macroeconomics with Financial Frictions: A Survey*, Cambridge University Press, vol. 2 of *Econometric Society Monographs*, 3–94.

Brunnermeier, M. K. and Y. Koby (2019): "The Reversal Interest Rate," Working paper.

Caballero, R. and A. Simsek (2018): "A Risk-Centric Model of Demand Recessions and Macroprudential Policy," BIS Working Papers 733.

Carlstrom, C. T. and T. S. Fuerst (2014): "Adding Double Inertia to Taylor Rules to Improve Accuracy," *Economic Commentary*, 2014-08.

Cecchetti, S. G. and E. Kharroubi (2015): "Why Does Financial Sector Growth Crowd Out Real Economic Growth?" BIS Working Papers 490.

Cerra, V. and S. C. Saxena (2008): "Growth Dynamics: The Myth of Economic Recovery," *American Economic Review*, 98, 439–457.

Claessens, S. and M. A. Kose (2017): "Macroeconomic Implications of Financial Imperfections: A Survey," BIS Working Papers 677.

Claessens, S., M. A. Kose, and M. E. Terrones (2012): "How Do Business and Financial cycles Interact?" *Journal of International Economics*, 87, 178–190.

Coibion, O. and Y. Gorodnichenko (2012): "Why Are Target Interest Rate Changes So Persistent?" *American Economic Journal: Macroeconomics*, 4, 126–162.

Coimbra, N. and H. Rey (2017): "Financial Cycles with Heterogeneous Intermediaries," NBER Working Papers 23245.

Cúrdia, V. and M. Woodford (2016): "Credit Frictions and Optimal Monetary Policy," *Journal of Monetary Economics*, 84, 30–65.

De Fiore, F. and O. Tristani (2011): "Credit and the Natural Rate of Interest," *Journal of Money, Credit and Banking*, 43, 407–440.

Dell'Ariccia, G., L. Laeven, and G. A. Suarez (2017): "Bank Leverage and Monetary Policy's Risk-Taking Channel: Evidence from the United States," *Journal of Finance*, 72, 613–654.

Disyatat, P. (2011): "The Bank Lending Channel Revisited," *Journal of Money, Credit and Banking*, 43, 711–734.

Drehmann, M., C. Borio, and K. Tsatsaronis (2012): "Characterising the Financial Cycle: Don't Lost Sight of the Medium Term!" *BIS Working Papers no. 380.*

Drehmann, M., M. Juselius, and A. Korinek (2017): "Accounting for Debt Service: The Painful Legacy of Credit Booms," BIS Working Papers 645.

Filardo, A. and P. Rungcharoenkitkul (2016): "A Quantitative Case for Leaning Against the Wind," BIS Working Papers 594.

Gabaix, X. (2019): "A Behavioral New Keynesian Model," Working Paper 22954, National Bureau of Economic Research.

Gaiotti, E. and A. Secchi (2006): "Is There a Cost Channel of Monetary Policy Transmission? An Investigation into the Pricing Behavior of 2,000 Firms," *Journal of Money, Credit and Banking*, 38, 2013–2037.

Garga, V. and S. R. Singh (2016): "Output Hysteresis and Optimal Monetary Policy," Mimeo.

Gertler, M. and N. Kiyotaki (2010): "Financial Intermediation and Credit Policy in Business Cycle Analysis," in *Handbook of Monetary Economics*, ed. by B. M. Friedman and M. Woodford, Elsevier, vol. 3 of *Handbook of Monetary Economics*, chap. 11, 547–599.

Gertler, M., N. Kiyotaki, and A. Prestipino (2017): "A Macroeconomic Model with Financial Panics," NBER Working Papers 24126.

Haubrich, J. G. (2015): "How Cyclical Is Bank Capital?" Working Papers (Old Series) 1504, Federal Reserve Bank of Cleveland.

Jakab, Z. and M. Kumhof (2018): "Banks Are Not Intermediaries of Loanable Funds – Facts, Theory and Evidence," Bank of England working papers 761, Bank of England.

Jordà, Ò., M. Schularick, and A. M. Taylor (2016): "The Great Mortgaging: Housing Finance, Crises and Business Cycles," *Economic Policy*, 31, 107–152.

