Review of Economic Dynamics 54 (2024) 101236

Contents lists available at ScienceDirect

Review of Economic Dynamics

journal homepage: www.elsevier.com/locate/red

Full Length Article

Coordinating in financial crises ☆

Caio Machado

Instituto de Economía, Pontificia Universidad Católica de Chile, Santiago, Chile

ARTICLE INFO

JEL classification:
D84
E32
G10

Keywords:
Coordination failures
Financial frictions
Financial crises

ABSTRACT

Why do some financial crises lead to macroeconomic disasters, while others barely affect the real economy? This paper proposes a model to study unusually deep financial crises. Deep crises arise from the interplay of demand-driven coordination failures on the productive sector and weak banks' balance sheets. There is a dynamic feedback between banks' balance sheets and coordination. Coordination failures happen when banks suffer large losses and substantially reduce asset prices and welfare, even if the economy is in good times and they rarely happen. Financial crises that start from similar initial shocks can feature very heterogeneous real effects.

1. Introduction

Some financial crises have a small effect on economic activity, while others are followed by economic disasters, with sudden and persistent collapses in output. Two years after the start of a financial crisis, output is on average 6.5% lower than what it would have been if the economy had grown according to its average growth rate.[1] However, the shortfall in output is very heterogeneous across episodes: For about 20% of the financial crises, this two-year output loss is above 12%, while this number is below 0.5% also for approximately 20% of the episodes. The standard deviation of two-year output losses is 51% above the average loss. Financial crises are rare events that, by definition, feature a lot of financial distress, but are not necessarily followed by large and persistent downswings in economic activity.

Motivated by those facts, in this paper I propose a model in which financial crises that start off with similar shocks can feature very heterogeneous real effects. To do so, I augment an otherwise standard model of financial crises with a novel two-way feedback between financial frictions and demand-driven coordination failures. It is shown that large shocks that depress bank lending can lead to a self-reinforcing coordination trap where firms scale down production, in which case the path of the economy becomes very dependent on whether it can leave such state soon enough. Through the lens of the proposed model, I then study how deep financial

☆ I am very grateful to Bernardo Guimaraes and Nobuhiro Kiyotaki for fruitful discussions and invaluable advice. I also thank anonymous referees, Luis Araujo, Braz Camargo, Daniel Carroll, Andre Diniz, Elton Dusha, Alexandre Janiak, Remzi Kaygusuz, Marco Lyrio, Ana Elisa Pereira, Joao Paulo Pessoa, Marcel Ribeiro, Hikaru Saijo, Klaus Schmidt-Hebbel, and seminar participants at the Central Bank of Chile, Cleveland Fed, Insper, ITAM, LAMES, LubraMacro, PUC-Chile, Midwest Macro Meeting, NYU Abu Dhabi, RIDGE Workshop of Financial Stability, Sao Paulo School of Economics, SBE Meeting, Tilburg University, UC Santa Cruz, University of Chile, UFPE and University of Sao Paulo. In addition, I am grateful to Matthew Baron, Emil Verner and Wei Xiong for sharing data used in this paper. I gratefully acknowledge financial support from the São Paulo Research Foundation (Fapesp) through grants 2012/23222-6 and 2013/22873-6, and from ANID/Conicyt through grants Fondecyt Iniciación 11180046 and Fondecyt Regular 1220615. A previous version of this paper circulated under the title "Financial Crises, Coordination Failures and Disasters".

E-mail address: caio.machado@uc.cl.

[1] Those numbers use the definition of financial crises of Baron et al. (2021). See Online Appendix C.1 for further details on the heterogeneous real effects of financial crises.

https://doi.org/10.1016/j.red.2024.101236
Received 15 May 2023; Received in revised form 10 July 2024

Available online 15 July 2024

crises (those leading to extreme output losses) affect the economy not only during those episodes, but also in good times, when those events are relatively unlikely.

The environment consists of a standard stochastic growth model, with two additional ingredients that are at the heart of the proposed mechanism: financial frictions and coordination frictions. Financial frictions force financial intermediaries (banks) to invest in risky assets while funding themselves through risk-free debt. That implies that they use leverage in equilibrium, making their balance sheets exposed to aggregate risk. Coordination frictions directly affect the productive sector. Due to demand externalities, a firm's decision to increase production depends on its expectation of aggregate demand, which in turn depends on the production of other firms. Firms are subject to fixed costs, and may want to scale up production (increase capacity utilization in the model) only if they believe others will do the same. This opens up the possibility of the economy getting trapped in a self-fulfilling regime with low economic activity and low aggregate demand.

How do financial and coordination frictions interact? Banks have expertise in intermediating capital, and due to their superior screening technology they can avoid lending capital to unproductive firms that create no aggregate demand. Thus, when banks have strong balance sheets and intermediate a lot of capital, productive firms are confident that aggregate demand will be sufficiently high, increasing their incentives to scale up production. Firms then manage to coordinate on a good equilibrium where output is high and firms operate at full capacity. On the other hand, when banks cut down intermediation, households directly lend capital to the real sector. Due to their inferior screening technology, they grant loans to unproductive firms (which simply default), reducing expectations about aggregate demand and possibly leading to a situation where all firms operate below full capacity. Hence, in equilibrium, demand-driven coordination failures arise as a result of insufficient bank lending.

Importantly, coordination failures also affect the financial sector, creating an endogenous two-way feedback between banks' balance sheets and firms' decisions. When the economy enters a coordination trap, the marginal productivity of capital falls (as firms operate with idle capacity), reducing the equilibrium rental rate of capital. Banks' capital returns are more affected than households', since the fraction of households' capital that is lent to unproductive firms is not affected by the coordination failure (as those firms would default anyway). This implies that the allocation of capital in the economy is re-balanced towards households as a coordination failure approaches. Fearing a coordination failure is possible in a near future, banks start to unload capital, which is precisely what causes the coordination trap in the first place. Coordination failures among firms then reinforce what causes them, the insufficient level of bank lending, and banks' precautionary behavior precipitates a coordination failure.

Two types of financial crises emerge from the model, depending on how much banks' balance sheets are affected during a crisis. The first type is interpreted as "mild" financial crises. Banks' balance sheets are mildly weak and some of the capital in the economy does not find its first-best use, but firms are still able to coordinate on the high regime. The second kind, labeled "deep" financial crises, is associated to larger losses in the banking sector and triggers demand-driven coordination failures. Banks strongly reduce their investments in risky assets, capital is poorly allocated, firms have low demand expectations and operate below its full potential.

In order to investigate how deep financial crises—understood as demand-driven coordination failures caused by insufficient bank lending—affect the economy not only during those episodes but also in good times, the model is solved globally, building on techniques similar to those in Brunnermeier and Sannikov (2014) and Phelan (2016). On a technical note, since firms' capacity decisions are subject to multiple equilibria, I use the global games approach of Morris and Shin (2001) to select equilibrium among firms. Differently from standard global games settings, here the fundamental that determines equilibrium outcomes is endogenous and depends on banks' decisions. As explained in the text, banks are also subject to self-fulling prophecies, and for some regions of the state space there may still be multiple levels of bank intermediation consistent with equilibrium. I numerically verify that, for reasonable parameters, this multiplicity is not quantitatively relevant.

The model is then calibrated for the US economy, as to match usual unconditional moments, as well as the response of key macroeconomic aggregates during the Great Recession of 2008. By changing some parameters of the model, one can turn off the coordination friction, turning the model into an extension of Brunnermeier and Sannikov (2014) augmented with monopolistic competition and labor. The calibrated model is then compared to a model with the same parameters, except for the fact that coordination failures are shut down. The model without coordination failures is also re-calibrated. I now describe the set of results that emerge from my numerical exercises.

First, I show that deep financial crises have large effects on the economy not only during those episodes but also in good times. The economy with coordination failures features significantly lower asset prices, and consequently lower economic growth, even in states where coordination traps are not expected to happen in almost 80 years. During coordination traps, asset prices decline significantly, as banks fire sell capital to households. Since coordination failures happen because banks are under-capitalized (and therefore cutting down lending), they hit banks in states where they assign a high value to their net worth. Therefore, although unlikely, coordination failures affect banks in states where it hurts. Banks fear coordination failures, which reduces their incentives to hold capital even in good times, reducing asset prices substantially. When coordination failures are turned off, the economy grows on average 2.4%, while with coordination failures this number is 1.8%, even though they are rare events.

Second, I show that the interaction of coordination and financial frictions produces a powerful amplification mechanism if at least one of two conditions is satisfied: (i) Shocks are sufficiently large; (ii) Banks are sufficiently under-capitalized when shocks hit. If those conditions do not apply, the model with coordination failures has a similar response to shocks as the model with coordination frictions turned off. However, the economies with and without coordination frictions behave very differently for large shocks, with the main model featuring output losses that can be more than five times larger than those in the model without coordination problems. If shocks are not so large, the two economies still behave very differently if banks have little net worth when the shock hits. This large amplification is explained by firms operating with idle capacity and by the fact that coordination failures depress asset prices and therefore economic growth after the shock. It is also shown that only the model with coordination frictions can match well the output losses observed during financial crises in the data, for shocks of a size roughly consistent with the observed frequency of those events.

Third, the model with coordination failures can generate a lot of heterogeneity in the real effects of financial crises: After a sequence of bad shocks, the fate of the economy is very dependent on the realization of subsequent uncertainty. If, say, a large shock has thrown the economy into a coordination trap, the economy may recover quickly in case positive shocks hit, leaving the coordination trap. Negative shocks, on the other hand, further reduce asset prices and bank intermediation, exacerbating the initial impact of the shock. The interplay of coordination and financial frictions makes the economy very sensitive to small shocks after a financial crisis hit. In contrast, for the economy without coordination failures, the effect of a shock of the same relative size is almost independent of the subsequent history.

Fourth, the welfare costs of deep financial crises are large even in good times. For instance, even in a good scenario where deep crises are not expected in almost 80 years, they lead to consumption-equivalent losses of around 5%. This is largely explained by the effect that those events have on banks' capital demand, which reduces asset prices even in good times. It is also shown that once one shuts down the two-way feedback between coordination and financial frictions, the welfare costs of coordination failures almost vanish. In other words, what explains the large welfare losses of coordination failures is not their direct effect on output, but their effect on banks' behavior, which harms asset prices and economic growth.

Fifth, on the policy front, I study how leverage constraints affect welfare and financial stability. I show that coordination failures can make tight leverage constraints welfare improving at the stochastic steady state. By preventing states in which banks suffer larger losses, it helps to reduce the probability of coordination failures (and hence helps to avoid large output losses), improving welfare, despite making banks constrained in good times. If leverage constraints are not tight enough, they do not bind before a coordination failure hits and are not effective in avoiding them. As a result, loose leverage constraints tend to worsen welfare around the stochastic steady state. Leverage constraints that bind in good times (before a coordination failure hits) tend to improve welfare.

Finally, and still on the normative front, I study how financial innovation that improves household efficiency when intermediating capital affects welfare. It is shown that when households become better at intermediating capital, the probability of coordination failure can increase, which tends to reduce welfare. When coordination failures are shut down, however, financial innovation improves welfare. In other words, the "intermediation paradox" of Phelan (2016) only arises in the model with coordination failures. Intuitively, when households become better at intermediating capital, banks earn less excess returns, as now households are a closer substitute to banks. This negatively affects banks net worth, making them less resilient to bad shocks and more likely to substantially cut down intermediation, which causes coordination failures.

While the macroeconomic literature has made progress incorporating different types of financial frictions that can amplify the response of real variables to shocks, two characteristics of financial crises remain puzzling through the lens of standard models (such as Kiyotaki and Moore, 1997 and Brunnermeier and Sannikov, 2014, to name a few). The first is the sudden and discrete nature of deep financial crises such as the Great Recession of 2008. The second is the large heterogeneity in output losses in the aftermath of crises that seem similar at first. To deal with the former, models that combine banking panics (that lead to runs) and a usual financial accelerator have been proposed (Gertler and Kiyotaki, 2015, Gertler et al., 2020). Similarly, others have proposed that regime switches in information production about collateral lead to run-like episodes that collapse economic activity (Gorton and Ordonez, 2014, Gorton and Ordonez, 2020). However, as documented by Baron et al. (2021), not all financial crises feature banking panics, and those crises can lead to large collapses in economic activity even in episodes where no bank run was observed. Hence, this paper also contributes to that literature by providing a theory that can rationalize why some financial crises turn into real economic disasters and others do not, even in the absence of bank runs.[2]

This paper is also related to several other strands of literature. The kind of coordination problem studied here is present in the seminal contributions of Kiyotaki (1988), Cooper and John (1988) and Murphy et al. (1989). Recent papers along this line, such as Guimaraes and Machado (2018) and Schaal and Taschereau-Dumouchel (2018), use equilibrium selection techniques and are able to generate strong and long-lasting recessions in dynamic models.[3]

In Schaal and Taschereau-Dumouchel (2018), there is a feedback between capital accumulation and coordination. When savings are low, firms are less likely to coordinate on high capacity, which further decreases capital accumulation. In Guimaraes and Machado (2018), firms adjust their capacity sequentially, and the economy may get stuck in a coordination trap simply because no firm wants to be the first to switch to a high regime. Schaal and Taschereau-Dumouchel (2019) explore the interaction of demand externalities and labor market frictions: high aggregate demand leads to more vacancy posting, which leads to lower unemployment and higher demand. Here, the feedback comes from banks' balance sheets to coordination, which is motivated by the goal of understanding deep financial crises.

A key assumption in all those papers, including this one, is that firms face some non-convexity, such as a fixed cost, in their production technology. This is necessary to create enough feedback from aggregate demand to firms' decisions. Non-convexities played an important role in the 2008 crisis. For instance, using Greek data, Fakos et al. (2018) document that the share of inactive firms (measured as firms with near-zero investment) almost doubled from 2008 to 2011, and that credit conditions explain a lot of this variation. Crane et al. (2020) documents that temporary establishment closures are common, and that establishment closures increased significantly around the Great Recession, mostly driven by small firms. As argued by Schaal and Taschereau-Dumouchel (2019), there

[2] Other examples of papers that combine financial accelerator models with bank runs are Clymo (2016) and Gertler et al. (2016). Differently from those papers, that rely on sunspots to generate deep crises, here those events happen as the result of fundamental shocks only. For a different but broadly related approach, see Boissay et al. (2016), who offer a macroeconomic model of banking crises based on interbank market freezes.

[3] See also Guimaraes et al. (2016) for a static macroeconomic model with coordination failures.

is evidence that firms are subject to non-convexities when adjusting different factors of production, such as labor (Caballero et al., 1995), capital (Cooper and Haltiwanger, 2006) and capacity utilization (Hall, 2000). In particular, for the automobile industry, Hall (2000) shows that accounting for the fact that firms adjust the number of production shifts and the number of operating plants discretely is essential to understand observed production patterns, a message that is reinforced by Copeland and Hall (2011). Also, as shown in Online Appendix C.2, aggregate capacity utilization featured sudden and large movements during the 2008 crisis, being able to explain most movements in output in that period. Consistently with the aforementioned evidence and following Guimaraes and Machado (2018) and Schaal and Taschereau-Dumouchel (2019), this paper introduces a non-convexity in capacity utilization.

The consequences of coordination frictions and multiple equilibria in macroeconomics have been explored in different contexts. A branch of the literature has relied on "sunspots" to drive agents' expectations in real business cycle models, creating the possibility of agents coordinating on an inefficient equilibrium. Early contributions along those lines include Benhabib and Farmer (1994) and Farmer and Guo (1994).[4] Multiple equilibria in macroeconomics can arise from the interplay of productivity growth and aggregate demand (Benigno and Fornaro, 2016), Taylor rules and the zero lower bound (Benhabib et al., 2001), decentralized trading and precautionary savings (Chamley, 2014), to name a few examples.

As previously mentioned, this paper naturally relates to the large literature that incorporates financial frictions in macroeconomic models, as in the seminal papers of Bernanke and Gertler (1989) and Kiyotaki and Moore (1997). As in that literature, my model features a friction preventing agents from issuing fully state-contingent financial claims. More specifically, banks are not allowed to issue new equity and need to fund themselves exclusively through debt. Due to the highly non-linear nature of coordination failures, however, the model needs to be solved away from steady state, and a continuous-time approach is employed, in the spirit of Brunnermeier and Sannikov (2014).[5]

The remainder of this paper is organized as follows. Section 2 presents the model, Section 3 discusses its solution, with the calibration being discussed in Section 4. Section 5 presents the main positive results. Section 6 discusses welfare and policy implications. Section 7 concludes. All proofs are relegated to Online Appendix A.

