Risk to Control Risk

Fernando Mendo†
Princeton University

This version: November 19, 2018

Download latest version here

Abstract

Can low risk be risky? In this continuous-time macroeconomic model, low measured volatility may disguise the buildup of hidden systemic run risk. Runs trigger large drops in asset prices and real production and propel the economy into a highly unstable crisis regime. Agents take on too much leverage in the hidden risk regime because they disregard their contribution to the economy’s exposure to systemic runs. Surprisingly, a leverage cap can increase hidden run risk by deepening crises that follow runs. Economies exposed to less volatile fluctuations due to real shocks are more prone to systemic runs: stability breeds instability. This trade-off suggests that full stabilization of business cycles is not necessarily optimal once runs are taken into account. Instead, policymakers may consider allowing risk to control risk.

JEL classification: E13, E32, E44, G01, G12, G21, G33.

*I am deeply indebted to Markus Brunnermeier for his invaluable guidance and support, as well as to Nobuhiro Kiyotaki and Mark Aguiar for many helpful conversations and advice. I thank Joseph Abadi for detailed comments and corrections on preliminary versions of the manuscript. I also thank Oleg Itskhoki, Moritz Lenel, Chris Sims, Césaire Meh, Sebastien Merkel, Motohiro Yogo, Paul Ho, Julius Vutz, Nick Huang, Yann Koby, Ezra Oberfield, Simon Schmickler, and seminar participants at Princeton University and at the 2018 Macro Financial Modeling Summer Meeting for their comments and suggestions. Financial support of the Alfred P. Sloan Foundation and the Macro Financial Modeling (MFM) project is gratefully acknowledged.

†Email: fmendo@princeton.edu.
1 Introduction

Fighting fire with fire, a principle effectively exploited in fields such as medicine (vaccines) or firefighting (controlled burns), can be the key to controlling risk in economics. The hypothesis that prolonged periods of stability incite the buildup of hidden risks is a first-order concern among policymakers and academics, yet there is no framework that illustrates the internal consistency of this concern or the mechanism underlying it. Such a framework is necessary to incorporate its effects into risk measurement and policy design.

The seemingly paradoxical idea that stability breeds instability is widespread in policy and academic discussions. However, we have limited knowledge about the interactions between the different forms of risk. Several questions regarding these interactions remain unresolved. Do prolonged periods of macroeconomic stability (e.g. the Great Moderation) disguise the risk of sudden collapses such as the last financial crisis? If so, how can we identify this vulnerability in seemingly safe low-volatility environments? Is there a trade-off between business cycle stabilization and financial stability, i.e., do less volatile real fluctuations foster financial fragility? How effective are macroprudential policies to simultaneously control different forms of risk?

To study these questions, I set up a continuous-time macro model with risk-averse banks and households. Banks obtain higher returns from managing capital than households and these returns are subject to small exogenous shocks to capital quality which are the only real shocks in the economy. There are no equity markets, so banks finance capital holdings beyond their net worth by issuing short-term non-contingent debt (deposits) to households. Households occasionally consider whether or not to roll over deposits (according to a sunspot that follows a Poisson process). If banks are unable to meet their deposit obligations by liquidating capital, households refuse to roll over their deposits, triggering a systemic run. Bank runs wipe out large parts of the banking system and trigger large collapses in asset prices and output.

The economy exhibits two endogenous instabilities: 1) amplification of real shocks on output and asset prices, and 2) exposure to systemic runs. The main contribution of this paper is to include both instabilities in a tractable framework in order to characterize their joint dynamics and study how they respond to changes in the underlying volatility of real shocks.

The presence and intensity of the endogenous instabilities vary according to the key endogenous state variable: the ratio of the banking sector’s wealth to the aggregate value of capital (or banks’ wealth share). In particular, the joint dynamics unveil three different risk regimes: 1) a highly unstable crisis regime characterized by the amplification of real shocks and occasional runs, 2) a safe regime with no endogenous
risks, and 3) a *hidden risk regime* in which real shocks are not amplified but systemic runs can occur.

The crisis regime corresponds to situations in which the wealth share of the banking sector is low enough to prevent banks from holding all capital in the economy. This misallocation depresses output and asset prices. Managing the entire stock of capital would require such high leverage and risk exposure that banks refrain from doing so despite their productivity advantage. Then, banks’ capital demand depends on their risk-bearing capacity, which in turn depends on their wealth. Real shocks are amplified in this regime precisely because of their impact on banks’ risk-bearing capacity, which translates into changes in allocative efficiency.

Within the crisis regime, the economy is exposed to systemic runs only if the banking sector’s wealth share is sufficiently large. For runs to be possible, the difference between contemporaneous asset prices and their liquidation value (i.e., their price just after a run when households manage almost all capital) needs to be large enough to imply bankruptcy. When banks’ wealth share is exceptionally low, this difference is minuscule, and the economy is shielded from systemic runs. This region with amplification risk but no run risk is the exact opposite of the hidden risk regime in terms of the endogenous instabilities present.