Juselius, M., C. Borio, P. Disyatat, and M. Drehmann (2017): "Monetary Policy, the Financial Cycle, and Ultra-Low Interest Rates," *International Journal of Central Banking*, 13, 55–89.

Juselius, M. and M. Drehmann (2015): "Leverage Dynamics and the Real Burden of Debt," BIS Working Papers 501.

Kaplan, G., B. Moll, and G. L. Violante (2018): "Monetary Policy According to HANK," *American Economic Review*, 108, 697–743.

Karlan, D. S. and J. Zinman (2008): "Credit Elasticities in Less-Developed Economies: Implications for Microfinance," *American Economic Review*, 98, 1040–1068.

Kindleberger, C. P. (2000): *Manias, Panics and Crashes: A History of Financial Crises*, Palgrave Macmillan, 4 ed.

Kiyotaki, N. and J. Moore (2002): "Evil Is the Root of All Money," *American Economic Review*, 92, 62–66.

Krishnamurthy, A. and T. Muir (2017): "How Credit Cycles across a Financial Crisis," Research Papers repec:ecl:stabus:3579, Stanford University, Graduate School of Business.

Lagos, R. and G. Rocheteau (2008): "Money and Capital as Competing Media of Exchange," *Journal of Economic Theory*, 142, 247–58.

McLeay, M., A. Radia, and R. Thomas (2014): "Money Creation in the Modern Economy," *Bank of England Quarterly Bulletin*, 54, 14–27.

Mian, A. and A. Sufi (2015): *House of Debt*, University of Chicago Press.

Mian, A., A. Sufi, and E. Verner (2017): "Household Debt and Business Cycles Worldwide," *The Quarterly Journal of Economics*, 132, 1755–1817.

Minsky, H. P. (1982): *Can "IT" Happen Again? Essays on Instability & Finance*, Armonk N.Y.: M.E. Sharpe & Co.

——— (1986): *Stabilising an Unstable Economy*, Yale University Press.

Rajan, R. G. (1994): "Why Bank Credit Policies Fluctuate: A Theory and Some Evidence," *The Quarterly Journal of Economics*, 109, 399–441.

Ravenna, F. and C. Walsh (2006): "Optimal Monetary Policy with the Cost Channel," *Journal of Monetary Economics*, 53, 199–216.

Reifschneider, D., W. Wascher, and D. Wilcox (2015): "Aggregate Supply in the United States: Recent Developments and Implications for the Conduct of Monetary Policy," *IMF Economic Review*, 63, 71–109.

Reinhart, C. M. and K. S. Rogoff (2009): *This Time Is Different: Eight Centuries of Financial Folly*, Princeton, New Jersey: Princeton University Press.

Samuelson, P. A. (1958): "An Exact Consumption-Loan Model of Interest With or Without the Social Contrivance of Money," *Journal of Political Economy*, 66, 467–82.

Schularick, M. and A. Taylor (2012): "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008," *American Economic Review*, 102, 1029–61.

Sheedy, K. (2018): "Taking Away the Punch Bowl: Monetary Policy and Financial Instability," mimeo.

Tobin, J. (1965): "Money and Economic Growth," *Econometrica*, 33, 671–84.

Vines, D. and S. Wills (2018): "The Rebuilding Macroeconomic Theory Project: an Analytical Assessment," *Oxford Review of Economic Policy*, 34, 1–42.

Woodford, M. (2018): "Monetary Policy Analysis when Planning Horizons are Finite," NBER Working Papers 24692.

Wright, R. (2018): "On the Future of Macroeconomics: a New Monetarist Perspective," *Oxford Review of Economic Policy*, 34, 107–131.

Appendix A Balance sheet mechanics

This appendix details the book-keeping of all agents in the simple financing model. Figure 11 shows entries in a given period, assuming a steady state. We consolidate old and young firms' balance sheets for brevity.