2. Model

2.1. Preliminaries

Time is continuous and indexed by t. There are three types of agents: banks (financial intermediaries), households and firms. Households and banks are in a unit-measure continuum each. They can hold capital and a risk-free bond in their portfolios. At the initial date, each household and bank holds no bonds and has a given amount of capital k_0^h and k_0^b, respectively. Banks and households can rent capital to firms in a competitive market. There are two types of firms: a unit-measure continuum of intermediate goods producers, and a competitive final good producer. Intermediate goods firms use capital and labor to produce differentiated varieties, which are aggregated by the competitive final good producer.

Capital held by banks (k_t^b) and households (k_t^h) evolves according to

$$dk_t^i = g_t^i k_t^i dt + \sigma k_t^i dZ_t,$$

where $i \in \{h, b\}$, g_t^b and g_t^h are the capital growth rates for banks and households, respectively, dZ_t are aggregate Brownian shocks (capital quality shocks), and $\sigma > 0$ is the exogenous volatility. Increasing capital at rate g_t^i entails standard convex investment costs $\iota(g_t^i)$. For notational convenience, I include the capital depreciation rate in the function $\iota(\cdot)$. Total capital in the economy is $K_t = k_t^b + k_t^h$.

I restrict attention to equilibria where the price of capital q_t follows a diffusion process:

$$dq_t = \mu_t^q q_t dt + \sigma_t^q q_t dZ_t, \tag{1}$$

where μ_t^q and σ_t^q are endogenous equilibrium objects.

2.2. Households

Households own banks and firms. They are risk neutral and can have negative consumption, so the equilibrium risk-free rate is constant and equal to households' discount rate ρ.[6] Households can buy capital and rent it to firms at a rental rate R_t or invest in the risk-free bond that pays an interest rate r_f. To capture the fact that households are not as good as banks at intermediating capital, I assume that for each unit of capital households lend to firms, only a fraction $\beta \in (0, 1)$ pays the rental rate. The justification is as follows: There is a fraction β of productive firms that pay back the promised rental rate, and a fraction $1 - \beta$ of unproductive firms that deliver no output and then default on their promised rental rate—bad firms do not run away with the capital rented, they only

4 Angeletos and La'O (2010) and Angeletos and La'O (2013) formalize the idea of shifts in expectations (akin to sunspots) in unique equilibrium models.

5 See also He and Krishnamurthy (2019) and He and Krishnamurthy (2011). For comprehensive surveys on the recent continuous-time macrofinance literature, see Isohätälä et al. (2016) and Brunnermeier et al. (2013). For a study of leverage constraints in a model similar to Brunnermeier and Sannikov (2014), see Phelan (2016). For a recent contribution incorporating household heterogeneity in a continuous-time macrofinance model, see Fernández-Villaverde et al. (2022).

6 This assumption follows Brunnermeier and Sannikov (2014) and ensures households can always rebuild their net worth if it drops to zero (which never occurs in equilibrium). One can think that households have access to some technology to produce the final good themselves, and negative consumption is the disutility incurred by households when doing so. In equilibrium, consumption is always positive.

do not pay the agreed rental rate as they deliver no output. Households are unable to screen firms and end up granting bad loans. The household rental rate is then $R_t^h \equiv \beta R_t$, since that is the effective rate they receive. I call β the direct finance efficiency.

Using Ito's Lemma, we can write the households' realized capital returns dr_t^{kh} as

$$dr_t^{kh} = \underbrace{\frac{R_t^h - \iota_t^h}{q_t}}_{\text{dividend yield}} dt + \underbrace{\left(g_t^h + \mu_t^q + \sigma\sigma_t^q\right) dt + \left(\sigma + \sigma_t^q\right) dZ_t}_{\text{capital gains rate } \left(dq_t k_t^h / q_t k_t^h\right)}.$$

The households' problem is given by

$$\begin{aligned} \max_{c_t, \varphi_t^h, g_t^h, l_t} \quad & \mathbb{E}_\tau \left[\int_\tau^\infty e^{-\rho(t-\tau)} \left(c_t - \xi l_t\right) dt \right] \\ \text{s.t.} \quad & dn_t^h = \varphi_t^h n_t^h dr_t^{kh} + (1-\varphi_t^h) n_t^h r_f dt + \left(\zeta_t + \Pi_t + w_t l_t - c_t\right) dt, \\ & n_t^h \geq 0, \end{aligned}$$

where c_t denotes consumption, l_t denotes hours worked, w_t is the wage, φ_t^h is the share of the net worth n_t^h invested in capital, ζ_t and Π_t denote dividends paid by banks and firms, respectively, and $\xi > 0$ is the marginal disutility of labor. Since agents are risk neutral, any inefficiency that arises in equilibrium comes from distortions in production, capital allocation and/or investment, not from consumption smoothing. However, in Online Appendix D.1, I propose an extension of the model where households discount for risk in their utility and show that the main results are similar.

2.3. Banks

Each bank is managed by a banker. Bankers maximize the discounted dividends they pay during their lifetime, but are subject to mortality risk: in each infinitesimal interval of length dt, a banker exits with probability λdt. When a banker exits, it is immediately replaced by a new banker who inherits her assets. As usual in the macrofinance literature, this kind of assumption is necessary to prevent banks from accumulating too much net worth.[7]

Banks' dividend payouts ζ_t cannot be negative, that is, banks cannot issue new equity. The only way banks can obtain extra funds to acquire capital is through risk-free debt, having to absorb all the risk in their balance sheets. This restriction is standard in macrofinance models and it can be justified by an incentive problem of bankers that requires them to have some "skin in the game", as extensively shown in the corporate finance literature.[8]

Banks have a superior intermediation technology and do not waste capital (they only rent it to productive firms). Therefore, they earn a capital rental rate $R_t^b \equiv R_t$. Similarly to households, banks' realized capital returns are

$$dr_t^{kb} = \frac{R_t^b - \iota_t^b}{q_t} dt + \left(g_t^b + \mu_t^q + \sigma\sigma_t^q\right) dt + \left(\sigma + \sigma_t^q\right) dZ_t.$$

The banks' problem is given by

$$\begin{aligned} \max_{\zeta_t, \varphi_t^b, g_t^b} \quad & \mathbb{E}_\tau \left[\int_\tau^\infty e^{-(\rho+\lambda)(t-\tau)} \zeta_t dt \right] \\ \text{s.t.} \quad & dn_t^b = \varphi_t^b n_t^b dr_t^{kb} + (1-\varphi_t^b) n_t^b r_f dt - \zeta_t dt, \\ & n_t^b \geq 0, \zeta_t \geq 0, \end{aligned}$$

where φ_t^b is the share of the net worth n_t^b invested in capital. Besides the fact that households have additional sources of income (dividends and wages), the main difference between the banks' and the households' problems is the non-negativity constraint $\zeta_t \geq 0$. One can interpret it as a solvency constraint: given that the bank cannot issue new equity, whenever a banker's net worth drops to zero she will be wiped out.

A notation remark: For future reference, I refer to the drift of banks' and households' realized capital returns as $\hat{r}_t^b$ and $\hat{r}_t^h$, respectively, that is, $\hat{r}_t^j = \frac{R_t^j - \iota_t^j}{q_t} + \left(g_t^j + \mu_t^q + \sigma\sigma_t^q\right)$, for $j = h, b$. I refer to those variables as banks' and households' expected capital returns, as usual, as they reflect the expected cash flow obtained per unit of capital.

[7] Mortality risk guarantees a non-degenerate stationary distribution of wealth.

[8] See Sannikov (2013) for a survey of this literature.

2.4. Firms

There are two types of productive firms: a continuum of intermediate goods producers $j \in [0,1]$, each producing a differentiated good using capital and labor; and a competitive final good producer that uses intermediate goods as inputs. Moreover, there is a mass $(1-\beta)/\beta$ of unproductive firms. The only role of unproductive firms is to use a fraction β of households' capital in each period without actually producing anything. As will be clear, this implies that when households intermediate capital, part of the aggregate demand that could be generated will be lost, making the economy prone to demand-driven coordination failures.

2.4.1. Final good producer

The competitive final good firm uses intermediate goods to produce the final good Y_t, according to the production function

$$Y_t = \left(\int_0^1 y_{j,t}^{\frac{\varepsilon-1}{\varepsilon}} dj \right)^{\frac{\varepsilon}{\varepsilon-1}},$$

where $\varepsilon > 1$ is the elasticity of substitution and $y_{j,t}$ is the quantity of good j used as input. This technology implies the usual demand for each variety:

$$p_{j,t} = \left(\frac{Y_t}{y_{j,t}} \right)^{\frac{1}{\varepsilon}} P_t, \tag{2}$$

where P_t is the price of the final good, which is normalized to one.

2.4.2. Intermediate goods producers

Intermediate goods firms have a standard Cobb-Douglas production function:

$$y_{j,t} = u_{j,t} A k_{j,t}^{\alpha} l_{j,t}^{1-\alpha}, \tag{3}$$

where $u_{j,t}$ is a capacity utilization variable of firm j, $k_{j,t}$ and $l_{j,t}$ are the quantities of capital and labor used by firm j, respectively, and $\alpha \in (0,1)$ is the capital elasticity.

To generate a coordination problem among firms, one must introduce some non-convexity in the model, such as a fixed cost. As discussed in the introduction, it has been documented that firms face non-convexities in different types of decisions, such as investment, labor and capacity utilization. Capacity utilization was a particularly important margin of adjustment for firms during the 2008 financial crisis. As shown in Online Appendix C.2, capacity utilization featured a large and sudden decline in 2008, and a simple growth-accounting exercise reveals that this variable alone can explain most of the movements in output during the Great Recession. Hence, I add a non-convexity to the model via firms' capacity utilization decisions.

Following Guimaraes and Machado (2018) and Schaal and Taschereau-Dumouchel (2018), it is assumed that capacity can only take two values, $u_t \in \{u_L, u_H\}$, with $u_H > u_L$. Choosing the low capacity u_L can be interpreted as inaction. Adopting high capacity entails a fixed cost $f_t > 0$ (in units of the final good). Consistently with the fact that capacity featured a quick fall during the beginning of the Great Recession, followed by a relatively quick partial recovery in 2009 and 2010, I follow Schaal and Taschereau-Dumouchel (2018) and assume that capacity decisions are made every period, so that capacity is not by assumption a slow-moving variable as in Guimaraes and Machado (2018).

Since the model presented here features endogenous growth, one needs fixed costs to increase as the economy grows. Otherwise, firms would grow out of the fixed cost eventually and capacity decisions would become trivial. Hence, it is important to eliminate size effects, as in Cooper and Haltiwanger (2006). There are different assumptions that could deliver that, but a particularly tractable one is to assume that fixed costs scale with aggregate capital. I then assume that the fixed cost firms pay to increase capacity is given by $f_t = \chi K_t$, where $\chi \geq 0$ denotes the scaled fixed cost. That is, fixed costs are proportional to firms' size. This linear specification is particularly useful since it preserves the scale-invariance properties of the economy, and it is common in the investment and corporate finance literature.[9]

Firms take the wage w_t and the rental rate R_t as given. They choose capacity $u_{j,t}$, capital $k_{j,t}$ and labor $l_{j,t}$ to maximize profits

$$\pi_t = p_{j,t} y_{j,t} - w_t l_{j,t} - R_t k_{j,t} - \mathbb{1}_{\{u_{j,t}=u_H\}} f_t, \tag{4}$$

subject to the demand schedule (2) and the production function (3). To ease the exposition, the firms' problem is split in two stages. In a first stage, firms choose capacity $u_{j,t}$. In a second stage, taking as given the fraction of firms with high capacity, denoted by x_t, they choose capital and labor.

[9] See Cooper and Haltiwanger (2006), Bolton et al. (2011) and Décamps et al. (2016).

3. Equilibrium

Having described each agent's problem, the equilibrium definition is standard. An equilibrium is a set of stochastic processes for prices and quantities defined on some filtered probability space adapted to the Brownian motion $\{Z_t, t \geq 0\}$ such that: (i) banks, households and firms solve their problems, and (ii) markets clear.

Given banks' and households' capital (k_t^b and k_t^h), one can find the equilibrium on the real sector. The equilibrium characterization is then divided in two blocks. First, for a given k_t^b and k_t^h, one can solve for firms' decisions. Then, using the map from banks' and households' capital to firms' decisions as an input, one can solve for banks' and households' optimal decisions and asset prices.

3.1. Firms equilibrium

In this section, I describe how to derive the equilibrium of firms for a given level of banks' and households' capital. The equilibrium is solved backwards. First, firms' production decisions are computed for a given proportion x_t of firms with high capacity (second stage). Then, firms' capacity decisions are determined (first stage).

3.1.1. Second stage equilibrium

As firms with the same capacity make the same decisions in equilibrium, firm subscripts j are often replaced by L or H, which denote firms in the low and high regime, respectively.

Firms optimality Remember that firms take the wage w_t, the rental rate R_t and the aggregate demand Y_t as given. The aggregate demand Y_t shifts the demand schedule each firm faces. Let $mc_{j,t}$ denote a firm's marginal cost. It can be shown that[10]

$$mc_{j,t} = \frac{1}{u_{j,t}A}\left(\frac{R_t}{\alpha}\right)^{\alpha}\left(\frac{w_t}{1-\alpha}\right)^{1-\alpha}. \tag{5}$$

Firms with high capacity have a lower marginal cost (with $mc_{L,t} = \left(u_H/u_L\right) mc_{H,t}$), but have to pay the fixed cost f_t. Due to market power, firms optimally set their prices above the marginal cost, with $mc_{j,t} = \frac{\varepsilon-1}{\varepsilon} p_{j,t}$. Using the demand schedule (2), this implies that

$$mc_{j,t} = \frac{\varepsilon-1}{\varepsilon}\left(\frac{Y_t}{y_{j,t}}\right)^{\frac{1}{\varepsilon}}, \tag{6}$$

which, together with (5), pins down a firm's choice of $y_{j,t}$. Capital expenditures represent a share α of total costs, and therefore

$$k_{j,t} = \alpha\frac{y_{j,t}mc_{j,t}}{R_t} \quad \text{and} \quad l_{j,t} = (1-\alpha)\frac{y_{j,t}mc_{j,t}}{w_t}.$$

Market clearing Due to the linear disutility of labor, the labor supply is flat and the labor market always clears at the wage $w_t = \xi$. Capital and goods market clearing imply

$$x_t k_{H,t} + (1-x_t)k_{L,t} = k_t^b + \beta k_t^h$$

and

$$Y_t = \left(x_t y_{H,t}^{\frac{\varepsilon-1}{\varepsilon}} + \left(1-x_t\right) y_{L,t}^{\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{\varepsilon}{\varepsilon-1}}.$$

Using firms' optimality and market clearing, we find the equilibrium in stage two for a given proportion of firms with high capacity, x_t. Proposition 1 characterizes the second stage equilibrium. The variable ψ_t is the share of total capital intermediated by banks, $\psi_t \equiv k_t^b/K_t$.