The safe regime is associated with situations in which the banking sector’s wealth share is sufficiently large to allow banks to absorb all capital in the economy and yet maintain low leverage. The absence of misallocation implies no amplification. The soundness of banks’ balance sheets prevents exposure to systemic runs in spite of the large difference between contemporaneous high asset prices and their liquidation value.

The hidden risk regime corresponds to an intermediate range of the banking sector’s wealth share. It is high enough to allow banks to hold the entire capital stock but low enough to require them to choose weak balance sheets, i.e., high leverage, in order to do so. The latter combined with the high asset prices implied by the absence of misallocation exposes the economy to systemic runs, which immediately propel the system into the unstable crisis regime.

Importantly, prior to the realization of a run, the hidden risk regime is observationally equivalent to the safe regime in terms of the volatility of real and financial variables. In this sense, run risk is hidden.\(^1\) This illustrates the potential fragility of low-volatility environments and highlights the necessity of a *risk topography* to identify risks. In particular, the model suggests the following warning signals of run risk when measured volatility is low: the joint presence of high asset prices and a highly (or, at least, \(^1\)Run risk is not hidden for agents in the economy.)
moderately) levered banking sector, high excess returns per unit of measured volatility, and a low risk-free rate.

Systemic instabilities are an ergodic phenomenon in this economy. That is, the risk regimes associated with amplification and systemic runs are frequently visited by the system. In fact, the two peaks of the bimodal stationary distribution of banks’ wealth share may lie within these risky regimes. The highest peak corresponds to the stochastic steady state, i.e., the balance point to which the system tends to return, and it can be located within the hidden risk regime. The economy spends prolonged periods in this regime because, in contrast to amplification risk, run risk has a limited impact on excess returns on capital over the deposit rate. Then, wealth accumulation by levered banks is not particularly accelerated in the hidden risk regime. The second peak corresponds to the situation that follows a systemic run, which is characterized by a severely undercapitalized banking sector, i.e., it lies within the crisis regime. This peak reflects that the initial phase of the recovery after a run is slow.

A comparison across economies with different levels of exogenous risk (fundamental volatility of real shocks) unveils the key insight of the model: stability breeds instability or its contrapositive, risk controls risk. Economies with the lowest exogenous risk have the most severe endogenous instabilities. Lower fundamental risk prompts banks to issue more debt in order to expand capital holdings. This translates into a more efficient allocation of resources and therefore into higher asset prices. In other words, lower exogenous risk fuels high leverage and asset prices, which lead to both greater amplification risk and greater run risk.

Crucially, the increase in amplification risk is not enough to overcome the initial reduction of exogenous risk, so total volatility due to real shocks decreases alongside exogenous risk. Then, the main risk for economies with low fundamental volatility of real shocks is an elevated exposure to systemic runs. On the other end of the spectrum, economies exhibiting sufficiently high exogenous risk are completely shielded from run risk, i.e., risk controls risk. The weak reaction of amplification risk follows from the presence of an automatic stabilization mechanism, which is absent in the case of run risk. An increase in amplification risk feeds back into more prudent equilibrium portfolios for banks, while larger exposure to systemic runs has no effect on banks’ equilibrium leverage. This is ultimately related to the fact that banks disregard their contribution to run risk when making portfolio decisions (run externality).

An important insight is that the welfare ranking with respect to exogenous risk is non-monotonic. The key effect driving the result is that lower exogenous risk induces higher bank leverage (capital holdings). This has two opposing effects on welfare: 1) reduced misallocation, and 2) increased exposure to endogenous instabilities. Banks fail to adequately internalize these forces, so laissez-faire leverage can be excessive or overly
restrained with respect to its efficient level. The total effect on welfare depends on the relative strengths of externalities affecting banks’ portfolio decisions.

If the externality associated with misallocation is stronger than those related to instabilities, then banks’ capital holdings are inefficiently restrained, and a marginal reduction of exogenous risk helps minimize this inefficiency. Otherwise, a marginal reduction of exogenous risk aggravates the economy’s inefficiencies. Numerical results suggest that the run externality is the decisive factor in this comparison between externalities. This is the key behind the non-monotonicity since run risk is present only in economies where exogenous risk is sufficiently low.

Regarding policy interventions, the model reveals that a constant leverage constraint may fail to control endogenous instabilities. A leverage cap depresses the liquidation price of capital by deepening the recession that follows a run. This fuels run risk and is particularly detrimental for the economy since leverage is limited when it is needed the most. This force dampens the reduction of run vulnerability due to lower leverage in the crisis regime and, most importantly, it leads to greater exposure to runs in the hidden risk regime (unless the constraint is tight enough to completely rule out runs). A constant leverage constraint also increases amplification risk in situations where it is not binding and households still hold some capital. Despite these shortcomings, welfare analysis suggests that a constant leverage cap can benefit both agents as long as it reduces the economy’s overall exposure to runs.