Figure 11: Balance sheet entries

Banks

	Asset	Liability
①	RWL	$R^d(WL-C_1)$ $NII=0$ C_1
②	WL	WL
③	$-R^d(WL-C_1)$ $-C_1$	$-R^d(WL-C_1)$ $-C_1$

Firms (consolidated)

	Asset	Liability
①	AL^α	RWL π^F
②	WL	WL
③	$-AL^\alpha$	$-R^d(WL-C_1)$ $-C_1$ $-\pi^F$

Young Households

	Asset	Liability
①	0	0
②	WL	WL
③	$-C_1$	$-C_1$

Old Households

	Asset	Liability
①	$R^d(WL-C_1)$ π^F	$R^d(WL-C_1)$ π^F
②	0	0
③	$-R^d(WL-C_1)$ $-\pi^F$	$-R^d(WL-C_1)$ $-\pi^F$

Note: This figure traces steady-state balance sheet entries of all agents in a given period. The first-row entries, marked by ①, are initial balances at the beginning of each period. Entries ② correspond to new bank lending and creation of new bank deposits. Entries ③ document the book entries after the goods market opens, relating to the payments of goods by households to firms and settlement of bank loans by firms to banks. As the entries represent steady-state transactions, the aggregate positions net out to equal the initial conditions in row ① after taking interest rates into account.

The initial positions at the beginning of the period are marked by ①. Banks' assets are loans extended last period plus interest, RWL, which also appear as firms' liabilities. Banks owe deposit plus interest to old households worth $R^d(WL-C_1)$ (appearing in old households' books as equity), where the principal is $WL-C_1$ because households withdrew C_1 last period for immediate consumption. These outstanding claims and liabilities earn banks net interest income of NII, which is zero in equilibrium. Recall from equation 2.9 that NII is defined as $(R-1)WL-(R^d-1)(WL-C_1) = RWL-R^d(WL-C_1)-C_1$. Equating it to zero implies $RWL = R^d(WL-C_1)+C_1$, namely that banks have a residual liability term of C_1 as of stage ①. This represents bank equity which will always be carried forward. It arises because, in any period, part of newly created deposit is used to extinguish old loans, so that bank assets exceed liabilities.[29] Finally, firms have output AL^α coming due as their assets, and owe dividend π^F

[29] This can also be viewed as a consequence of the timing feature of overlapping generations setups. In these models, there is always a question of how the young households obtain the goods at the very first date when there is not yet any production. One solution is to assume that banks supply these initially out of their

to old households.

In step ②, banks grant new loans to firms worth WL, creating new financial claims in the economy. Banks' assets represent 1-period loans to firms, while liabilities represent obligations to depositors. Young firms transfer this deposit to young households for labour services. Accrual accounting convention means interest rates are not included until next period, while young firms' assets are booked at cost WL for the moment as production only finishes next period. Young households now have WL deposits, representing their equity.

In the final step ③, the goods market opens and loan repayments take place. Old households transfer the remaining deposits of $R^d(WL-C_1)$ to firms in exchange for goods, while young households similarly transfer C_1. Households consume these goods (plus dividend π^F in the case of old households). Firms receive the deposit transfers totalling $R^d(WL-C_1)+C_1$, just enough to repay maturing bank loans of RWL (again using the fact that $R^d(WL-C_1)+C_1 = RWL$ when $NII=0$). Firms distribute the output AL^α to households corresponding to deposit transfers, and rebate the remaining π^F to old households as dividend. Entries at step ③ exactly cancel the initial positions at step ①. For banks, all outstanding loans in ① are settled by firms' repayments in ③ leaving only new loans of WL, while deposit liability nets to $WL-C_1$. All agents' balance sheets contract at this final stage, as a round of financing runs its course.

The net balance sheet positions, when shifted forward one period, are exactly the initial positions at step ①, consistent with the steady-state situation. Banks' assets, once including interests, become RWL. For liability, banks now owe $R^d(WL-C_1)$ of deposits to households, earn NII and carry over equity of C_1. For firms, assets become output worth AL^α with profit π^F booked as the difference over liability plus interest of RWL. Previously young households become old, and their net positions plus interest becomes $R^d(WL-C_1)$ plus the firm dividend.

endowment, which is the source of capital C_1.