Proposition 1. *Let the (weighted) average capacity $\bar{u}_t$ and the endogenous TFP $\overline{A}_t$ be*

$$\bar{u}_t \equiv \left(x_t u_H^{\varepsilon-1} + \left(1-x_t\right)u_L^{\varepsilon-1}\right)^{\frac{1}{\varepsilon-1}} \quad \text{and} \quad \overline{A}_t \equiv A\bar{u}_t\left(\psi_t(1-\beta)+\beta\right)^{\alpha}.$$

Then, in equilibrium, wages and the rental rate of capital are

$$w_t = \xi \quad \text{and} \quad R_t = \alpha\left(\frac{\xi}{1-\alpha}\right)^{\frac{\alpha-1}{\alpha}}\left[\left(\frac{\varepsilon-1}{\varepsilon}\right)A\bar{u}_t\right]^{\frac{1}{\alpha}}.$$

Aggregate labor and gross output are

[10] The details are provided in the proof of Proposition 1.

$$L_t = \left[\left(\frac{1-\alpha}{\xi} \right) \left(\frac{\varepsilon - 1}{\varepsilon} \right) \overline{A}_t \right]^{\frac{1}{\alpha}} K_t \quad \textit{and} \quad Y_t = \overline{A}_t K_t^{\alpha} L_t^{1-\alpha}.$$

For $j \in \{L, H\}$, firms' production and profits are

$$y_{j,t} = \left(\frac{u_{j,t}}{\overline{u}_t} \right)^{\varepsilon} Y_t \quad \textit{and} \quad \pi_{j,t} = \frac{1}{\varepsilon} \left(\frac{u_{j,t}}{\overline{u}_t} \right)^{\varepsilon - 1} Y_t - \mathbb{1}_{\{u_{j,t} = u_H\}} f_t.$$

Note that I refer to Y_t as gross output since it ignores the fixed costs paid. Recall that the final good is used as an input to pay for firms fixed costs. Thus, the actual output available for investment and consumption is given by $Y_t - x_t f_t$.

Proposition 1 shows that the endogenous TFP of the economy depends on the average capacity $\overline{u}_t$ and on the share of capital intermediated by banks ψ_t. The effect of ψ_t on $\overline{A}_t$ summarizes the role of capital allocation on output, as in the seminal contribution of Kiyotaki and Moore (1997). The effect of firms' capacity choices on output is summarized by the variable $\overline{u}_t$, which captures the coordination effect, as in Schaal and Taschereau-Dumouchel (2018).

Proposition 1 also shows that the rental rate of capital increases when firms coordinate on high capacity. Intuitively, when firms coordinate, they demand more capital, raising rental rates. Thus, coordination increases banks' and households' incentives to hold capital.

Define $\Delta\pi_t = \pi_{H,t} - \pi_{L,t}$ as the relative gain of being in the high regime (relative profits) and let $\Delta\hat{\pi}_t \equiv \Delta\pi_t / K_t$ be the scaled relative profits. Using Proposition 1, we can write it as

$$\Delta\hat{\pi}_t = \mathcal{B}\overline{u}_t^{\Omega} \left(\psi_t (1-\beta) + \beta \right) - \chi, \tag{7}$$

where $\mathcal{B} \equiv \left[(1-\alpha)(\varepsilon - 1)/\xi \right]^{\frac{1-\alpha}{\alpha}} (A/\varepsilon)^{\frac{1}{\alpha}} \left(u_H^{\varepsilon-1} - u_L^{\varepsilon-1} \right)$ and $\Omega \equiv 1 + 1/\alpha - \varepsilon$. One can see that the return of paying the fixed cost is affected by two endogenous variables, ψ_t and $\overline{u}_t$. Therefore, we can write $\Delta\hat{\pi}_t = \Delta\hat{\pi}(\psi_t, \overline{u}_t)$. Note that relative profits are increasing in ψ_t. The more capital banks intermediate, the more aggregate demand there is in equilibrium for a given average capacity, as capital flows to productive firms. This increases incentives for firms to pay the fixed cost and scale up their production.

As for average capacity, $\Delta\hat{\pi}_t$ increases with $\overline{u}_t$ when $\Omega > 0$. In this case, firms' capacity decisions are *strategic complements*, that is, a firm has higher incentives to increase its production when others do the same. When $\Omega < 0$, instead, firms' capacity decisions are *strategic substitutes*. This ambiguity arises because there are two effects at play. On the one hand, an increase in average capacity raises the aggregate demand Y_t, shifting up the demand schedule each firm faces. This high demand provides incentives to pay the fixed cost and scale up production. On the other hand, an increase in capacity raises the demand for capital, and consequently its rental rate, reducing firms' absolute profits and their incentives to expand. Which effect dominates depends on the sign of Ω. This tension between demand externalities and competition in the market for factors is a common characteristic of models with demand externalities. From now on it is assumed that $\Omega > 0$, since it holds in my calibration, and hence I focus on the case with coordination failures. The case with $\Omega \leq 0$ is discussed in Online Appendix D.3.

3.1.2. *First stage equilibrium*

In the first stage, firms play a capacity game. Proposition 2 characterizes firms' capacity decisions in equilibrium for different values of banks' capital share ψ_t.

Proposition 2. *Let $\underset{\sim}{\psi}$ and $\widetilde{\psi}$ be such that $\Delta\hat{\pi}(\underset{\sim}{\psi}, u_H) = \Delta\hat{\pi}(\widetilde{\psi}, u_L) = 0$. Then:*

1. *If $\psi_t < \underset{\sim}{\psi}$, there is a unique equilibrium and all firms choose low capacity ($x_t = 0$);*
2. *If $\psi_t > \widetilde{\psi}$, there is a unique equilibrium and all firms choose high capacity ($x_t = 1$);*
3. *If $\psi_t \in \left[\underset{\sim}{\psi}, \widetilde{\psi}\right]$, there are three equilibria: $x_t = 0$, $x_t = 1$ and $x_t = h(\psi_t) \in [0,1]$, where $h(\cdot)$ is a decreasing function of ψ_t.*

It is easy to see that $\widetilde{\psi} > \underset{\sim}{\psi}$. To guarantee that the three cases in Proposition 2 can happen, I impose the following parametric condition throughout the paper: $\underset{\sim}{\psi} > 0$ and $\widetilde{\psi} < 1$. For $\psi_t < \underset{\sim}{\psi}$, choosing low capacity is dominant, and for $\psi_t > \widetilde{\psi}$, high capacity is the dominant action for firms. For $\psi_t \in \left[\underset{\sim}{\psi}, \widetilde{\psi}\right]$, there is no dominant strategy and firms are willing to scale up their production only if they are *confident* that a sufficient number of firms will do the same. Choosing low capacity in the interval $\left[\underset{\sim}{\psi}, \widetilde{\psi}\right]$ is a coordination failure, as firms would be better off if they coordinated on the high regime (and doing so is consistent with equilibrium).

In fact, choosing low capacity can be inefficient even when it is a dominant action for firms ($\psi_t < \underset{\sim}{\psi}$). The reason is that firms do not internalize that when they choose high capacity they increase the demand for others. Hence, even when the equilibrium is unique it can be inefficient. For instance, if ψ_t is slightly below $\underset{\sim}{\psi}$, all firms would happily sign a contract committing to choose high capacity. However, no firm has incentives to honor that contract, since choosing low capacity is a dominant choice in those circumstances. Next proposition provides a sufficient condition for firms to always prefer a regime where all choose high capacity ($x_t = 1$) versus a low regime ($x_t = 0$), regardless of the amount of bank intermediation ψ_t. In the calibrated model (Section 4), that condition will hold.

Proposition 3. *Suppose that the following condition holds:*

$$\frac{\chi - \beta\mathcal{D}}{\mathcal{D}(1-\beta)} < 0, \text{ where } \mathcal{D} \equiv \frac{1}{\varepsilon} A^{\frac{1}{\alpha}} \left[\left(\frac{1-\alpha}{\xi}\right)\left(\frac{\varepsilon-1}{\varepsilon}\right)\right]^{\frac{1-\alpha}{\alpha}} \left(u_H^{\frac{1}{\alpha}} - u_L^{\frac{1}{\alpha}}\right). \tag{8}$$

Then, for all $\psi_t \in [0,1]$, firms' equilibrium profits are higher if $x_t = 1$ than if $x_t = 0$.

Equilibrium selection

Firms play a coordination game with multiple equilibria. In some regions of the state space, firms may fail to coordinate on high capacity. A central message of the global games literature is that multiplicity in coordination games usually relies on strong assumptions about agents' knowledge of the behavior of others in equilibrium.[11] Some equilibria can only be sustained if agents face no strategic uncertainty about the action of others. A small departure from this benchmark may lead to equilibrium uniqueness.[12]

The global game approach of the continuous-time debt-run model of Morris and Shin (2001) is applied to uniquely pin down firms' capacity choices. I explain the procedure in Online Appendix A.8. The main idea is to consider a slightly perturbed version of the model where firms observe banks' capital share ψ_t with a small amount of idiosyncratic noise.[13]

In the unique equilibrium selected by global games, firms choose high capacity whenever banks intermediate a large enough fraction of the capital in the economy: There is a cutoff ψ_{GG} such that all firms choose high capacity if $\psi_t \geq \psi_{GG}$ and low capacity if $\psi_t < \psi_{GG}$.[14] Intuitively, if a large fraction of capital is channeled to unproductive firms (ψ_t is low), productive firms anticipate that gross output Y_t will be lower and therefore they will face a low demand (see the demand schedule (2)). These low demand expectations reduce firms' incentives to expand production. Further anticipating that other firms also have low incentives to pay the fixed cost, all firms in the economy refrain from choosing high capacity and the economy ends up in a coordination trap. On the other hand, when banks intermediate a lot of capital (ψ_t is high), firms anticipate a relatively high aggregate demand, even if the economy fails to coordinate, which induces them to choose high capacity. Hence, by avoiding the destruction of part of the aggregate demand, bank intermediation can help firms to coordinate.

The threshold ψ_{GG} is given by the indifference condition of a firm that has a uniform belief over the proportion of firms choosing high capacity (Morris and Shin, 2003). It is given by

$$\psi_{GG} = \left(\frac{1}{1-\beta}\right)\left[\frac{\chi}{\alpha}\left(\frac{\xi}{1-\alpha}\right)^{\frac{1-\alpha}{\alpha}}\left(\frac{1}{A}\frac{\varepsilon}{\varepsilon-1}\right)^{\frac{1}{\alpha}}\left(u_H^{\frac{1}{\alpha}} - u_L^{\frac{1}{\alpha}}\right)^{-1} - \beta\right], \tag{9}$$

which is increasing on the fixed cost (the higher that cost, the more difficult it is to coordinate).

Different selection devices would lead to the same qualitative conclusions. For instance, if we select the best equilibrium, where firms choose high capacity whenever $\psi_t \geq \underset{\sim}{\psi}$, the equilibrium can still be inefficient, as shown in Proposition 3. The dynamics of the economy would be similar, but coordination failures would happen less often. In the calibration, the scaled fixed cost χ is chosen to match a given frequency of coordination failures. Different selection devices that still select a cutoff strategy increasing in χ will only lead to different choices of χ in the calibration. Also, the fact that all firms take the same action in equilibrium is not crucial for the results. In Online Appendix D.2, I present an extension of the model where heterogeneity in fixed costs among firms leads them to choose different actions in equilibrium, and switches between high and low regimes are more gradual. The dynamics and the results are similar, but the model becomes computationally much more demanding.

A key implication of the global games equilibrium, characterized by the cutoff (9), is that what happens in the financial block of the model affects coordination outcomes. Unlike in standard global games, where the fundamental that ultimately determines equilibrium outcomes is exogenous, here, it is going to be determined in equilibrium by banks' and households' dynamic decisions. Banks' and households' decisions, in turn, are affected by coordination outcomes, as those impact the rental rate of capital.

To summarize the discussion so far, the firms equilibrium unravels a two-way feedback that is at the heart of the mechanism proposed: (i) Firms coordination increases their demand for capital, affecting banks' and households' rental rates and their incentives to hold capital; (ii) When banks intermediate more capital, they help firms coordinate on high economic activity by increasing their expectations about aggregate demand.

3.2. Households and banks equilibrium

Given the firms equilibrium, we can solve for households' and banks' decisions. Coordination affects banks' and households' portfolio decisions through its effect on the rental rate of capital. Moreover, portfolio decisions matter for the allocation of capital, creating a two-way feedback between banks' balance sheets and coordination.

---

[11] For a survey on global games, see Morris and Shin (2003).

[12] For instance, consider a firm making its choice knowing that ψ_t is slightly above $\underset{\sim}{\psi}$. Suppose it thinks other firms may have observed ψ_t with a small error and is thus not sure that others know that $\psi_t > \underset{\sim}{\psi}$. If others think that $\psi_t < \underset{\sim}{\psi}$, they will certainly choose low capacity. Thus, it seems unreasonable that a firm will choose high capacity in this case, as it has a lot to lose if others do not follow, but not much to gain if they do.

[13] A technical requirement for this procedure to work is the existence of dominance regions, which is why $\underset{\sim}{\psi} > 0$ and $\widetilde{\psi} < 1$ was assumed.

[14] Agents' actions are not determined when they are indifferent between the two actions, which happens with zero probability. Without loss of generality, I assume that agents choose high capacity whenever $\psi_t = \psi_{GG}$.

As usual in macrofinance models that built upon Brunnermeier and Sannikov (2014), the state of the economy can be summarized by the banks' share of total net worth:

$$\eta_t \equiv \frac{n_t^b}{n_t^b + n_t^h}.$$

As will be shown, one can construct an equilibrium in which all endogenous variables only depend on η_t (possibly after scaling those variables by total capital K_t). Given the assumption that banks and households start off with some capital only, and no positions in bonds, $\eta_0 = k_0^b / \left(k_0^b + k_0^h\right)$, and therefore η_t is a predetermined variable. Since we are looking for equilibria in which asset prices q_t have continuous paths, banks' and households' net worth must also have continuous paths and therefore, by Ito's Lemma, we can write:

$$d\eta_t = \mu_t^\eta \eta_t dt + \sigma_t^\eta \eta_t dZ_t,$$

where μ_t^η and σ_t^η are determined in equilibrium. I look for Markov equilibria where banks always borrow from households, and prices and allocations are functions of η_t: $\psi_t = \psi(\eta_t)$ and $q_t = q(\eta_t)$.

The non-standard feature of the model is the endogeneity of the capital rental rates R_t^b and R_t^h, since those depend on how much capital banks and households have on their balance sheets. This will create complementarities in banks' capital holdings, leading to multiple equilibria, even when the firms equilibrium is unique, as will be further discussed.[15] Let $\underline{R}$ and $\overline{R}$ be given by the equilibrium rental rates defined in Proposition 2 when $\overline{u}_t = u_L$ (low regime) and $\overline{u}_t = u_H$ (high regime), respectively. Given the firms equilibrium, $R_t = \overline{R}$ if $\psi_t \geq \psi_{GG}$ and $R_t = \underline{R}$ otherwise. In equilibrium, R_t will be then a function of the state η_t, so I write $R_t = R(\eta_t)$.

In what follows, I present some optimality conditions and then discuss how to compute the equilibrium. To ease notation, time subscripts may be omitted if no confusion arises.

Optimal investment Let g_t^{i*}, $i \in \{h, b\}$, denote banks' and households' choices of g_t. It is given by the marginal Tobin's Q relationship: $\iota'(g_t^{i*}) = q_t$. Investment only depends on the asset price q_t, and we can write $g_t^{i*} = g_t^{i*}(q_t)$. Since investment is the same for households and banks, superscripts are dropped from now on and I sometimes abuse notation and simply write $g(\eta) = g^*(q(\eta))$ and $\iota(\eta) = \iota(g(\eta))$, when no confusion arises.

From the optimality condition $\iota'(g_t^*) = q_t$, note that the growth rate of capital in equilibrium, g_t^*, is an increasing function of the price of capital q_t (remember that $\iota'(g)$ is increasing in g). Since gross output Y_t scales with capital K_t in equilibrium, the growth rate of output $Y_t - x_t f_t$ will be roughly equal to g_t^* if fixed costs f_t are small relative to gross output (which is the case under the calibrated parameters to be presented). Hence, ultimately, what determines the growth rate of the economy is the equilibrium asset price q_t. Asset prices, in turn, will be affected by financial frictions, coordination frictions and their interaction, as will be further discussed when the numerical solution of the model is presented.