The optimal state-contingent leverage cap overcomes the mentioned drawbacks by restricting banks’ portfolio decisions only when the economy would be exposed to runs, i.e., in an intermediate range of banks’ wealth share. Interestingly, a binding optimal leverage cap in the hidden risk regime can limit run risk but introduces amplification risk. This emphasizes the relative importance of the run externality.

I also explore a different approach to control run risk. In particular, I study how a benevolent policymaker would adjust the complete markets prescription of full stabilization when considering exposure to runs. The exercise ignores any cost of stabilization by assuming the policymaker can directly adjust fundamental volatility conditional on the strength of banks’ balance sheets. This allows me to focus on the optimal adjustment given the presence of run risk. Results suggest that it is optimal to allow risk associated with real shocks if the economy is exposed to runs and to implement complete stabilization if it is not. The exercise illustrates the direction in which the possibility of runs should affect stabilization policy; however, it is not a direct policy recommendation since the model does not capture known benefits of stabilizing the economy.

**Literature.** The idea that *stability breeds instability* has been discussed at least since Minsky [1986]. He argues that instability is a normal result of modern financial capital-
ism and casts doubts on the “fine-tuning” approach to policy. Minsky conjectures that even if policy does manage to achieve transitory stability, this stability would render financial markets complacent and susceptible to the reemergence of instability. This paper formalizes one such mechanism: a larger exposure to systemic runs.

The literature has extensively explored, separately, the two endogenous risks studied in this paper. In the case of amplification risk, the financial accelerator mechanism behind it was first discussed by Kiyotaki and Moore [1997] (hereafter KM), and Bernanke et al. [1999] (hereafter BGG). Later work by Brunnermeier and Sannikov [2014] (hereafter BruSan) identifies that the intensity of this risk varies in a highly nonlinear fashion along the business cycle. In the case of run risk, there is an ample literature studying bank runs following the seminal contribution of Diamond and Dybvig [1983]. This literature works with partial equilibrium or short-horizon macroeconomic frameworks, which are not suitable for studying endogenous risk dynamics.

Only after both risks arguably played an important role during the last financial crisis did economists start to explore them together within macroeconomic models. Prominent examples are Gertler and Kiyotaki [2015] (hereafter GK) and Gertler et al. [2017]. The main interest in these papers is to compare the propagation mechanisms of real shocks with and without bank runs. In contrast, this paper focuses on the interactions between these risks and their implications for macroeconomic stability. Moreover, while the literature concentrates on the economy’s response to shocks in some specific situations, this paper provides a full characterization of the endogenous risks necessary to uncover the different risk regimes.

The “volatility paradox” result in BruSan shares with this paper the general idea that low exogenous volatility of real shocks can fuel endogenous risks. However, their results differ from mine along several dimensions. First, their model predicts that every time the economy is subject to risks associated with the financial sector, the risks translate into observable volatility of financial and macroeconomic variables. That is, their model is unable to identify the buildup of hidden risks. Second, this paper raises much stronger concerns about low exogenous volatility. In BruSan, endogenous amplification persists as exogenous volatility vanishes, but this effect is not enough to increase overall volatility, so welfare improves with lower exogenous volatility. This paper shows that the welfare ranking reverses once run risk is considered since this risk is magnified as exogenous fluctuations are reduced.

The modeling approach of this paper builds on BruSan and GK. The continuous-time techniques used to characterize the dynamics of risks expand work by BruSan, while the approach to include systemic runs in an infinite-horizon economy follows GK. In an international context, Brunnermeier and Sannikov [2015] study amplification risk together with a different class of discrete collapses: sudden stops of international capital
flows. Their work is technically the closest to this paper since they also consider a continuous-time framework with self-fulfilling runs; yet they only study unanticipated runs, i.e., their analysis disregards any feedback effects from run risk to agents’ decisions or equilibrium dynamics. These feedback effects are key in this paper.

From a technical point of view, this paper expands the literature on continuous-time macro-finance models, e.g., BruSan and Hansen et al. [2018], by considering (anticipated) aggregate jump risk. This is necessary to include systemic runs, and it is technically challenging because aggregate jump risk breaks the key feature that makes continuous-time models tractable: their focus on local perturbations.

2 Model

2.1 Environment

Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space that satisfies the usual conditions, and assume all stochastic processes are adapted. The economy evolves in continuous time with \(t \in [0, \infty)\) and is populated by a continuum of banks and households as well as a government. There are two goods: the final consumption good and physical capital.