Appendix B Financing with fiat money

This appendix examines an economy without banks, and where the role of financing is played by fiat money. Assume that the central bank can issue cash that is accepted by households as a trusted form of payments for labour. The central bank endows any newly printed cash to young firms, who use it together with any existing cash to finance production. When old, firms have incentives to exchange goods for cash, as it is the only means to start a new production round (we assume a bequest motive if generations of firms are different identities). Fiat money is thus valued by all in equilibrium and can function as both the medium of exchange and a store of value.

Let the face value of cash be the nominal unit of account, with P_t denoting the goods price at time t, and $\mathcal{W}_t$ the nominal wage rate (the real wage is $W_t \equiv \mathcal{W}_t/P_t$). The household problem is then given by

$$\max_{C_{1t},C_{2,t+1},L_t} \frac{C_{1t}^{1-\sigma}}{1-\sigma} + \rho\frac{C_{2,t+1}^{1-\sigma}}{1-\sigma} - \frac{\Phi}{1+\chi}L_t^{1+\chi} \tag{B.1}$$

$$\text{subject to} \qquad \mathcal{W}_t L_t + P_{t+1}\pi_{t+1} \geq P_t C_{1t} + P_{t+1}C_{2,t+1} \tag{B.2}$$

where the notation follows the original model. The budget constraint specifies nominal incomes on the left hand side, and the nominal expenditure on the right. Firms choose L_t to maximise profit

$$\pi_{t+1} = AL_t^{\alpha} - \frac{\mathcal{W}_t L_t}{P_{t+1}} \tag{B.3}$$

The cost is the amount of goods the firms will have to give up in the second period to reacquire the same amount of cash needed to hire L_t. We close the model in the traditional way by specifying a money supply growth chosen by the central bank

$$\frac{M_{t+1}}{M_t} = 1+m \tag{B.4}$$

Solving the model produces the first-order conditions

$$\rho\frac{P_t}{P_{t+1}} = \left(\frac{C_{2,t+1}}{C_{1t}}\right)^{\sigma} \tag{B.5}$$

$$\Phi L_t^{\chi} = \frac{W_t}{C_{1t}^{\sigma}} \tag{B.6}$$

$$\alpha A L_t^{\alpha-1} = W_t\frac{P_t}{P_{t+1}} \tag{B.7}$$

Note that the (inverse) inflation rate P_t/P_{t+1} is the real rate of return to holding cash, the economy's only monetary asset. To determine its equilibrium value, consider the moment when goods market opens in period t (focusing on the moment of financing gives rise to the same condition). Let x_t denote the fraction of nominal wage income $\mathcal{W}_t L_t = M_t$ that young households born in period t would like to spend immediately. Then the money market equilibrium is given by

$$P_t A L_{t-1}^{\alpha} = x_t M_t + (1 - x_{t-1}) M_{t-1} \tag{B.8}$$

where the left hand side is the demand for money coming from old firms who seek to sell goods and obtain cash, and the right hand terms aggregate the supply of cash coming from young and old households who seek consumption goods. In the steady state of constant L and x, this condition simplifies to

$$\frac{P_{t+1}}{P_t} = \frac{M_{t+1}}{M_t} = 1 + m \tag{B.9}$$

As in standard OLG models, money supply growth dictates the inflation rate and hence the interest rate on cash.

The first-order conditions B.5-B.7 are then all but isomorphic to those of the original model except the latter features two interest rates R_t^d and R_t on deposits and loans. Denoting the rate of return to cash by $R_t^c \equiv P_t/P_{t+1}$, solving for the steady-state equilibrium gives (suppressiong time subscript)

$$\Phi L^{1+\chi-\alpha(1-\sigma)} = \alpha A^{1-\sigma} \left[1 + \rho^{\frac{1}{\sigma}} R^{c\,\frac{1}{\sigma}-1}\right]^{\sigma} \tag{B.10}$$

$$C_1 = \frac{A L^{\alpha}}{R^c + (\rho R^c)^{\frac{1}{\sigma}}} \tag{B.11}$$

$$C_2 = A L^{\alpha} \left[\frac{(\rho R^c)^{\frac{1}{\sigma}}}{R^c + (\rho R^c)^{\frac{1}{\sigma}}}\right] \tag{B.12}$$

Result 2 of the bank financing model then extends to the current setting, except for the implementation of monetary policy. The social planner outcome is attained if $R^c = 1$, which the central bank can achieve by keeping money stock constant and set $m = 0$.