Households optimality Households are happy to hold capital whenever their expected capital returns equal the risk-free rate: $\hat{r}_t^{kh} = r_f = \rho$. They must either be indifferent between capital and consumption, or not holding capital:

$$\frac{\beta R(\eta) - \iota(\eta)}{q(\eta)} + g(\eta) + \mu^q(\eta) + \sigma^q(\eta)\sigma = r_f \text{ if } 1 - \psi(\eta) > 0, \tag{10}$$

$$\frac{\beta R(\eta) - \iota(\eta)}{q(\eta)} + g(\eta) + \mu^q(\eta) + \sigma^q(\eta)\sigma \leq r_f \text{ if } 1 - \psi(\eta) = 0. \tag{11}$$

Banks optimality Let $v(n^b, \eta)$ denote banks' value function. It depends on the aggregate state η and on individual net worth n^b. The homogeneity of the problem implies linearity in net worth. We can then write $v(n^b, \eta) = n^b \theta(\eta)$, where the function $\theta(\eta)$ represents the banks' marginal utility of net worth. Notice that $\theta(\eta) \geq 1$, since banks can always pay their total net worth out as dividends at a given date. Writing down the banks' Hamilton-Jacobi-Bellmann (HJB) equation, we get that banks are happy to hold capital whenever

$$\underbrace{\frac{R(\eta) - \iota(\eta)}{q(\eta)} + g(\eta) + \mu^q(\eta) + \sigma\sigma^q(\eta)}_{\text{expected capital returns}} = r_f \underbrace{- \eta(\sigma + \sigma^q(\eta))\sigma^\eta(\eta)\frac{\theta'(\eta)}{\theta(\eta)}}_{\text{risk premium}}. \tag{12}$$

Note that although banks are risk-neutral, they require a risk premium to hold capital if the marginal utility $\theta(\eta)$ is time-varying.[16] Moreover, the dividend policy must satisfy:

$$\zeta(\eta) \begin{cases} = 0 & \text{if } \theta(\eta) > 1, \\ \geq 0 & \text{if } \theta(\eta) = 1. \end{cases} \tag{13}$$

[15] The equilibrium set could further expand if we allowed for non-Markov equilibria, sunspot equilibria or equilibria where banks are not always leveraged. A full characterization of the equilibrium set is beyond the scope of this paper and thus I restrict attention to the equilibrium class most studied in the literature.

[16] This happens because banks fear losing future investment opportunities if their net worth drops to zero.

Therefore, whenever $\theta(\eta) = 1$, banks start paying dividends, which implies a reflecting boundary η^* on banks' wealth share ($\eta_t \leq \eta^*$ for every t).

Let $\underline{q}$ denote the price of capital if banks are wiped out and households have to hold all the capital forever. This is given by the Gordon growth formula:

$$\underline{q} = \max_{g<\rho} \left\{ \frac{\beta \underline{R} - \iota(g)}{\rho - g} \right\}.$$

As usual, it is assumed that parameters are such that $\underline{q}$ is a positive finite number. Proposition 4 shows how the equilibrium can be characterized as the solution of a boundary value problem (BVP).

Proposition 4. *The economy is in equilibrium if firms act according to the firms equilibrium and the following conditions hold:*

1. *For a given $\theta(\eta)$, $\theta'(\eta)$, $q(\eta)$ and $q'(\eta)$, the functions $\mu^q(\eta)$, $\sigma^q(\eta)$, $\mu^\eta(\eta)$, $\sigma^\eta(\eta)$, $\psi(\eta)$ and $R(\eta)$ satisfy equations (A.3) to (A.7) in Online Appendix A and they are consistent with the firms equilibrium:*

$$R(\eta) = \begin{cases} \overline{R} & \text{if } \psi(\eta) \geq \psi_{GG}, \\ \underline{R} & \text{otherwise.} \end{cases} \tag{14}$$

2. *The functions $\theta(\eta)$ and $q(\eta)$ satisfy the following BVP with unknown parameter η^*:*

$$\lambda\theta(\eta) = \eta\mu^\eta(\eta)\theta'(\eta) + \frac{1}{2}\left(\eta\sigma^\eta(\eta)\right)^2 \theta''(\eta), \tag{15}$$

$$\mu^q(\eta) = \eta\mu^\eta(\eta)\frac{q'(\eta)}{q(\eta)} + \frac{1}{2}\left(\eta\sigma^\eta(\eta)\right)^2 \frac{q''(\eta)}{q(\eta)}, \tag{16}$$

with boundary conditions given by (A.10) in Online Appendix A.

Equation (14) was previously discussed and follows from the firms equilibrium. Equation (15) is obtained from banks' HJB equation and banks' optimality (12), and (16) is obtained by applying Ito's Lemma to $q(\eta)$. The other variables are obtained using households' and banks' optimality conditions (10) to (12), and the laws of motion for η_t and q_t given by Ito's Lemma.[17] I refer to the set $[0, \eta^*]$ as the support of an equilibrium, as η always moves within that set.

If the function $R(\eta)$ were known, Proposition 4 would give us an easy way to compute the equilibrium. One would only have to plug equations (A.3) to (A.7) (defined in the Online Appendix) in the BVP and solve it using a standard BVP solver, as in Brunnermeier and Sannikov (2014). However, the rental rate $R(\eta)$ depends on the equilibrium share of bank intermediation $\psi(\eta)$, since both functions must be consistent with the firms equilibrium.

It turns out that both a high and a low rental rate may be consistent with equilibrium in part of the state space, as will be formalized in the next proposition. Before that, some terminology is useful: I say that an equilibrium characterized by the conditions in Proposition 4 is a *good* (respectively, *bad*) equilibrium if for a given η, $\theta(\eta)$, $\theta'(\eta)$, $q(\eta)$ and $q'(\eta)$, $R(\eta) = \underline{R}$ (respectively, $R(\eta) = \overline{R}$) if, and only if, there is no solution that satisfies the conditions in Proposition 4 with $R(\eta) = \overline{R}$ (respectively, $R(\eta) = \underline{R}$). In simpler terms, the good equilibrium assumes there is no coordination failure whenever that is consistent with the equilibrium conditions, and the bad equilibrium does the opposite.

Proposition 5. *Suppose $\beta\overline{R} < \iota(\rho)$. Then, for any good equilibrium E, there is another equilibrium E' and a non-empty interval I in the support of both equilibria, such that for all $\eta \in I$, $\psi(\eta) \geq \psi_{GG}$ under E and $\psi(\eta) < \psi_{GG}$ under E'.*[18]

Intuitively, if banks expect other banks to hold a lot of capital, then they also expect firms to coordinate on the high regime. Coordination increases the capital rental rate for banks and households, but it turns out that banks' incentives to hold capital are more affected by coordination failures than households'. A coordination failure then reinforces what caused it in the first place: the insufficient aggregate demand brought about by an insufficient level of bank intermediation.

To understand why that is the case, first note that when a coordination failure hits, it must be that $\psi_t < 1$, that is, both banks and households are holding capital. Therefore, it must be that for both types of agents:

expected capital returns = risk-free rate + risk premium,

where the risk premium is zero for households due to risk-neutrality. All else constant, when a coordination failure hits, banks' capital returns fall by $\left(\overline{R} - \underline{R}\right)/q_t$ while households' capital returns fall by $\beta\left(\overline{R} - \underline{R}\right)/q_t$ (see equations (10) and (12)). Households initially have their capital returns less affected by the coordination failure precisely because the fraction of capital channeled to bad loans does not have its rental rate reduced, as those firms default anyway. After the coordination failure hits, the equilibrium drift

[17] See the proof in the Online Appendix for details.

[18] The condition $\beta\overline{R} < \iota(\rho)$ is sufficient but probably not necessary for the result to hold.

and volatility of asset prices will have to adjust to make banks and households hold capital again. However, since the initial wedge between the actual and the required expected capital returns is smaller for households, it is easier for our Walrasian auctioneer to convince households to hold a bit more capital. Therefore, the capital allocation in this economy is re-balanced towards households due to coordination failures: ψ_t jumps downward when a coordination failure hits (see Proposition 6 below for a formalization of this result).[19] But if a coordination failure did not hit, a high ψ_t could still be sustained, justifying coordination among firms.

To summarize, in some regions of the state space, if banks expect other banks to hold a lot of capital, a coordination failure is avoided, which makes banks earn high capital returns, justifying banks' optimistic expectations about other banks' capital holdings. If, on the other hand, banks expect other banks not to hold much capital, a coordination failure hits, which reduces banks' capital returns, justifying initial beliefs. Notice that the self-fulfilling argument does not rely on multiplicity of equilibria in firms' decisions (that was solved with global games). Instead, it arises from the two-way feedback between firms decisions and banks' capital returns.

Fortunately, during the numeric integration of (15) and (16), for a given η there are only two possibilities for $R(\eta)$, it is either $\overline{R}$ or $\underline{R}$. This, as detailed in Online Appendix A.7, allows me to construct an objective criterion to select equilibria. As mentioned, in what I call the good equilibrium, I select the high rental rate $\overline{R}$ whenever it is consistent with equilibrium. In what I call the bad equilibrium, I select the low rental rate $\underline{R}$ whenever it is consistent with equilibrium. To exhaust all possibilities, one could also apply the good equilibrium selection procedure in a given region of the state space and the bad equilibrium procedure in the complementary region. Online Appendix A.7 details how it was numerically verified that the equilibrium selection procedure chosen does not affect the results for the parameters used in the calibration: the good and bad equilibrium are almost identical for the calibrated parameters. Hence, the results presented hereafter simply focus on the good equilibrium.

Before turning to the numerical solution of the model, I partially characterize the local dynamics of the economy when a coordination failure hits analytically. The intuition for the dynamics of ψ_t and expected capital returns around a coordination failure were already discussed above to shed light on the multiplicity of equilibria. I relegate the discussion of the other results to Section 5.2, where those and other results are graphically represented with a numerical solution.

Proposition 6 *(Local dynamics around a coordination failure). Suppose that for $t < \hat{\tau}$ the economy is not in a coordination failure ($R_t = \overline{R}$) and it enters a coordination failure at date $\hat{\tau}$ ($R_{\hat{\tau}} = \underline{R}$). Then, at $t = \hat{\tau}$:*

1. *Capital is re-balanced towards households: ψ_t jumps downward;*
2. *Output, consumption and labor discretely fall.*

Moreover, if households hold capital as the coordination failure approaches ($\psi_t < 1$ for all $t < \hat{\tau}$ in a neighborhood of $\hat{\tau}$), then, at $t = \hat{\tau}$:

3. *Banks earn less expected returns on capital: $\hat{r}^b_t$ jumps downward.*

A technical remark is in order: The condition that households hold capital as we get close to a coordination failure is a sufficient condition for the third statement to hold. Of course, although natural, it is not obvious that this condition will hold in equilibrium, but all numerical simulations indicate that this is always the case, and hence the statement is not vacuous.

4. Calibration

Before analyzing the real effects of deep financial crises, I calibrate the model for the US to ensure it is consistent with usual unconditional moments in the data and with the response of key real variables during the Great Recession of 2008, which I take as the leading example of a deep financial crisis. In what follows, I give an overview of the calibration strategy, relegating a more detailed discussion for Online Appendix B.1, with data sources detailed in Online Appendix C.3. Table 1 shows the calibrated parameters and a summary of the calibration strategy.

Investment is subject quadratic adjustment costs with a capital depreciation rate δ:

$$\iota(g) = g + \delta + \frac{\kappa}{2}(g+\delta)^2, \text{ with } \kappa > 0. \tag{17}$$

The discount rate ρ is fixed to the standard value of 0.05. The marginal disutility of labor (ξ) and the capacity utilization in the low regime (u_L) are normalized to 1. The choice of the remaining parameters is divided in two steps. First, I choose the parameters that individually pin down some of the chosen targets, which are either moments computed using the data presented in Online Appendix C.3 or some statistic estimated by previous work. I then jointly calibrate the remaining parameters using a simulated method of moments. Whenever needed to compute moments, the model is simulated three hundred times for two thousand years using an Euler scheme (initial conditions are randomly drawn according to the stationary distribution of the state), and the simulated data is converted to quarterly frequency.[20] The parameters that are jointly calibrated are χ (scaled fixed cost), β (direct finance efficiency), λ (exit rate of banks), A (exogenous productivity) and σ (volatility of capital quality shocks). After the numerical minimization of the distance between the targets and their counterparts in the model, the relative distance between each targeted moment and its model counterpart is always less than 0.1%, and therefore this statistic is not reported.

[19] The dynamics of ψ_t will be further discussed when the numerical solution of the model is presented.

[20] Some moments can be directly computed using the stationary distribution given by Kolmogorov forward equations, which is less computationally expensive.

Table 1
Calibrated parameters. Time is expressed in years.

Parameter		Value	Main associated target
Panel (a): Individually calibrated parameters			
κ	Adjustment costs	5.46	Investment sensitivity to marginal Q (Gilchrist and Himmelberg, 1995)
δ	Depreciation rate	0.052	Average depreciation of fixed assets
ε	Elasticity of substitution	5	Markup ratio of 1.25 (see Basu, 2019)
u_H	TFP shifter (high regime)	1.0165	Δ Capacity in the Great Recession
α	Capital elasticity	0.24	Average labor share
Panel (b): Jointly calibrated parameters			
χ	Scaled fixed cost	0.0084	Frequency of coordination failures
β	Direct finance efficiency	0.9590	Δ Output-to-capital in the Great Recession
λ	Bankers' exit rate	0.0015	Average banks' leverage
A	Productivity	1.3053	Long-run output growth
σ	Capital quality shocks	0.0129	GDP volatility
Panel (c): Remaining parameters			
ρ	Discount rate	0.05	Standard value
ξ	Marginal disutility of labor	1	Normalization
u_L	TFP shifter (low regime)	1	Normalization

I now present an overview of the targets and refer the reader to Online Appendix B.1 for details. The adjustment cost κ is chosen to match the sensitivity of investment to Tobin's Q estimated by Gilchrist and Himmelberg (1995), and the depreciation rate δ is obtained from the average depreciation of private fixed assets. The elasticity of substitution ε is chosen to match a markup ratio consistent with usual estimates. Capacity utilization in the high regime, u_H, is chosen to capture the response of the Fed index of capacity utilization during the Great Recession. The capital elasticity α is chosen to match the average labor share for the US economy.[21] The jointly calibrated parameters (χ, β, λ, A and σ) are chosen to match: (i) A frequency of coordination failures consistent with the interpretation that since 1929 there were two coordination failures (deep financial crises); (ii) The drop in the output-to-capital ratio observed during the Great Recession; (iii) The average leverage ratio of the banking sector; (iv) The average growth rate of the US economy; and (v) The volatility of output growth. While it was not targeted during the calibration, it was also verified that the obtained direct finance efficiency parameter β is consistent with the differential recovery and default rates of bonds (direct finance) and bank loans (intermediated finance) reported in the literature.

In order to assess the effects of coordination failures, I compare the main model to two benchmark models that do not feature coordination failures, which are detailed below.

No coordination failures benchmark I construct a benchmark with the same parameters as the calibrated model, but in which firms always choose high capacity. This is justified by the fact that condition (8) holds for the obtained parameters, and therefore at any given date all firms would be better off if they could collectively agree to choose high capacity. This is equivalent to setting $u_L = 0$ in the calibrated model, but the results would be similar if we simply fixed the scaled fixed cost χ to zero.[22] I call this benchmark the model with *no coordination failures*, and it should be interpreted as a counterfactual where the coordination mechanism presented in this paper does not operate.

Calibrated benchmark In order to give the best chance for a model without coordination failures to match the data, I also construct a benchmark where coordination failures are shut down by fixing the scaled fixed cost χ to zero, and re-calibrate it using an analogous procedure. The calibration of this benchmark is also detailed in Online Appendix B.1. The parameters obtained are $\lambda = 0.0012$, $A = 1.3032$ and $\sigma = 0.01836$ (the remaining parameters are the same as in the main model). I call this parametrization the *calibrated benchmark*.

4.1. Untargeted moments

I now check whether the model can replicate some untargeted moments in the data, which are reported in Table 2. To make capacity in the model comparable to the Fed capacity utilization index, the capacity variable here refers to a monotonic transformation of the capacity variable in the model: It takes value one whenever the economy is in the high regime and $(1/u_H)^{\varepsilon}$ otherwise (see the discussion on the calibration of u_H in Online Appendix B.1).

The calibration targets the volatility of output, but the main model almost perfectly matches the volatility of consumption, hours and investment, and captures a significant part of the volatility in bank equity. Not surprisingly, the volatility of capacity utilization is lower in the model than in the data. Since capacity is a binary variable in the model, its growth rate equals zero in many quarters

[21] Note that, with market power, the labor share of output is approximately equal to $1-\alpha$ divided by the markup ratio. See Online Appendix B.1 for details.

[22] Under the calibration, the only difference between the model with $\chi = 0$ and the model with $u_L = 0$ is that output is always a bit lower in the latter due to the fixed cost.