A. Technology

The production technology is linear and reflects that bankers are more efficient than households at managing capital. In particular, capital held by bankers produces

\[ y_t = a k_t \]

while households’ production function is \(y_t = a k_t\) with \(a > a^\prime\). In general, I use the underbar notation for parameters and variables associated with the household sector.

Physical capital evolves according to\(^2\)

\[ dk_t = (\Phi(\iota_t) - \delta) k_t dt + \sigma k_t dZ_t, \tag{1} \]

where \(\iota_t\) is the investment per unit of capital, \(\delta\) is the depreciation rate, \(\sigma > 0\) is the exogenous volatility of capital growth, and \(Z_t\) is a Brownian motion associated with capital quality shocks. The investment technology allows for transforming \(\iota_t k_t\) units of final goods into \(\Phi(\iota_t) k_t\) units of capital. Function \(\Phi\) represents technological illiquidity or adjustment costs and satisfies \(\Phi(0) = 0, \Phi'(0) = 1, \Phi'(\cdot) > 0, \text{ and } \Phi''(\cdot) < 0. \)

\(^2\)As is standard in the literature, this law of motion does not include capital purchases.
Brownian shock is the only source of exogenous aggregate uncertainty in the model. It affects capital quantity but not its productivity, which can be motivated by stochastic depreciation or news about capital quality. Further discussion follows in subsection 2.4.

B. Preferences

Preferences are symmetric for both types of agent. In particular, the utility that agents obtain from a stochastic consumption path \( \{c_t\}_{t \geq 0} \) is

\[
E_0 \left[ \int_0^\infty e^{-\rho t} \log(c_t) dt \right],
\]

where \( \rho \) is the preference discount rate.

C. Switching identities

Agents experience idiosyncratic Poisson shocks that switch their types, i.e., turn bankers into households and vice versa. Let \( \lambda \) be the intensity of the shock that turns bankers into households and \( \Lambda \) be the intensity of the inverse process. These dynamics prevent either sector from taking over the entire wealth of the economy indefinitely and therefore ensure the existence of a non-degenerate stationary distribution.

D. Markets and financial friction

There are markets for physical capital and final goods. Deposit contracts are the only financial assets in this economy. A deposit contract ensures the repayment of principal and interest as long as the borrower remains solvent, and it transforms into a claim on borrowers’ assets in the case of default. Note that deposits held at different banks are distinct assets because banks’ exposures to default and recovery values may differ. The absence of equity markets represents a financial friction that can be motivated by a “skin in the game” constraint.

Also, I assume that banks that underperform with respect to their peers incur in a prohibitively high cost \( \phi \) (measured in final goods). This allows me to focus on symmetric equilibriums among banks. Further discussion of these assumptions is presented in subsection 2.4.

E. Bank runs

Households are occasionally concerned about the soundness of banks. In particular, they “wake up” to consider whether or not to roll over their deposits according to a Poisson process \( \hat{J}_t \) with arrival rate \( \hat{p}_t \). The timing of events within the “\( dt \) period” is as follows.
1. Households decide whether or not to roll over their deposits. Banks sell their assets to repay deposits not rolled over. The liquidation price of capital is the one at which capital can be sold at the end of the $dt$ period.

2. Banks that are not able to fully repay deposits go into bankruptcy. Failed banks’ assets are liquidated and the resources are used to pay depositors. Depositors who decided not to roll over deposits are served first.\(^3\)

3. Failed banks receive a small transfer, which is financed by taxes on the household sector. This transfer allows them to resume operations.

4. Agents optimally rebalance their portfolios.

Whenever a bank run equilibrium is feasible, there is equilibrium multiplicity since the no-run equilibrium is always a possibility. I assume depositors coordinate on a bank run using a sunspot $\omega_t \in \{0, 1\}$. If $\omega_t = 1$, then the run equilibrium is realized. This can be interpreted as a panic among households. Since the no-run equilibrium is equivalent to households not evaluating the soundness of banks, the only relevant wake-up calls are the ones linked to panics, i.e., to a run equilibrium. I define the Poisson process $J_t$ associated with wake-up calls linked to panics as

$$d\tilde{J}_t = \mathbb{1}_{\{\omega_t = 1\}} d\tilde{J}_t$$

with arrival rate $\tilde{p}_t = \Pr(\omega_t = 1)\tilde{p}_t$. In general, the probability of coordinating on a run equilibrium $\Pr(\omega_t = 1)$ can be a function of endogenous variables. From now on, I work with Poisson process $\tilde{J}_t$ and later I discuss the equilibrium selection mechanism directly in terms of $\tilde{p}_t$.

\(F.\) Transfer policy

Let $T_t K_t$ be the total transfer from the household sector to banks that go bankrupt due to a systemic run, where $K_t$ is the aggregate capital in the economy and $T_t$ is chosen by the government. The transfer is financed via taxes levied on households, and the government runs a balanced budget at each point in time.