Appendix C Banks' market shares

This appendix provides a microfoundation for the bank-firm matching assumption under imperfect lending competition, deriving banks' market shares in equations 3.8 and 3.9. We present two complementary approaches to motivate market power that allows each bank to charge a different interest rate from its competitors.

C.1 Preferences for diversity

The first approach follows closely the modelling of pricing power in the New Keynesian literature and is grounded on firms' preferences for bank diversity. Let a representative firm have a CES preference over services from different banks, reflecting firms' desire for a diverse set of relationships with banks. Consider first the case of productive firms. Let M_{it} denote the amount of loans a firm obtains from bank i, and R_{it} the corresponding lending interest rate. Total loan $L_t W_t$ and average interest rate R_t are aggregate indices of the bank-level variables:

$$L_t W_t \equiv \left(\int_0^1 M_{it}^{\frac{\theta_1 - 1}{\theta_1}} di \right)^{\frac{\theta_1}{\theta_1 - 1}} \tag{C.1}$$

$$R_t \equiv \left(\int_0^1 R_{it}^{1-\theta_1} di \right)^{\frac{1}{1-\theta_1}} \tag{C.2}$$

where $\theta_1 > 1$ represents the elasticity of substitution, and the number of banks is normalised to one (hence the integral limits). The total cost of borrowing for a given firm is

$$\bar{R} = \int_0^1 R_{it} M_{it} di \tag{C.3}$$

Firms decide how much to borrow from each bank by maximising equation C.1 subject to C.3. Solving this problem gives rise to the standard demand equation for bank i's loan:

$$M_{it} = \left(\frac{R_t}{R_{it}} \right)^{\theta_1} L_t W_t \tag{C.4}$$

which depends on the relative price R_t / R_{it} and the degree of bank market power θ_1. This establishes the first part of bank profit function in equation 3.8. Note that the restriction $\theta_1 > 1$ ensures that the interest rate markup, $\theta_1/(\theta_1 - 1)$ in equation 3.11, is sensibly positive.

Identical derivation applies in the case of unproductive firms, except the aggregate loan is given by $\bar{N}/R_t^{\omega}$, and $\theta = \theta_2$. In this case, we assume in addition that unproductive firms value bank diversity less when the average lending rate R_t is higher – in other words, when credit is scarce, they care more about obtaining cheapest possible loans than having an access to a diverse group of banks. For simplicity, the diversity preference parameter θ_2 is set at a constant θ_H for higher values of R_t, and θ_L otherwise, where $\theta_H > \theta_L > 1$.

C.2 Preferred habitat

The second approach motivates market power directly as arising from firms' preference to be matched with their chosen banks. Assume that each firm is initially assigned to its first-choice bank i, but has an option to switch to a second-choice bank j – call this firm (i, j). When firm (i, j) decides to move to bank j, it incurs a switching cost $c \in [0, \infty)$ (in utility unit) and receives a utility $v(R_j/R_i) \geq 0$ that is strictly positive if bank j offers a cheaper loan, with $v(1) = 0$ and $v' < 0$. Firm (i, j) will thus leave bank i for bank j if and only if

$$v\left(\frac{R_j}{R_i}\right) > c \tag{C.5}$$

This assumption endows bank i with market power, as firm (i, j) will stay with bank i even if it can obtain a lower rate from bank j, provided the rate differential is not too large. We assume that for each bank i there is a continuum of firms (i, j), where j is uniformly distributed across all non-i banks and the switching costs c are heterogeneous with density function $h(c)$.

With these assumptions, one can derive the size of market share that bank i will capture as a function of its lending rate R_i and the market rate R. For a market lending rate R and an aggregate loan amount LW (or $\bar{N}/R^\omega$ for unproductive firms), loan demand facing bank i is $s_i LW$ (or $s_i \bar{N}/R^\omega$), where market share s_i is given by

$$s_i\left(\frac{R}{R_i}, h\right) = \begin{cases} 1 - \int_{\forall j} \int_0^{v\left(\frac{R_j}{R_i}\right)} h(c)\, dc\, dj & = 1 - \int_0^{v\left(\frac{R}{R_i}\right)} h(c)\, dc \quad \text{for } R_i > R \\ 1 + \int_{\forall j} \int_0^{v\left(\frac{R_i}{R_j}\right)} h(c)\, dc\, dj & = 1 + \int_0^{v\left(\frac{R_i}{R}\right)} h(c)\, dc \quad \text{for } R_i < R \end{cases} \tag{C.6}$$