Table 2
Unconditional moments. All variables refer to the annualized growth rate between two consecutive quarters. Targeted moments are marked with an asterisk.

	Output	Consumption	Hours	Investment	Bank equity	Capacity
Panel (a): Standard deviations (%)						
Data	3.1*	2.7	3.2	8.3	40.0	6.0
Main model	3.1*	2.7	3.3	8.5	31.0	2.2
No coordination failures	2.2	2.0	2.2	3.9	26.8	0.0
Calibrated benchmark	3.1*	2.9	3.1	4.7	29.0	0.0
Panel (b): Correlation with output						
Data	1.00	0.67	0.70	0.70	0.18	0.71
Main model	1.00	0.97	1.00	0.79	0.38	0.64
No coordination failures	1.00	0.99	1.00	0.90	0.44	NA
Calibrated benchmark	1.00	0.99	1.00	0.93	0.45	NA

and therefore the model misses a lot of high-frequency movements in capacity. Note that coordination failures introduce a good amount of endogenous volatility in the model, as the standard deviation of all variables is considerably lower in the model without coordination failures. By definition, the calibrated benchmark matches the volatility of output, but it generates much less volatility in investment than observed in the data. As for the correlations, all models predict higher correlations than those observed in the data (except for capacity in the main model).

5. The effects of deep financial crises

In this section, I study the effect that deep financial crises—interpreted as demand-driven coordination failures caused by insufficient bank lending—have on equilibrium outcomes.

First, I further discuss how financial crises may propagate differently to the real sector due to coordination failures and depending on how much those affect banks' balance sheets. Second, I analyze how the global dynamics of the economy are affected by those coordination failures, by comparing the equilibrium in the main model to the one where coordination failures are turned off. Third, I focus on how the real economy responds during those episodes, emphasizing how much amplification the interaction between coordination and financial frictions generates. Fourth, I study whether the proposed mechanism can account for the fact that financial crises have very heterogeneous real effects.

5.1. Two types of financial crises

I now further discuss how the model generates different types of financial crises, depending on how banks balance sheets are affected. To that end, Fig. 1 shows some equilibrium functions (for now, the discussion focuses only on the solid lines and on a subset of those graphs). The cutoff $\hat{\eta}$ below which coordination failures happen in the main model is represented by the dashed vertical line, and the terminal points of each curve denote the reflecting boundary η^* at which banks pay dividends. If an economy starts off with $\eta_0 > \eta^*$, banks pay dividends at an infinite rate and the economy instantaneously moves to $\eta = \eta^*$.

The solid line in Fig. 1b—which represents banks' share of capital in the economy where coordination failures are possible—shows that the model generates different types of financial crises, depending on how banks' balance sheets are affected. First, we have "mild" financial crises, in which banks' net worth share η is not high enough to sustain $\psi = 1$, but it is still above $\hat{\eta}$, thus sustaining coordination among firms. Second, we have "deep" financial crises, in which banks are sufficiently under-capitalized ($\eta < \hat{\eta}$) to induce a demand-driven coordination failure in the real sector. Hence, as was already hinted from the firms equilibrium, very poor balance sheets in the banking sector (very low η) cause demand-driven coordination failures in the real sector, while moderately bad balance sheets do not.

Fig. 1i illustrates how output per unit of capital reacts as banks intermediate less capital. As the economy enters the coordination failure region, output suddenly falls 5.4% due to firms switching to the low-capacity regime. Since firms use less capacity once a coordination failure hits, they demand less labor, which explains why labor discretely falls as well (Fig. 1j). Investment (Fig. 1k) moves continuously around a coordination failure, as it only depends on asset prices, which are forward looking and do not jump once the coordination failure hits. Since investment does not jump and output does, consumption must adjust for market clearing to hold in the goods market (Fig. 1l).

5.2. Global responses to deep financial crises

To understand how coordination failures affect equilibrium outcomes, both locally and globally, this section compares the main model to the no coordination failures benchmark (in which u_L is set to zero and firms always choose high capacity). That is, we now compare the solid and dashed lines of Fig. 1.

To get a sense of how likely each state is, Fig. 1a shows the stationary distribution of the state variable η under each parametrization. In the main model, the unconditional probability of the economy reaching a coordination failure is 11.5%, even though the

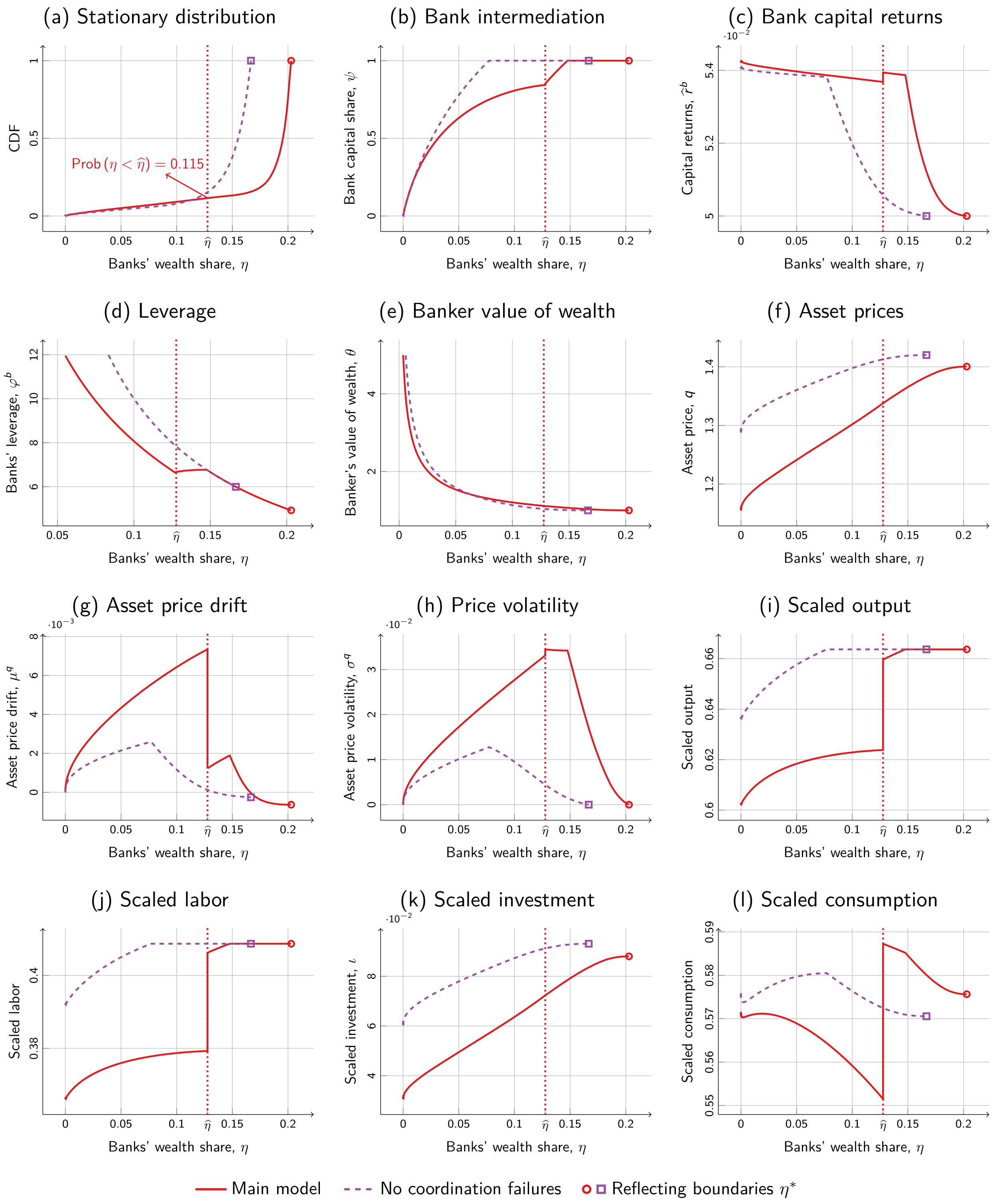

Fig. 1. Equilibrium functions for the model with and without coordination failures.

expected time to reach a coordination failure is 78 years on the reflecting boundary η^*. Although those events are very unlikely in good times, once the economy enters a coordination trap, the expected time it spends in such situation can still be high.

One can see that the possibility of coordination traps makes banks more prone to accumulate net worth, as captured by the higher cutoff η^* with coordination failures. When coordination traps are not possible, banks pay dividends when their share of net worth reaches 16.7%, while this number is 20% in the main model. Given this endogenous response of banks, it is not obvious that the economy with coordination failures will be more volatile and prone to crisis. In fact, from the volatility paradox in Brunnermeier and Sannikov (2014), we know that, all else constant, events that make banks more exposed to risk can reduce overall volatility, once one takes into account the endogenous precautionary behavior of banks.

However, Fig. 1b suggests this is unlikely to be the case. Without coordination failures, banks are much more willing to hold capital. With coordination failures, banks start to unload part of their capital on households soon after bad shocks hit (when their net worth share falls below 14.7%). But this behavior is precisely what causes the coordination failure: If banks managed to intermediate more capital as their net worth dropped, a coordination failure would not happen. Remember that coordination failures also harm households: Once the economy enters a coordination trap, the marginal productivity of capital is reduced (since firms operate with idle capacity), and the rental rate falls for both banks and households. However, as discussed in Section 3.2, the coordination failure affects banks more than households, since the fraction of capital that households lend to firms that default is not affected by a lower equilibrium rental rate. This helps explain why in equilibrium banks' expected capital returns fall once the economy enters a coordination trap, as reported in Fig. 1c.

Importantly, in the main model, banks start to preemptively reduce their capital holdings, even before the coordination failure starts, as reflected by the fact that the solid curve in Fig. 1b is below the dashed curve for many values of $\eta > \hat{\eta}$. Putting it differently, banks want to avoid being highly leveraged in case a coordination failure hits, as depicted by the almost-flat region in Fig. 1d. This is partly because banks have their capital rental rate more affected by coordination failures than households. Intuitively, if both agents feared coordination failures equally, both would try to unload capital if they anticipated a coordination trap, and the bank share of capital ψ_t should not move much. However, there is more to the story.

Fig. 1e shows banks' marginal utility of wealth. As banks' net worth falls, the marginal value of each dollar of net worth increases for them, for two mains reasons: (i) As a bank's *individual* net worth falls, it gets closer to bankruptcy, which bankers want to avoid; (ii) In equilibrium, when *aggregate* bank net worth goes down, capital is being fire sold to households at low prices, which provides good investment opportunities for a rich bank that can buy cheap capital. Households, on the other hand, have a constant marginal utility of wealth. Therefore, the fact that coordination failures are events that happen when banks have a high marginal utility of wealth also contributes for them to act more preemptively than households.

Fig. 1f shows how equilibrium asset prices are affected. First, notice that asset prices in the coordination failure region are much smaller than in good times, which is a result of banks fire selling capital to households (who obtain a lower rental rate), and also a result of coordination failures reducing the rental rate of capital. Second, notice that even in good states, when coordination failures are very unlikely, asset prices are considerably lower in the economy with coordination failures. As previously discussed, although very unlikely, coordination failures hit when bankers assign a high value to their wealth, which reduces their demand for capital and asset prices. Low asset prices reduce investment in equilibrium, which helps explain the differential growth rates of output: On average, the economy with coordination failures grows 1.8% per year, while the economy without coordination failures grows 2.4%. Also, the economy with coordination failures is much more volatile, as shown by the higher volatility of output reported in Table 2.

Figs. 1g and 1h report the drift and volatility of asset prices. As expected from the previous discussion, the economy with coordination failures features more asset price volatility. Also, when the economy enters a coordination trap, the drift of asset prices increases substantially: To convince households to hold capital (as banks have been more affected by the coordination trap), they need to expect asset prices to increase. Since, when in a coordination trap, positive shocks take the economy out of it and negative shocks keep the economy in the low regime, such expectations are consistent with equilibrium.

Asset price volatility initially rises as the economy starts to get bad shocks and a coordination failure approaches, to later decrease if the bad scenario realizes and the crisis deepens. This is not inconsistent with what happened in the 2008 crisis. For instance, the VIX index slowly increased in the second semester of 2007, peaking in October of 2008 and quickly falling afterwards. In the model, price volatility decreases as banks' wealth share becomes very small as to convince banks to still hold some capital. As banks become under-capitalized and get closer to bankruptcy, they require a large risk premium to hold even a small amount of assets.

As a final comment, note that the jump in banks' capital share when the economy enters a coordination failure (predicted by Proposition 6) is not very high (in fact, it is hard to notice given the scale of Fig. 1b). This is partly due to the fact that banks already reduce ψ_t preemptively as they anticipate a coordination failure.

5.3. Impulse responses

The interaction between coordination and financial frictions has the potential to greatly amplify the response of endogenous variables to exogenous shocks. To study how much amplification this mechanism can generate, in this section 1 present impulse response functions (IRFs) to illustrate the dynamics of the model. I define IRFs in the following way. Let X_t be some variable of interest and let "shock" and "no-shock" denote two events. The impulse response of X at horizon k, denoted by $IRF\left(X_{t+k}\right)$, is given by

$$IRF\left(X_{t+k}\right) = \mathbb{E}_t\left[X_{t+k}|\text{shock}\right] - \mathbb{E}_t\left[X_{t+k}|\text{no-shock}\right]. \tag{18}$$

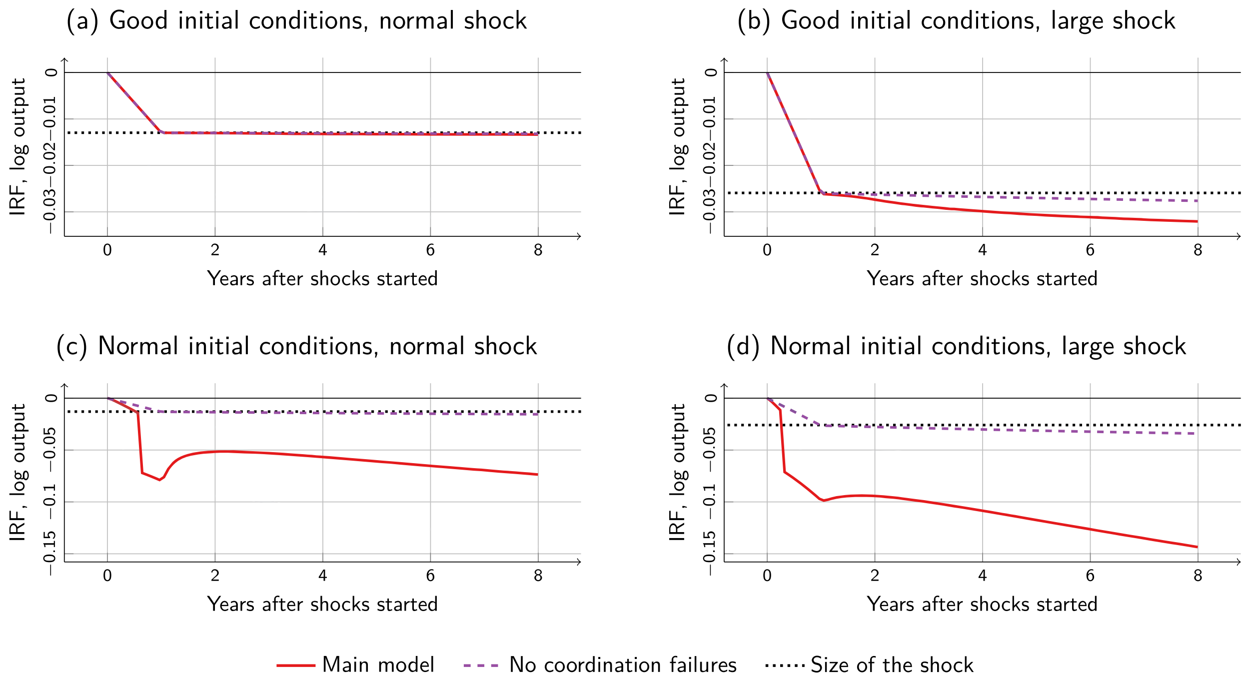

Fig. 2. Impulse responses with and without coordination failures ($u_L = 1$ and $u_L = 0$). Normal shock and large shock refer to a 1-sd and 2-sd negative shock, respectively. Normal and good initial conditions refer to $\eta_0 = 0.2$ and $\eta_0 = 0.15$, respectively.