Taxes and transfers are proportional to the net worth agents had before the systemic run in order to preserve tractability. Denote $\tau_t$ as the transfer per unit of net worth to

---

\(^3\)If depositors recover the same value independently of their rollover decision, they are always indifferent as to whether banks fail (no loss on deposits’ value) or not (recovery rate depends on liquidation value). There are several (off-equilibrium) assumptions that break the indifference if the bank will fail. I assume sequential services in two stages: depositors who do not roll over deposits are paid before the rest. Alternatively, I could assume that banks can divert a fraction of deposits that are rolled over just before collapse.
banks and \( \tau \), the tax per unit of net worth levied on households. The transfer can be interpreted as a direct bailout of the banking sector or as recognition that banks are able to rescue some value after a collapse.

2.2 Agents’ problems

A. Debt rollover decision and run vulnerability

**Liquidation value.** The liquidation value of capital corresponds to the price of capital after a systemic run, i.e., the price when almost all capital is managed by unproductive households (except for the capital banks can acquire using the small transfer). I denote the liquidation value of capital in terms of the numeraire (final goods) as \( \tilde{q}_{t-} \).

I use notation \( z_{t-} \equiv \lim_{s \to t} z_{s} \) to refer to the value of variables just before a jump.\(^4\) The liquidation price is known before the uncertainty regarding the Poisson shock is unveiled. The liquidation price is equal to the capital price \( q_{t} \) if the systemic run is realized, i.e., \( \tilde{q}_{t-} = q_{t} \) if \( d\tilde{J}_{t} = 1 \) and the economy is vulnerable to runs (see condition below).

**Debt rollover decision.** If a bank is not able to meet its deposit obligations at the liquidation price for its assets, i.e.,

\[
\tilde{q}_{t-} k_{t-} < b_{t-},
\]

where \( b_{t-} \) is the value of deposits at the bank and \( k_{t-} \) are the banks’ capital holdings, then households decide not to roll over their deposits at this bank. In this case, households recover \( \tilde{q}_{t-} k_{t-}/b_{t-} \) per deposit. If households had rolled over deposits, they wouldn’t have been able to recover any value from their deposits since there are no resources left after serving depositors who decided not to roll over deposits. Importantly, households are only willing to pay the liquidation value \( \tilde{q}_{t-} \) to banks selling capital to meet their deposit obligations. If this run condition is not met, then households roll over their deposits in this bank after the wake-up call.

The run condition can also be written in terms of the percentage loss on deposits in case of a run, i.e.,

\[
\ell_{t-}^{d}(k_{t-}, b_{t-}) \equiv \left[ 1 - \frac{\tilde{q}_{t-} k_{t-}}{b_{t-}} \right]^{+} > 0,
\]

where \([\cdot]^{+} \equiv \max \{\cdot, 0\} \). I use \( \ell_{t-}^{d} \) to denote the loss on deposits evaluated at banks’ equilibrium portfolio \((k_{t}, b_{t})\). For convenience, also define the capital price drop following

\(^4\)This means \( z_{t-} \) is measurable with respect to the filtration associated with \( \{Z_{s}, J_{s} : s \in [0, t)\} \).

9
a systemic run as
\[ l^d_{t-} = 1 - \frac{\tilde{q}_t - q_t}{q_t}, \]
where \( q_t \) is the contemporaneous capital price.

**Run vulnerability.** The liquidation price \( \tilde{q}_t \) at which banks’ solvency is evaluated assumes that the entire banking sector fails (and resumes operation only with the net worth provided by the transfer). Therefore, consistency requires that this is indeed the case, i.e., that all banks collapse in case of a run. This condition can be characterized in terms of the aggregate percentage loss on deposits after a systemic run\(^5\)

\[ L^d_{t-} \equiv \left[ 1 - \frac{\tilde{q}_t - K^b_t}{B_t} \right]^+ > 0, \]

where \( B_t \) is the aggregate value of deposits at the banking sector and \( K^b_t \) represents banks’ aggregate capital holdings.\(^6\) Systemic runs occur if and only if there is a wake-up call linked to a panic \((d\tilde{J}_t = 1)\) and the banking sector is sufficiently vulnerable \((L^d_{t-} > 0)\).

**Equilibrium selection mechanism.** I assume

\[ \tilde{p}_{t-} = \Gamma(L^d_{t-}), \quad (3) \]

where \( \Gamma(.) \) is continuous and satisfies \( \Gamma(0) = 0 \) and \( \Gamma'(\cdot) \geq 0 \), i.e., households are more likely to coordinate on the bank run equilibrium if potential losses for them are higher. Note that \( L^d_{t-} \) is an aggregate variable, which cannot be influenced by any individual agent alone. The equilibrium selection mechanism chosen has a key role for welfare analysis and numerical exercises, but qualitative features about joint dynamics of instabilities will be independent from it. Further discussion about this assumption is presented in subsection 2.4.