The second equalities are obtained from integrating over evenly distributed j using $R_j = R$. Note the symmetry of s_i as a function of R/R_i around $R/R_i = 1$. For a small deviation of R_i around R, we can then approximate the market share function s_i for any given h by

$$s_i\left(\frac{R}{R_i}, h\right) \approx \left(\frac{R}{R_i}\right)^{\theta(h)} \tag{C.7}$$

by choosing an appropriate $\theta > 0$. To motivate the two regimes $\theta_2 \in \theta_L, \theta_H$, unproductive firms are assumed to have their utility $v()$ or switching cost density $h(c)$ being regime specific and dependent on R, such that θ_2 is a step function of R.

Note that this modelling approach permits any $\theta \geq 0$, and is more general than the first approach which requires $\theta > 1$. This provides a foundation for the case of $\theta_2 = \theta_L = 0$ considered in the main text. The case corresponds to $h(c)$ being non-zero only at very high c, i.e. the switching cost is prohibitively high such that all firms stay with their first-choice banks.

Appendix D Proofs

Result 3. *When $\theta_L = 0$ and $R_t^d = 1$, sufficient conditions for a financial cycle to exist are*

$$\theta_H > \theta_1 - 1 \tag{3.23}$$

$$\kappa > 0 \tag{3.24}$$

$$\theta_1 > \theta_1^* \tag{3.25}$$

where $\kappa \equiv \omega - \frac{1+\chi}{\chi+\sigma+(1-\alpha)(1-\sigma)}$ *and* θ_1^* *solves*

$$\frac{(\theta_1^* - 1)^{\kappa+1}}{\theta_1^{*\kappa}} = \frac{1}{\bar{N}} \left[(\alpha A)^{1+\chi} \left(\frac{\psi_1^\sigma}{\Phi} \right)^\alpha \right]^{\frac{1}{\chi+\sigma+(1-\alpha)(1-\sigma)}}$$

Proof. With a constant policy rate, all endogenous variables are constant in any given regime, and hence we drop the time subscript, with an understanding that L, W and R depend on the prevailing regime. The existence of a financial cycle at $R^d = 1$ requires

$$\Pi(R_H, R_H) > 0 > \Pi(R_L, R_L)$$

At $R^d = 1$, we have

$$\psi_1 = \alpha(1 + \rho^{\frac{1}{\sigma}})$$
$$\psi_2 = 1$$

implying bank profit and lending rate of

$$\Pi(R, R) = (R-1)LW - \frac{\bar{N}}{R^\omega}$$
$$R = \frac{\theta_1}{\theta_1 - 1} + \frac{\theta_2 \bar{N}}{(\theta_1 - 1)LWR^\omega}$$

Bank profit $\Pi(R, R)$ is positive when

$$(R-1)LWR^\omega > \bar{N}$$

Substituting for R, the left-hand side can be expanded to

$$\begin{aligned}(R-1)LWR^\omega &= \left(\frac{\theta_1}{\theta_1 - 1} - 1 + \frac{\theta_2 \bar{N}}{(\theta_1 - 1)LWR^\omega} \right) LWR^\omega \\ &= \left(\frac{1}{\theta_1 - 1} \right) \left(LWR^\omega + \theta_2 \bar{N} \right)\end{aligned}$$

so that bank profit is positive when

$$LWR^\omega > (\theta_1 - 1 - \theta_2)\bar{N} \tag{D.1}$$

and negative when inequality is reversed.