In other words, the IRF represents how much X is expected to change due to the shock, relative to the no-shock scenario. Instead of considering one-time shocks, I consider shocks that hit the economy continuously during one year. The advantage of that approach is that it gives a clear interpretation of the size of the shock in terms of standard deviations.[23] More precisely, I define a shock of size s starting at date t as the event $\{\sigma dZ_t = s, \text{ for all } t \in [t, t+1]\}$, where $s = \sigma(Z_{t+1} - Z_t)$ is approximately the fraction of the capital stock directly destroyed by the shock.[24] In the no-shock economy, it is assumed that $\sigma dZ_t = 0$ for all $[t, t+1]$, i.e., the event no-shock represents a shock of size zero.[25] Given the non-linearities of the model, one cannot simply set shocks to zero after impact to compute IRFs. Hence, the model must be simulated several times with an Euler scheme to compute the conditional expectations in (18).

A remark about impulse responses is in order. As usual in the macrofinance literature in the tradition of Brunnermeier and Sannikov (2014), exogenous shocks are by assumption permanent. While the growth rate of output is stationary in my setting, output (or capital) does not revert to some trend after a shock. If a shock destroys 5% of the capital in the economy, this automatically reduces output by 5% in a permanent way, even if financial or coordination frictions are not present. Therefore, impulse responses here are meant to capture the endogenous response of the economy to *permanent* shocks, and we are interested in checking how much additional output losses and persistence the model generates beyond what would automatically happen in a frictionless environment. That is why I always report the output change that would happen in a frictionless model by a dotted curve in Fig. 2, which is simply the size of the shock in absolute terms (not in standard deviation units). Any fall in output beyond that is due to the endogenous propagation mechanisms of the model.

Fig. 2 compares the IRFs of the economies with and without coordination failures (i.e., the economies with $u_L = 1$ and $u_L = 0$) under different initial conditions. In the top panels, both economies start with $\eta_0 = 0.2$, which is the reflecting boundary of the main model. Since the reflecting boundary of the economy without coordination failures is smaller, at date 0 banks pay a lot of dividends and move to the reflecting boundary, and the impulse responses take that into account. The bottom panels show impulse responses with worse initial conditions, assuming $\eta_0 = 0.15$ in both economies. The left panels consider a one-standard-deviation negative shock, while the right panels consider a two-standard-deviation negative shock.

As Fig. 2a shows, when initial conditions are good and the shock is not too large, both economies behave very similarly to a frictionless economy: The expected output loss is roughly proportional to the size of the shock. A small shock does not make banks cut down intermediation, as they are initially as well capitalized as they can be. However, as shown in Fig. 2b, a larger shock creates some amplification even if initial conditions are strong. Both economies feature larger output losses, but the one with coordination failures faces considerably larger losses for larger horizons, as the shock increases the risk of banks cutting down intermediation fearing a coordination failure, and the risk of a coordination failure itself.

23 In continuous-time settings with Brownian shocks, a shock in Z_t of one standard deviation in an infinitesimal interval $(t, t+dt)$ has negligible size.

24 Due to the geometric structure assumed for the process of K_t, s represents the log change in K_t directly induced by the shock in an interval of one year.

25 Alternatively, one could define the no-shock event as the empty set, in which case the IRF would be simply $IRF(X_{t+k}) = \mathbb{E}_t[X_{t+k}|\text{shock}] - \mathbb{E}_t[X_{t+k}]$. The results presented here are virtually the same under this alternative definition of the no-shock economy, but the interpretation of results is a bit more nuanced.

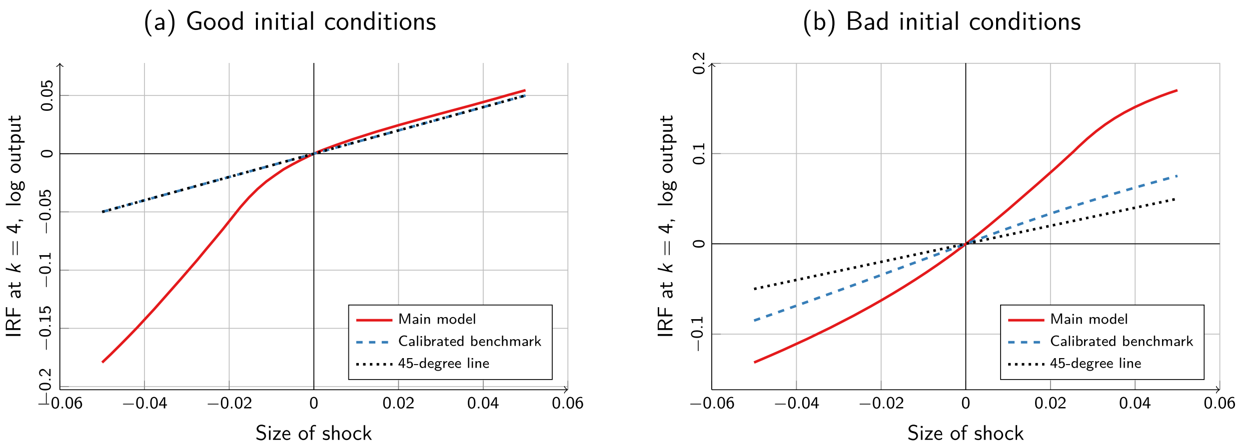

Fig. 3. Impulse responses with and without coordination failures ($u_L = 1$ and $u_L = 0$).

The behavior of the two economies becomes really different if banks are not so well capitalized when the shock hits, as reported in Figs. 2c and 2d. For a one-standard-deviation negative shock (Fig. 2c), the main model already displays a discrete and large collapse in output, while the economy without coordination failures barely shows any amplification. Losses increase through time, as coordination failures reduce asset prices and economic growth in subsequent periods, even in scenarios where the economy leaves the coordination trap but does not fully recover. As the economy enters the coordination trap, there is some expected recovery in the very short run due to an asymmetry: Positive shocks take the economy out of the coordination trap, while negative shocks keep the economy in the low regime. The message is similar with larger shocks (Fig. 2d), and while the economy without coordination failures generates some amplification, it is still much smaller than what the main model delivers.

To summarize, the message here is that the interaction of coordination failures and financial frictions creates powerful and persistent amplification, but only under certain conditions: when shocks are large and/or banks are initially not so well capitalized.

To better understand the non-linearities of the model, Fig. 3 reports the IRF evaluated at horizon $k = 4$ for different sizes of shocks and initial conditions. Instead of simply comparing our economy to the benchmark without coordination failures, as in Fig. 2, it compares the main model to the re-calibrated economy with $\chi = 0$, so that we are comparing two economies that were already mapped to the data. Good initial conditions refer to $\eta_0 = 0.18$, which is slightly above the average value of η in both economies, and bad initial conditions refer to $\eta_0 = 0.08$, so that the main model is already in a coordination trap. The 45-degree line represents the response that would be expected in a frictionless economy.

Fig. 3a shows that, if initial conditions are good, both economies behave similarly to a frictionless one for small negative or positive shocks: Output changes are roughly proportional to the size of the shock. However, for large negative shocks, the economy with coordination failures behaves very differently, with much larger output losses after 4 years, in expectation. The larger the shock, the larger the differences between the two economies. This non-linearity arises because, if the economy starts off at a good situation, positive shocks do not change the regime, while sufficiently large negative shocks drive the economy into a coordination trap.

Fig. 3b shows the same exercise, but starting at very bad conditions (the economy starts in a coordination trap in the main model). Note that the model with only financial frictions generates some amplification: The response to good and bad shocks is larger than in a frictionless economy. However, the response of the economy is still a linear function of shocks, financial frictions only change the slope of the response. The response of the coordination failures economy is again very non-linear, and the difference from the calibrated benchmark becomes larger for positive shocks, as those take the economy out of the coordination trap. As this exercise shows, the interplay of coordination and financial frictions can possibly create very heterogeneous real effects of financial crises, as it not only amplifies the effects of shocks, but also makes the change in output a very non-linear function of shocks.

5.4. Heterogeneous real effects of financial crises

As discussed in the Introduction, and further detailed in Online Appendix C.1, financial crises have very heterogeneous effects on output. As the discussion of Figs. 2 and 3 has suggested, the two-way feedback between banks' balance sheets and firms' coordination failures can potentially explain this heterogeneous propagation of financial crises to the real sector, since the effect of shocks on output is not only amplified in my model, but it is also a highly non-linear function of the size of shocks. I now further investigate how much heterogeneity in propagation to the real sector the model yields and whether it generates enough amplification to explain output losses observed in financial crises with reasonably sized shocks. I compare the main model with the calibrated benchmark that does not feature coordination failures, emphasizing how much heterogeneity in output losses each one can generate after a sequence of negative shocks. Before, I need to introduce some concepts.

Fix a Brownian path Z_t and initial conditions, and let X_t^{shock} and $X_t^{\text{no-shock}}$ be the realization of log output conditional on the shock happening or not happening, respectively. The change in X_t implied by the shock at some horizon k is given by $\Delta X_{t+k} = X_{t+k}^{\text{shock}} - X_{t+k}^{\text{no-shock}}$, and the output loss is defined as $-\Delta X_{t+k}$. The IRFs of Section 5.3 simply computed the expected change $\mathbb{E}_t\left[\Delta X_{t+k}\right]$.

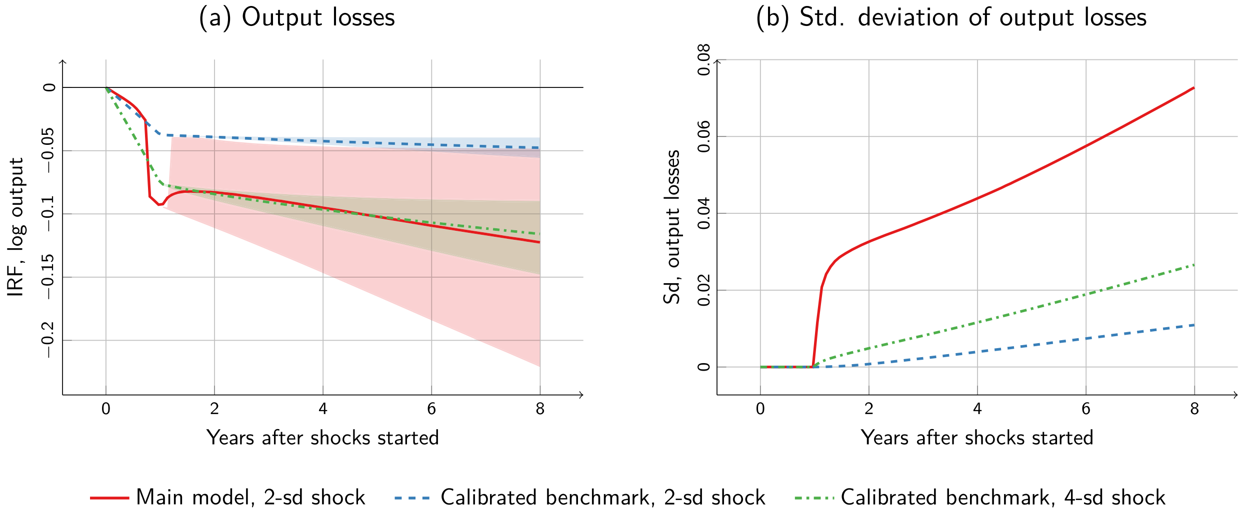

Fig. 4. The heterogeneous effects of financial crises. All shocks are negative shocks. Shaded areas denote 68% confidence bands.

Now, I will also compute other moments of ΔX_{t+k}, in particular the standard deviation and quantiles. Those will help to understand how heterogeneous are the output losses induced by an initial negative shock.

I consider the same type of shock used in Section 5.3, a negative shock that unfolds during one year. For the size of the shock to be roughly consistent with the frequency of financial crises, I consider shocks of 2 and 1.5 standard deviations (sd). Since Brownian shocks are normally distributed, a negative shock larger than 2 sd happens 2.25 times every 100 years, while this number is 6.5 for a 1.5-sd shock. Those numbers are in the ballpark of the frequency of financial crises in the US, as according to the Baron et al.'s (2021) list there were 4 financial crises in the US from 1930 to 2016. As for initial conditions, I assume that banks' wealth share η is initially close to its average. In the main model this number is 17.7%, and in the calibrated benchmark 17.3%, so I set $\eta_0 = 0.175$. Results are similar for any $\eta_0 \in [0.173, 0.177]$. Fig. 4 shows the results (I postpone the discussion of the dash-dotted curve for later).

Fig. 4a shows that a 2-sd (negative) shock in the main model generates output losses close to the numbers calculated using the data in Baron et al. (2021) and presented in Figure C.1 in Online Appendix C.1. The model without coordination failures, on the other hand, generates too little amplification overall. Since I am fixing the frequency of the shocks, the impulse responses already take into account that exogenous volatility σ is about 50% larger in the calibrated benchmark, so the shock simulated in that model is about 50% larger than in the main model, in absolute terms.

The shaded areas in Fig. 4a show 68% confidence bands, computed by estimating the 16th and 84th quantiles of the output loss experienced after the shock. In the benchmark, the effect of the shock is almost deterministic after it happens: Output will remain low, with some additional losses, but regardless of whether the economy experiences good or bad realizations of capital quality shocks dZ_t after the initial shock, there is not much variation in output losses (the upper and lower quantiles are very close). In the economy with coordination failures, the story is very different. A 2-sd shock throws the economy in a coordination trap, but what happens next depends a lot on the realization of subsequent uncertainty. If good shocks realize after the initial bad shock, the economy leaves the coordination trap soon, and output losses look like those in the model without coordination failures. However, if the economy further experiences bad shocks, output keeps falling, banks intermediate even less capital, asset prices and investment fall, and the economy gets stuck in the coordination trap, with very large output losses. This heterogeneous propagation of an initial negative shock is illustrated by the large distance between the lower and upper quantiles in the main model.

The message that the model with coordination failures can generate very heterogeneous real effects is reinforced in Fig. 4b, which reports the standard deviation of the output losses induced by the shock. This is much larger in the model with coordination failures. As suggested by the discussion in Online Appendix C.1 and its Figure C.1, in the data, a simple measure of output losses following a financial crisis can have a standard deviation larger than the mean. While the model with coordination failures still delivers a standard deviation below the mean for the shock considered, it is able to fill a large part of this gap (relative to the benchmark).

Now I ask: If we consider extremely large (and unlikely) shocks in the model without coordination failures, can we obtain a heterogeneous propagation close to the main model? I simulate a 4-sd negative shock in the calibrated benchmark, and the results are represented by the dash-dotted curves in Fig. 4. Comparing a 2-sd shock in the main model with a 4-sd shock in the benchmark is not a fair comparison, as 4-sd sized shocks are much more infrequent. Also, such shocks are inconsistent with the frequency of financial crises in the data. Still, this exercise serves to illustrate how difficult it is to explain the heterogeneous effects of financial crises without the mechanism proposed in this paper. As one can see, with a 4-sd shock, the calibrated benchmark roughly matches the average fall in output of the main model with a 2-sd shock, but is still far from it in terms of the heterogeneous propagation to the real sector, as illustrated by the confidence bands in Fig. 4a and the dash-dotted curve in Fig. 4b.

If one considers a 1.5-sd negative shock for both models, and also a 3-sd shock for the calibrated benchmark, the results are very similar (with the main message being actually reinforced), so I relegate the presentation of those additional results to Online Appendix B.2.

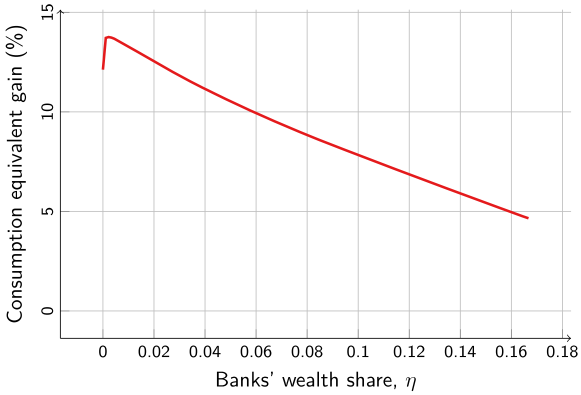

Fig. 5. Consumption equivalent gain of avoiding coordination failures.