**B. Households**

Households choose their consumption rate \( c_t \), investment rate \( \iota_t \), capital holdings \( k_t \), and deposits \( b_t \). Let \( n_t \) denote a household’s net worth, which is the only individual state variable for this problem.\(^7\) Also, let \( T \) be the stochastic arrival time of the idiosyncratic shock that turns a household into a bank. For a given \( n_s \), a household solves

---

\(^5\)Given that in equilibrium all banks will choose the same portfolio, the average loss will be equal to the individual loss of each agent.

\(^6\)The formal definition of aggregate variables is provided in subsection 2.3.

\(^7\)The shock structure (capital quality shocks instead of productivity shocks) and the linear production technology allow for reducing individual states from capital and deposits to only net worth.
\[ V_s(n_s) \equiv \max_{\{c_t, a_t \geq 0, k_t \geq 0, b_t \}} \mathbb{E}_s \left[ \int_s^T e^{-\rho(t-s)} \log(c_t) dt + V_T(n_T) \right] \tag{4} \]

s.t.

\[
dn_t = [(a - \lambda_t) b_t + r_t b_t - \zeta_t] dt + d(q_t k_t)|_{dJ_t=0} \\
- \mathbb{I}_{\{\mathcal{A}_t^d > 0\}} \left[ \mathbb{I}_{\{e_t^d - b_t \geq 0\}} (1 - \tau_{t-}) n_{t-} dJ_t + \mathbb{I}_{\{e_t^d - b_t \leq 0\}} \phi dt \right] \quad \forall t \leq T \\
n_t = q_t k_t - b_t,
\]

where \( V_t(.) \) is banks’ value function, defined below. The last term of the dynamic budget constraint represents the cost households bear in case of a systemic run: losses on deposits \( e_t^d b_t \), on capital value \( e_t^d q_t^d k_t \), and taxes levied to bail out banks \( \tau_t n_t \). The other terms are standard and correspond to production net of investment \( (a - \lambda_t) b_t \), return on deposits \( r_t b_t \), consumption expenditure \( c_t \) and capital gains absent runs \( d(q_t k_t)|_{dJ_t=0} \).

For simplicity, this formulation assumes that all banks choose the same portfolio and therefore there is a unique deposit contract available for households with return \( r_t \) conditional on no default. In general, deposits at banks with different portfolios render different returns, i.e., they are different assets, because the portfolio determines the recovery rate of deposits in case of a run.

### C. Banks

Banks choose their consumption rate \( c_t \), investment rate \( \iota_t \), capital holdings \( k_t \), and debt \( b_t \) (deposits). Let \( n_t \) denote the household’s net worth, which is the only individual state variable for this problem. Also, let \( T \) be the stochastic arrival time of the shock that turns the bank into a household. Given \( n_s \), a bank solves

\[ V_s(n_s) \equiv \max_{\{c_t, a_t \geq 0, k_t \geq 0, b_t \}} \mathbb{E}_s \left[ \int_s^T e^{-\rho(t-s)} \log(c_t) dt + V_T(n_T) \right] \tag{5} \]

s.t.

\[
dn_t = [(a - \lambda_t) k_t - r_t (k_t, b_t) b_t - c_t] dt + d(q_t k_t)|_{dJ_t=0} \\
- \mathbb{I}_{\{\mathcal{A}_t^d > 0\}} \left[ \mathbb{I}_{\{e_t^d (k_t - b_t) \geq 0\}} (1 - \tau_{t-}) n_{t-} dJ_t + \mathbb{I}_{\{e_t^d (k_t - b_t) \leq 0\}} \phi dt \right] \quad \forall t \leq T \\
n_t = q_t k_t - b_t
\]

The last term of the dynamic budget constraint is the one associated with runs, and it is relevant only if the economy is vulnerable to runs, i.e., \( \mathcal{A}_t^d > 0 \). In these situations,
the bank can choose a portfolio \((k_t, b_t)\) that implies bankruptcy and losses for depositors after a systemic run, i.e., \(t^d_t(k_t, b_t) > 0\). In this case, a run wipes out the bank’s net worth \(n_{t-}\) but allows it to receive a transfer \(\tau_{t-}\). Alternatively, the bank can choose a safe portfolio that allows it to avoid bankruptcy in case the run is realized \((t^d_t(k_t, b_t) \leq 0)\) and pay the underperforming cost \(\phi\). A safe portfolio implies underperforming with respect to banks choosing a risky portfolio and exposing the economy to a systemic run.