In the bust regime, $\theta_2 = \theta_H > \theta_1 - 1$ is sufficient to ensure D.1, thus proving the first inequality 3.23. In the boom regime where $\theta_2 = \theta_L = 0$ and $R = \theta_1/(\theta_1 - 1)$, a sufficient condition for the reverse of D.1 is

$$LW < (\theta_1 - 1)\bar{N}/R^{\omega} \tag{D.2}$$

We solve for LW under the boom by equating the labour demand and supply functions (equations 3.20 and 3.21), yielding

$$LW = \left[\left(\frac{\alpha A}{R}\right)^{1+\chi}\left(\frac{\psi_1^{\sigma}}{\Phi}\right)^{\alpha}\right]^{\frac{1}{\chi+\sigma+(1-\alpha)(1-\sigma)}}$$

Substituting this into D.2 and using $R = \theta_1/(\theta_1 - 1)$, we get

$$\frac{(\theta_1 - 1)^{\kappa+1}}{\theta_1^{\kappa}} > \frac{1}{\bar{N}}\left[(\alpha A)^{1+\chi}\left(\frac{\psi_1^{\sigma}}{\Phi}\right)^{\alpha}\right]^{\frac{1}{\chi+\sigma+(1-\alpha)(1-\sigma)}} \tag{D.3}$$

where

$$\kappa \equiv \omega - \frac{1+\chi}{\chi+\sigma+(1-\alpha)(1-\sigma)} \tag{D.4}$$

Differentiating the left-hand side of equation D.3 gives

$$\frac{\partial}{\partial\theta_1}\left[\frac{(\theta_1 - 1)^{\kappa+1}}{\theta_1^{\kappa}}\right] = \left[\frac{\theta_1 - 1}{\theta_1}\right]^{\kappa}\left(\kappa + \frac{1}{\theta_1}\right)$$

This is positive when $\kappa > -1/\theta_1$, or in the limit of large θ_1, when $\kappa > 0$. When satisfied, the left hand side of D.3 is both strictly increasing and continuous, so that there exists $\theta_1^* > 1$ such that $(\theta_1^* - 1)^{\kappa+1}/\theta_1^{*\kappa}$ is exactly equal to the right hand side of D.3. Together, $\kappa > 0$ and $\theta_1 > \theta_1^*$ imply that D.3 holds. Thus conditions 3.24 and 3.25 ensure that bank profit is indeed negative during the boom regime. This completes the proof. □

Previous volumes in this series

816 October 2019	The reaction function channel of monetary policy and the financial cycle	Andrew Filardo, Paul Hubert and Phurichai Rungcharoenkitkul
815 September 2019	Fragmentation in global financial markets: good or bad for financial stability?	Stijn Claessens
814 September 2019	Interest rate spillovers from the United States: expectations, term premia and macro-financial vulnerabilities	Aaron Mehrotra, Richhild Moessner and Chang Shu
813 September 2019	Modelling yields at the lower bound through regime shifts	Peter Hördahl and Oreste Tristani
812 September 2019	Steady-state growth	Emanuel Kohlscheen and Jouchi Nakajima
811 September 2019	Embedded supervision: how to build regulation into blockchain finance	Raphael Auer
810 September 2019	Spillovers of funding dry-ups	Iñaki Aldasoro, Florian Balke, Andreas Barth and Egemen Eren
809 September 2019	Inflation expectations anchoring: new insights from micro evidence of a survey at high-frequency and of distributions	Nikos Apokoritis, Gabriele Galati, Richhild Moessner and Federica Teppa
808 August 2019	A disaster under-(re)insurance puzzle: Home bias in disaster risk-bearing	Hiro Ito and Robert N McCauley
807 August 2019	Bank intermediation activity in a low interest rate environment	Michael Brei, Claudio Borio and Leonardo Gambacorta
806 August 2019	Geographic spread of currency trading: the renminbi and other EM currencies	Yin Wong Cheung, Robert N McCauley and Chang Shu
805 August 2019	Exchange rate puzzles: evidence from rigidly fixed nominal exchange rate systems	Charles Engel and Feng Zhu
804 August 2019	(Un)conventional Policy and the Effective Lower Bound	Fiorella De Fiore and Oreste Tristani
803 July 2019	Determinants of credit growth and the bank lending channel in Peru: a loan level analysis	José Bustamante, Walter Cuba and Rafael Nivin
802 July 2019	A loan-level analysis of bank lending in Mexico	Carlos Cantú, Roberto Lobato, Calixto López and Fabrizio López-Gallo

All volumes are available on our website www.bis.org.