I conclude this section with some remarks on the interpretation of the results. As shown, the model with coordination failures generates a lot of dispersion in output losses after bad shocks trigger a financial crisis, for shocks of a size consistent with the frequency of those events, and this is not the case once we take out of the picture the two-way feedback between coordination frictions and the financial block of the model. This high dispersion happens even though all shocks to banks' balance sheets are ultimately tied to the production side of the economy, as capital quality shocks are the unique exogenous source of volatility assumed. This is because an initial sequence of bad shocks throws the economy in a very fragile state, in which its fate (whether it will leave or not a coordination failure soon) is highly dependent on the shocks that hit afterwards.

Of course, one can think that in a model with more shocks, in which there are other sources of exogenous volatility, such heterogeneity in output losses could possibly also arise, although for more exogenous reasons. If there was another (non-fundamental) shock moving bank equity for reasons unrelated to the production side of the economy (say there are noise traders moving asset prices for non-fundamental motives, for instance), then perhaps output losses would be very dependent on what happens to those non-fundamental shocks after a sequence of bad shocks. But even then, it is not obvious that this would be the case, as adding other shocks could possibly reduce the variance of fundamental (capital quality) shocks needed to match the volatility of output. Part of the point of the exercise in this section is to show that the two-way feedback between banks' balance sheets and coordination can already generate a lot of dispersion in output losses, without the need to rely on extra sources of exogenous volatility.

6. Welfare and policy implications

I now discuss how coordination failures and different policies affect household welfare within the model, but before that, it is useful to discuss how welfare is computed. The welfare of the representative household is defined as

$$\mathcal{W}\left(K_t, \eta_t\right) = \mathbb{E}\left[\int_0^\infty e^{-\rho t}\left(C_t - \xi L_t\right) dt\right],$$

where C_t is the aggregate consumption. By market clearing, $C_t = Y_t - \iota_t K_t - x_t f_t$. Let Λ_t denote the household flow utility per unit of capital.[26] Using the firms equilibrium (Propositions 1 and 2), we can write Λ_t as a function of ψ_t and ι_t. From the banks and households equilibrium, ψ_t and ι_t are functions of η_t. Hence, we can write $\Lambda_t = \Lambda(\eta_t)$. Proposition A.2 in Online Appendix A.9 shows how to compute welfare as the solution of a boundary value problem.

6.1. *The welfare costs of deep financial crises*

I have shown that the interaction between coordination and financial frictions creates a powerful propagation mechanism *during* financial crises. Hence, by now it should not be a surprise that coordination failures strongly affect household welfare during financial crises, as they lead to much larger output losses following a sequence of bad shocks. However, as the discussion in Section 5.2 shows, the possibility of the economy entering a coordination trap affects the economy not only during those episodes, but also in good times, when coordination failures are very unlikely. For instance, if the economy is at the reflecting boundary η^*, coordination failures are expected to happen in 78 years, but still asset prices are considerably lower relative to the scenario where they never happen (see Fig. 1f). This section investigates how much coordination failures affect household welfare, both in good and bad times.

Fig. 5 reports the consumption-equivalent gain of avoiding coordination failures. That is, it compares welfare in the economy where firms always coordinate ($u_L = 0$) with the main (calibrated) model. As the figure shows, even in very good times, when coordination failures are not expected for a long time (78 years at η^*), they still cause welfare losses close to 5% in consumption

[26] Using the firms equilibrium and the market-clearing condition $C_t = Y_t - \iota_t K_t - f_t$, one can easily verify that the flow utility scales with capital.

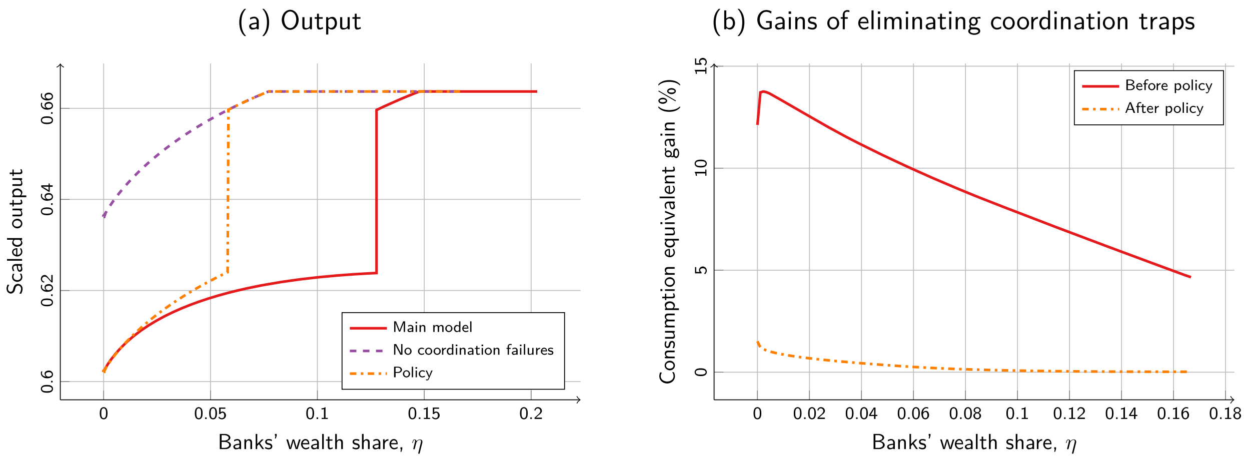

Fig. 6. Output and welfare gains before and after the policy.

equivalent units. This is because coordination failures affect asset prices, and therefore economic growth, even in good times, for reasons previously discussed in Section 5.2.

In what follows I study a policy that shuts down the two-way feedback between coordination and banks' balance sheets by subsidizing capital intermediation in bad times, ensuring that the rental rate is not affected by coordination failures. The exercise shows that the gains of avoiding coordination traps almost vanish once the two-way feedback is broken, although coordination failures still happen with non-negligible probability after the policy. Hence, what explains the large welfare costs of coordination failures is not their direct effect on output *per se*, but their effects on banks' behavior, which harms asset prices and output growth.

6.1.1. Breaking the two-way feedback

As shown in Section 6.1, the two-way feedback between coordination and banks' balance sheets can lead to very large welfare losses. In this section, I ask how much of those losses remain if the two-way feedback is eliminated, say, by protecting balance sheets during coordination traps. If the demand-driven coordination mechanism did not interact much with the financial block of the model, the losses induced by coordination traps should still remain large when banks are insulated from it.

More precisely, I consider a policy that guarantees that banks and households receive the rental rates $\overline{R}$ and $\beta\overline{R}$ (the equilibrium rental rates in the high regime), respectively, regardless of what happens in the real sector of the economy. This can be implemented by imposing a lump-sum tax on households and paying a subsidy per unit of capital lent to banks and households as to guarantee their target rental rate. One should think of this policy as a way to insulate banks and households from coordination traps: The banks and households equilibrium will be computed as if coordination failures never happened. Asset prices and banks' balance sheets will be identical to the benchmark with no coordination failures ($u_L = 0$).[27]

While this policy ensures the financial block of the model is not affected by coordination failures, in equilibrium they still happen, as shown in Fig. 6a, which compares the evolution of output after the policy in the main model and in the economy without coordination failures. Given that the preemptive behavior of banks is eliminated, coordination failures become less frequent: The expected time to reach a coordination failure in the reflecting boundary η^* is 314 years (as opposed to 78 years in the main model). Their unconditional probability goes from 11.5% to 4.7%, but it is still far from negligible.

Fig. 6b shows that the welfare costs of coordination failures almost disappear when the two-way feedback is shut down, even though their frequency has not vanished. For most initial conditions, the welfare gain of avoiding coordination failures is virtually zero, being still small even in cases where the economy starts off deep into a coordination trap. Therefore, what explains the large welfare losses of coordination failures is not their direct impact on output, but their effects on banks' behavior, which depresses asset prices and growth.

6.2. Other welfare and policy implications

6.2.1. Leverage constraints

Section 6.1.1 studied a policy that insulated banks and households from coordination failures by guaranteeing that the rental rate they receive from renting capital is not affected by those events. From a theoretical perspective, that exercise had not only a normative but also a positive goal, as it allowed us to understand how much of the welfare losses from coordination failures were driven by its two-way feedback with the financial block of the model. However, one could argue that, in practice, implementing such

[27] While the main objective of this exercise is to shut down the two-way feedback, and the policy is modeled as a simple subsidy to capital holdings (financed through lump-sum taxation), one can think of this policy more broadly as different measures taken by authorities to incentivize intermediation during financial crises, such as asset purchase programs (e.g., the Troubled Asset Relief Program), equity injections or a "Fed Put" (as in Chen and Phelan, 2023).

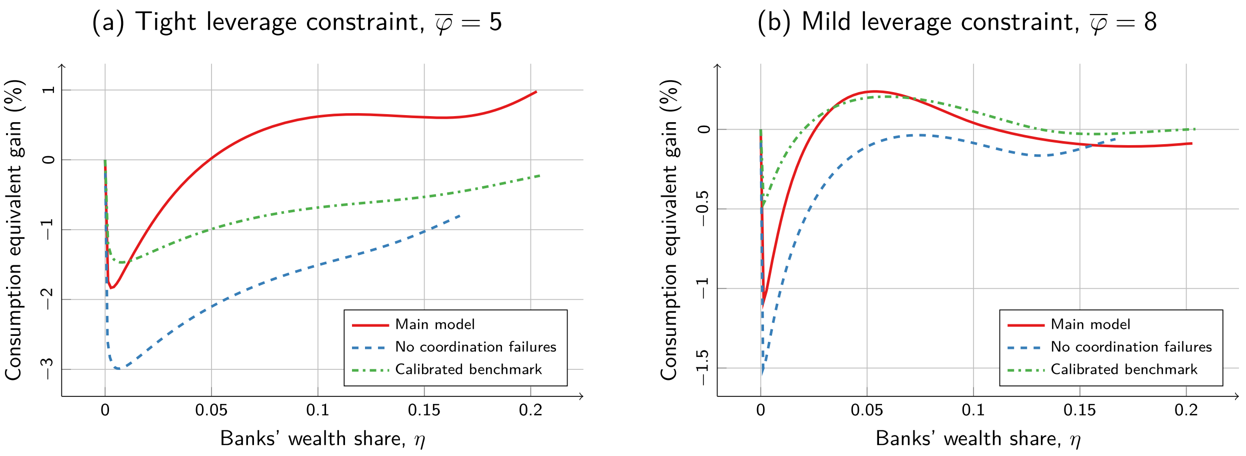

Fig. 7. Consumption equivalent welfare gain of implementing leverage constraint.

policy may be difficult, as it requires the policymaker to perfectly observe the rental rate earned by households and banks (beyond the fact that the subsidies are costly). In this section, I analyze within my model a policy that is commonly found in the regulatory toolkit: leverage constraints.

I consider a regulation that constraints the amount of leverage banks can have, as in Phelan (2016). In particular, I say that whenever leverage constraints are in place, each bank, individually, must satisfy the following constraint:

$$\varphi_t^b \leq \overline{\varphi}, \tag{19}$$

where $\overline{\varphi}$ is the maximum leverage allowed. For computational reasons, I assume that the policymaker lifts this restriction whenever banks' wealth share drops beyond some level $\overline{\eta}_{\text{lift}}$. However, in my computations I make it a very small number, always assuming that $\overline{\eta}_{\text{lift}} = 10^{-5}$ (results are not sensitive to this number, as long as it is sufficiently small).[28] One can think that there is some small cost associated with maintaining the institutional apparatus to enforce leverage constraints, and that the policymaker finds it optimal to cease the regulations when the banking sector becomes too small. The problem of banks is the same as before, except that now restriction (19) must be satisfied whenever $\eta > \overline{\eta}_{\text{lift}}$. Online Appendix A.10 discusses how the model is solved.

In the model with coordination failures, leverage is equal to 6.28 on average. I then simulate my model with leverage constraints (LC) in two scenarios, with leverage constraints above and below 6.28: $\overline{\varphi} = 5$ (tight leverage constraints) and $\overline{\varphi} = 8$ (mild leverage constraints).[29] I simulate this policy using the parameters of the main calibration (with coordination failures), the "no coordination failures benchmark" (in which coordination failures are shut down in the main calibration) and the "calibrated benchmark" (in which coordination failures are shut down and the model is re-calibrated).

Fig. 7 shows the consumption equivalent gains obtained from imposing leverage constraints, as a function of the state. As one can see, tight leverage constraints worsen welfare without coordination failures, which lends support to the claim of Brunnermeier and Sannikov (2014) that "*crude bounds on leverage are often counterproductive and reduce welfare*" in their setting. However, with coordination failures, tight leverage constraints can actually improve welfare in most regions of the state space, leading to consumption equivalent gains of almost 1% close to the stochastic steady state. This goes in line with the results obtained by Phelan (2016), who also finds that leverage constraints can increase welfare in a different setting.[30]

To better understand the mechanism through which leverage constraints improve welfare, Fig. 8 reports the outcome of some equilibrium variables with and without leverage constraints for the main model. Starting with the case of tight leverage constraints, one can see that those yield lower asset prices, for a *given* state, which suggests that such tight leverage constraints could be harmful to growth and welfare. However, as shown in Fig. 8c, tight leverage constraints strongly reduce the probability of bad states, as the probability of η being below $\hat{\eta}$ decreases with leverage constraints (as indicated by intersection of the vertical lines with the CDF). Avoiding such states is particularly important in the model with coordination traps, and imposing leverage constraints increases welfare at the stochastic steady state.

Intuitively, and as emphasized by Phelan (2016), leverage constraints have two effects on equilibrium outcomes. On one hand, by preventing banks from intermediating more capital, they can worsen the productive efficiency of capital. On the other hand, by preventing banks from facing large losses when negative shocks hit, they may reduce the probability of states in which banks'

[28] This is a sufficient (and probably not necessary) condition for the boundary conditions derived in the main model to continue to hold in the model with leverage constraints.

[29] I have explored different definitions of tight and mild constraints, and the main results were similar whenever the numerical algorithm used to compute the equilibrium converged.

[30] In Phelan (2016), there are two goods, and households are better at producing one good and banks are better at producing the other. Here, households are always worse than banks at intermediating capital, but there are coordination failures. Hence, the two models depart from Brunnermeier and Sannikov (2014) in different ways, but it turns out that both ways open room for leverage regulation to improve welfare.