The other terms in the constraint are standard and analogous to that described for the household. The only difference is that the banks’ portfolio decision affects the rate paid on deposits \(r_t(k_t, b_t)\) because it influences the depositors’ loss rate in case of a run. This function needs to be consistent with households’ asset-pricing conditions.

### 2.3 Equilibrium

**Notation.** Denote the set of banks by the interval \(\llbracket 0, \nu_t \rrbracket\) and index individual banks by \(i \in \llbracket 0, \nu_t \rrbracket\). Similarly, let the set of households be the interval \(\llbracket \nu_t, 2 \rrbracket\) and index individual households by \(j \in \llbracket \nu_t, 2 \rrbracket\). Let \(\mathbb{I}\) be the interval including the index of all agents who are banks at some point and \(\mathbb{J}\) be the corresponding interval for households.\(^8\)

The following notation is necessary because agents alternate types. Let \(T^u_n\) be the \(n\)th point in time in which agent \(u \in [0, 2]\) experiences an idiosyncratic shock that changes his type. Define \(\mathbb{B}^u \equiv \{[T^u_{n-1}, T^u_n] : u \text{ is a banker}\}\) as the set of all intervals in which agent \(u\) is a banker, and denote an element of this set as \(B^u_m\). Similarly, define set \(\mathbb{H}^u\) with elements \(H^u_m\) to be the set of time intervals in which the agent is a household. I abuse notation and also use \(\mathbb{B}^u = \cup_m B^u_m\) and \(\mathbb{H}^u = \cup_m H^u_m\).

**Equilibrium definition.**\(^9\) Fix a transfer policy \(\{\mathcal{T}_t\}\). For any initial endowment of capital \(\{k^i_0, k^j_0 : i \in \mathbb{I}_0, j \in \mathbb{J}_0\}\) such that

\[
\int_{\mathbb{I}_0} k^i_t dt + \int_{\mathbb{J}_0} k^j_t dj = K_0
\]

an equilibrium is a set of stochastic functions\(^10\) on the filtered probability space defined

---

\(^8\)The threshold \(\nu_t\) captures the transition between types and satisfies \(d\nu_t = (\lambda(1 - \nu_t) - \lambda\nu_t) dt\) with \(\nu_0 = 1\). This representation assumes agents don’t know their own index so they cannot anticipate their switching time. It also assumes only net flows between agents type take place. Alternative representations are possible, but imply reassigning indexes to keep intervals connected.

\(^9\)The definition includes the symmetry assumptions embedded in agents’ problems. In particular, when defining the asset prices available to an agent type, it assumes a symmetric portfolio decision of the other one. This simplifies considerable notation and exposition.

\(^10\)Stochastic functions on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) are mappings \(X : S \times \Omega \to \mathbb{R}\) such that \(X(s, \cdot)\) is random variable \(\forall s \in S\). Stochastic processes are the special case where \(S = \mathbb{R}\), which is usually interpreted as time. In this definition, \(\{q_t, \bar{q}_t, p_t, c^i_t\}\) are stochastic processes, but other objects are associated with different definitions of \(S\): \([V_t(\cdot), \underline{V}_t(\cdot)]\) with \(S = \mathbb{R}^2\), \([r_t(\cdot)]\) with \(S = \mathbb{R}\), \([c^i_t, i^i_t, x^i_{t-}]\) with \(S = \mathbb{B}\), and \([\underline{c}^j_t, \underline{i}^j_t, x^j_{t-}]\) with \(S = \mathbb{H}\).
by aggregate shocks \( \{ Z_t, J_t : t \geq 0 \} \): capital price \( \{ q_t \} \) and liquidation value \( \{ \bar{q}_t \} \), deposit rate function \( \{ r_t(\cdot) \} \), losses function \( \{ \ell_t^i \} \), arrival intensity \( \{ \tilde{p}_t \} \), transfers and taxes \( \{ \tau_t, \Sigma_t \} \), households’ decisions \( \{ c_i^h, \bar{c}_i^h, k_i^h, b_i^h \} \in H^h \) and value function \( \{ V^h_i(\cdot) \} \) for \( j \in J \), and banks’ decisions \( \{ c_i^b, \bar{c}_i^b, k_i^b, b_i^b \} \in B_m \) and value function \( \{ V^b_i(\cdot) \} \) for \( i \in I \) such that

1. Initial net worths satisfy \( n_0^i = q_0k_0^i \) and \( n_0^j = q_0k_0^j \) for all \( i \in I \) and \( j \in J \).

2. Agents optimize

   (a) Households: Given stochastic functions \( \{ q_t, \bar{q}_t, r_t, \ell_t^i, \bar{p}_t, V^i_t(\cdot) \} \), decisions \( \{ c_i^h, \bar{c}_i^h, k_i^h, b_i^h \} \in H^h \) solve households’ problem for all \( j \in J \), \( H^h_m \in H^J \).