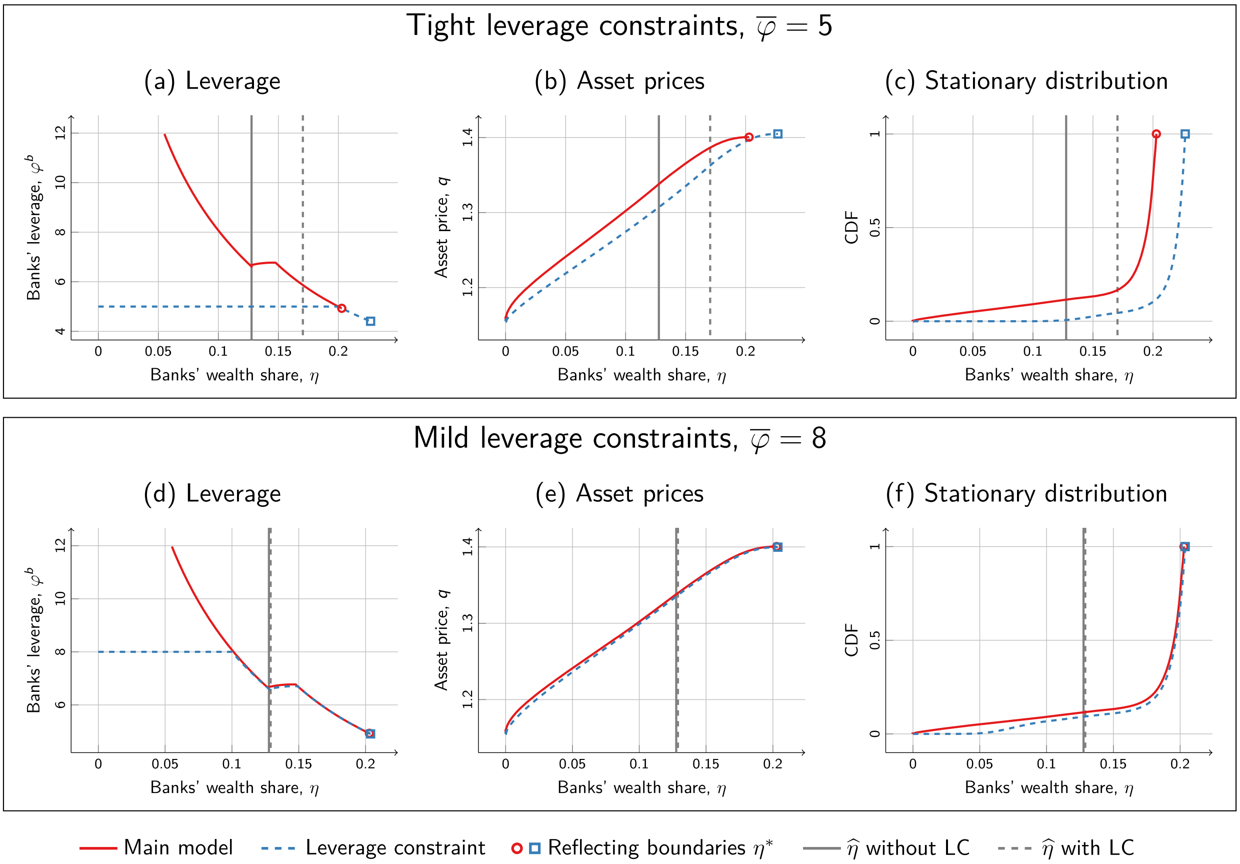

Fig. 8. Equilibrium functions for the model with and without leverage constraints.

net worth is very low. In a model with coordination failures, preventing states in which banks become very under-capitalized is particularly important, as those induce large output losses. This helps explain why tight leverage constraints are beneficial in the economy with coordination failures but are not in the economy without those.

As shown in the lower panels of Fig. 8, less strict leverage constraints are not effective in preventing the most relevant bad states (the probability of coordination failures is practically unchanged by the implementation of leverage constraints), since they do not bind before coordination failures hit, only after. This explains why they tend to worsen welfare around the stochastic steady state. They can still improve welfare if the economy starts off in very bad states, as they help prevent scenarios in which banks are almost wiped out (that is, states in which η is close to zero).

6.2.2. *Financial innovation*

Financial innovation, such as the emergency of fintechs, can create better ways through which non-traditional intermediaries can channel their funds to productive activities. To provide one example, the emergence of equity crowdfunding and peer-to-peer lending may allow non-traditional investors to allocate their funds to firms more efficiently (Cong and Xiao, 2024). In my model, this type of financial innovation can be captured by an increase in the efficiency of direct finance β. Taking it to an extreme, if $\beta = 1$, financial frictions would become irrelevant: There would be no reason for banks to exist in the model as households would be equally efficient at intermediating capital, but would not be subject to financial frictions.

An interesting finding of Phelan (2016) is that when households get better at intermediating capital, their welfare sometimes decreases in models with balance sheet recessions (the "intermediation paradox"). This is because, as households improve their intermediation skills, banks earns less excess returns on capital, making the financial sector less resilient and possibly increasing the probability of states in which banks are under-capitalized. In this section 1 ask whether increasing β is good or bad, and argue that coordination failures contribute to make financial innovation more destabilizing for the economy as a whole.

The effect of increasing household efficiency on banks' capital returns is not obvious in my setting. All else constant, an increase in β makes banks earn a lower excess return for a given level of "firm efficiency" (capital utilization). On the other hand, it affects the coordination failure cutoff ψ_{GG}, and coordination failures affect banks' capital returns. Using (9), we get

$$\frac{\partial \psi_{GG}}{\partial \beta} = -\frac{1 - \psi_{GG}}{1 - \beta},$$

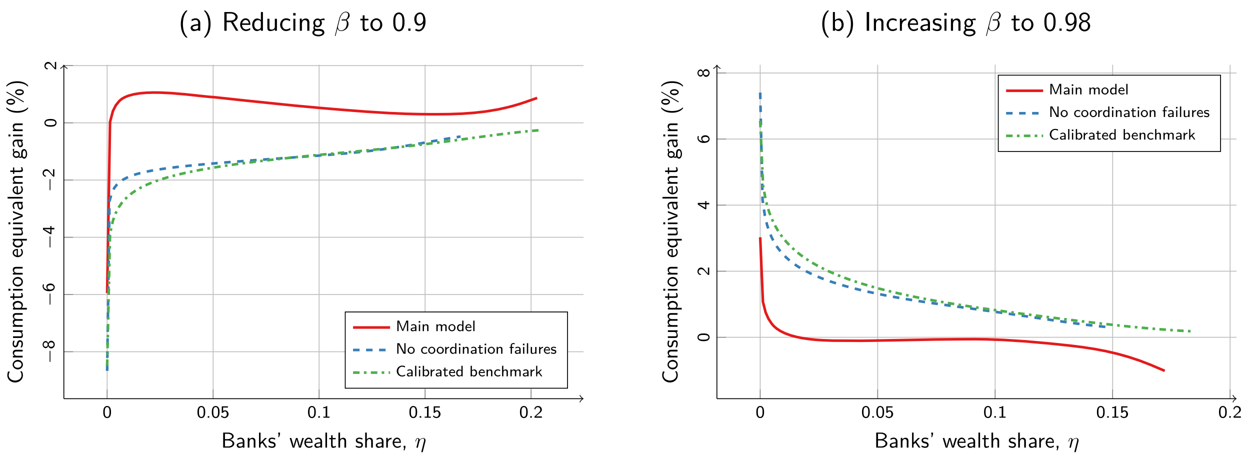

Fig. 9. Welfare gains of increasing and reducing β.

which is negative, since $\psi_{GG} \in (0, 1)$, given the assumption that dominance regions exist. That is, when households intermediate capital more efficiently, coordination failures become less prevalent, for *given* beliefs about bank intermediation. However, in my model, banks' intermediation decisions are endogenous and whether this lower cutoff will translate into a higher or lower probability of coordination failures will depend on how banks change their risk taking decisions, which in turn depends, among many things, on how their capital returns are affected by the change in β.

Fig. 9 presents the results for an increase and decrease of β, relative to its benchmark value of 0.959, for the main model, the "calibrated benchmark" and the "no coordination failures benchmark". As one can see, in this example, the intermediation paradox does not arise in the models without coordination failures, but it does arise in the model with coordination failures. The main reason is that, by increasing expected capital returns throughout the state-space, a lower β helps banks to maintain a healthy balance sheet after bad shocks, which is crucial to avoid a coordination failure.[31],[32]

Interestingly, in the numerical examples presented, the probability of a coordination failure increases with β, even though ψ_{GG} decreases with β.[33] This would not happen if ψ were an exogenous fundamental, as is usually the case in global games models, since in that case any decrease in ψ_{GG} would reduce the probability of the fundamental falling below that cutoff. However, in my setting the "fundamental" of the coordination game is endogenous: When β moves, ψ_{GG} moves, which, together with the fact that banks' capital returns change, makes banks adjust their capital holdings decisions, changing the distribution of the "endogenous fundamental" ψ. Hence, my global game with an endogenous fundamental leads to comparative statics that are very different from what one would obtain in standard global games settings.

Of course, a lower β has costs. In the event that the economy ends up in a situation where banks cannot intermediate all capital, the instantaneous output losses will be higher the lower β. In the economy with coordination failures, the lower probability of coordination failure events outweighs this cost in the proposed numerical example, but this is not always the case. If β falls sufficiently, welfare falls in almost all regions of the state space for the main model, but numerical computations indicate that β has to fall quite significantly (below 0.5) to get such result.[34]

7. Final remarks

This paper proposes a model to rationalize why some financial crises lead to sudden and persistent collapses in economic activity and others have mild real effects. There is an endogenous two-way feedback between banks' balance sheets and demand-driven coordination failures on the real sector. Low levels of bank lending reduce expectations about aggregate demand, which may cause firms to fail to coordinate on high economic activity, operating below full capacity. When firms operate with idle capacity, they demand less capital, affecting equilibrium capital returns and feeding back negatively into banks' balance sheets, further reinforcing what causes a coordination failure. Deep financial crises arise from a demand-driven coordination failure caused by insufficient bank lending.

The model can help explain why financial crises feature very heterogeneous real effects during those events and in its aftermath. Also, it is shown that deep financial crises have significant negative effects on the economy, not only during crises, but also in good times, when such episodes are not expected for a long time.

[31] The effect of β on expected capital returns is illustrated in Figure B.2 in Online Appendix B.2.

[32] Note that increasing β or shutting down coordination failures are two measures that reduce output losses in states where banks intermediate little capital (low ψ), but they have different welfare implications. This is in part explained by the impact of β on banks' capital returns already discussed.

[33] The unconditional probability of $\eta < \hat{\eta}$ with $\beta = 0.9$, $\beta = 0.959$ and $\beta = 0.98$ are 4.1%, 11.5% and 26.9%, respectively.

[34] Similarly, welfare should increase for a large enough increase in β, as setting $\beta = 1$ is equivalent to removing financial and coordination frictions at the same time (households would intermediate all capital in that case, but they would do that as efficiently as banks).

Data availability

I have shared replication files to be uploaded to the RED website.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.red.2024.101236.

References

Angeletos, G.-M., La'O, J., 2010. Noisy Business Cycles. NBER Macroeconomics Annual 2009, vol. 24. University of Chicago Press, pp. 319–378.
Angeletos, G.-M., La'O, J., 2013. Sentiments. Econometrica 81 (2), 739–779.
Baron, M., Verner, E., Xiong, W., 2021. Banking crises without panics. Q. J. Econ. 136 (1), 51–113.
Basu, S., 2019. Are price-cost markups rising in the United States? A discussion of the evidence. J. Econ. Perspect. 33 (3), 3–22.
Benhabib, J., Farmer, R., 1994. Indeterminacy and increasing returns. J. Econ. Theory 63 (1), 19–41.
Benhabib, J., Schmitt-Grohé, S., Uribe, M., 2001. The perils of Taylor rules. J. Econ. Theory 96 (1), 40–69.
Benigno, G., Fornaro, L., 2016. Stagnation traps.
Bernanke, B., Gertler, M., 1989. Agency costs, net worth, and business fluctuations. Am. Econ. Rev., 14–31.
Boissay, F., Collard, F., Smets, F., 2016. Booms and banking crises. J. Polit. Econ. 124 (2), 489–538.
Bolton, P., Chen, H., Wang, N., 2011. A unified theory of Tobin's q, corporate investment, financing, and risk management. J. Finance 66 (5), 1545–1578.
Brunnermeier, M., Eisenbach, T., Sannikov, Y., 2013. A Survey. Advances in Economics and Econometrics: Volume 2, Applied Economics: Tenth World Congress, vol. 50. Cambridge University Press, p. 3.
Brunnermeier, M., Sannikov, Y., 2014. A macroeconomic model with a financial sector. Am. Econ. Rev. 104 (2), 379–421.
Caballero, R.J., Engel, E., Haltiwanger, J.C., 1995. Aggregate employment dynamics: Building from microeconomic evidence.
Chamley, C., 2014. When demand creates its own supply: saving traps. Rev. Econ. Stud. 81 (2), 651–680.
Chen, W., Phelan, G., 2023. Should monetary policy target financial stability? Rev. Econ. Dyn. 49, 181–200.
Clymo, A., 2016. Growth, business cycles, and the fear of financial crises. Manuscript.
Cong, L.W., Xiao, Y., 2024. Information cascades and threshold implementation: theory and an application to crowdfunding. J. Finance 79 (1), 579–629.
Cooper, R., Haltiwanger, J., 2006. On the nature of capital adjustment costs. Rev. Econ. Stud. 73 (3), 611–633.
Cooper, R., John, A., 1988. Coordinating coordination failures in Keynesian models. Q. J. Econ. 103 (3), 441–463.
Copeland, A., Hall, G., 2011. The response of prices, sales, and output to temporary changes in demand. J. Appl. Econom. 26 (2), 232–269.
Crane, L.D., Decker, R.A., Flaaen, A., Hamins-Puertolas, A., Kurz, C., 2020. Business exit during the covid-19 pandemic: Non-traditional measures in historical context.
Décamps, J.-P., Gryglewicz, S., Morellec, E., Villeneuve, S., 2016. Corporate policies with permanent and transitory shocks. Rev. Financ. Stud., hhw07878.
Fakos, A., Sakellaris, P., Tavares, T., 2018. Investment slumps during financial crises: The role of credit constraints. Manuscript.
Farmer, R., Guo, J.-T., 1994. Real business cycles and the animal spirits hypothesis. J. Econ. Theory 63 (1), 42–72.
Fernández-Villaverde, J., Hurtado, S., Nuno, G., 2022. Financial frictions and the wealth distribution. Manuscript.
Gertler, M., Kiyotaki, N., 2015. Banking, liquidity, and bank runs in an infinite horizon economy. Am. Econ. Rev. 105 (7), 2011–2043.
Gertler, M., Kiyotaki, N., Prestipino, A., 2016. Wholesale banking and bank runs in macroeconomic modeling of financial crises. Handb. Macroecon. 2, 1345–1425.
Gertler, M., Kiyotaki, N., Prestipino, A., 2020. A macroeconomic model with financial panics. Rev. Econ. Stud. 87 (1), 240–288.
Gilchrist, S., Himmelberg, C.P., 1995. Evidence on the role of cash flow for investment. J. Monet. Econ. 36 (3), 541–572.
Gorton, G., Ordonez, G., 2014. Collateral crises. Am. Econ. Rev. 104 (2), 343–378.
Gorton, G., Ordonez, G., 2020. Good booms, bad booms. J. Eur. Econ. Assoc. 18 (2), 618–665.
Guimaraes, B., Machado, C., 2018. Dynamic coordination and the optimal stimulus policies. Econ. J. 128 (615), 2785–2811.
Guimaraes, B., Machado, C., Ribeiro, M., 2016. A model of the confidence channel of fiscal policy. J. Money Credit Bank. 48, 1363–1395.
Hall, G.J., 2000. Non-convex costs and capital utilization: a study of production scheduling at automobile assembly plants. J. Monet. Econ. 45 (3), 681–716.
He, Z., Krishnamurthy, A., 2011. A model of capital and crises. Rev. Econ. Stud., rdr 036.
He, Z., Krishnamurthy, A., 2019. A macroeconomic framework for quantifying systemic risk. Am. Econ. J. Macroecon. 11 (4), 1–37.
Isohätälä, J., Klimenko, N., Milne, A., 2016. Post-crisis macrofinancial modeling: continuous time approaches. In: The Handbook of Post Crisis Financial Modeling. Springer, pp. 235–282.
Kiyotaki, N., 1988. Multiple expectational equilibria under monopolistic competition. Q. J. Econ. 103 (4), 695–713.
Kiyotaki, N., Moore, J., 1997. Credit cycles. J. Polit. Econ. 105 (2), 211–248.
Morris, S., Shin, H.S., 2001. Rethinking Multiple Equilibria in Macroeconomic Modeling. NBER Macroeconomics Annual, vol. 15. MIT Press, 2000, pp. 139–182.
Morris, S., Shin, H.S., 2003. Global games: theory and applications. In: Advances in Economics and Econometrics, Theory and Applications, Eighth World Congress, vol. 1. Dewatripont, M., Hansen, L., Turnovsky, S. (Eds.). Cambridge University Press.
Murphy, K., Shleifer, A., Vishny, R., 1989. Industrialization and the big push. J. Polit. Econ., 1003–1026.
Phelan, G., 2016. Financial intermediation, leverage, and macroeconomic instability. Am. Econ. J. Macroecon. 8 (4), 199–224.
Sannikov, Y., 2013. Dynamic security design and corporate financing. Handb. Econ. Finance 2, 71–122.
Schaal, E., Taschereau-Dumouchel, M., 2018. Coordinating business cycles. Manuscript.
Schaal, E., Taschereau-Dumouchel, M., 2019. Aggregate demand and the dynamics of unemployment. Manuscript.