   (b) Banks: Given stochastic functions \( \{ q_t, \bar{q}_t, r_t(\cdot), \bar{p}_t, V^j_t(\cdot) \} \), decisions \( \{ c_i^b, \bar{c}_i^b, k_i^b, b_i^b \} \in B_m \) solve banks’ problem for all \( i \in I \), \( B^i_m \in B^J \).

3. Markets clear

   (a) Goods
   \[
   \int_{I_t} c_i^h dt + \int_{J_t} c_i^j dj + \phi \mathbb{1}_{\{ \mathcal{L}_t^h > 0 \}} \int_{I_t} \mathbb{1}_{c_i^j(\mathcal{L}_t^j, b_j)} < 0 \, dt = \int_{I_t} (a - \bar{c}_i^h) k_i^h dt + \int_{J_t} (a - \bar{c}_i^j) b_i^j dt
   \]

   (b) Capital
   \[
   \int_{I_t} k_i^h dt + \int_{J_t} b_i^j dt = K_t
   \]

   (c) Deposits
   \[
   \int_{I_t} b_i^h dt = \int_{J_t} b_i^j dt
   \]

4. The law of motion of aggregate capital is
   \[
   dK_t = \left( \int_{I_t} \Phi(\ell^i_t) k_i^h dt + \int_{J_t} \Phi(\ell^j_t) b_i^j dt - \delta K_t \right) dt + \sigma K_t dZ_t
   \]

5. Consistency conditions:
   - Value functions \( \{ V^i_t(\cdot), V^j_t(\cdot) \} \) are consistent with agents’ optimal decisions, i.e., they satisfy (4) and (5).
   - Deposit rate function \( r_t(k_t, b_t) \) is consistent with asset-pricing conditions of households.
   - Arrival intensity of Poisson process \( \{ \tilde{p}_t \} \) is consistent with the selection mechanism chosen.
   - Liquidation value \( \{ \bar{q}_t \} \) is consistent with the capital price process \( \{ q_t \} \), i.e. \( \bar{q}_t = q_t \) if \( \mathbb{1}_{\{ \mathcal{L}_t^{\bar{i}} > 0 \}} d\bar{J}_t = 1 \).
6. Total transfers are consistent with the government’s policy and its budget is balanced

\[ T_t K_t = \tau_t \int_{t_i}^{t_f} n_i^j di = \tau_t \int_{t_j}^{t_f} n_j^i dj \]

2.4 Discussion of assumptions

**Productivity difference.** Banks’ higher productivity captures the advantage that the banking sector has in allocating resources to productive agents. It summarizes the real contribution of the financial sector to the economy. Ideally, the banking sector would intermediate all resources to achieve the most efficient allocation.

**Capital quality shocks.** The model features shocks to the quantity of capital instead of the more standard productivity shocks. This assumption is made for tractability and is typical in the macro-finance literature. This formulation together with the linearity of the production function allows the economy to scale with the capital stock; i.e., it reduces the state space by eliminating the aggregate capital stock as a state variable. An interpretation is that capital is measured in efficiency units and the shocks are news about its quality. A positive capital quality shock behaves similarly to a very persistent positive productivity shock that induces a higher utilization rate.

**Financial friction.** The financial friction is that agents cannot issue equity to raise funds. This can be thought of as a limited case of a “skin in the game” constraint, which can be microfounded by an agency problem in which banks are able to divert a fraction of asset returns at the expense of households (see, for example, BruSan). Due to this constraint, markets are incomplete, so agents cannot write contracts conditional on the aggregate capital quality shock. Market incompleteness links the allocation of productive resources to the allocation of risk since agents need to bear the risk associated with capital in order to use it for production.

**Cost of underperforming the market.** If the economy is exposed to systemic runs, banks’ portfolio choices will have two local maxima: 1) one that exploits excess returns but exposes them to bankruptcy in case the run is realized, and 2) one that allows them to survive a run but forgo high contemporaneous excess returns. Given that systemic runs are rare events, consistently choosing the second investment strategy implies underperforming the market for a prolonged period, e.g., 40 years with baseline parameters. The cost of underperformance captures the fact that no asset manager can stay in his job long following such a strategy. Recent history is replete with examples of financial institutions that have chosen to underperform the market while waiting for a large collapse but were not able to sustain the strategy long enough, such as Tiger Management funds\(^\text{11}\) during the dot-com bubble.

\(^{11}\)In March 2000, Tiger Management LLC closed its six hedge funds after poor return performances
The underperformance cost does not force the economy to be exposed to systemic runs. It simply allows me to focus on symmetric equilibria among banks. Absent this cost, the economy would be exposed to runs in the same situations, but only a fraction of the banking sector would fail after a run.

Transfer policy. The transfer policy ensures that, after a bankruptcy, banks are able to recover some value at the expense of their creditors. This policy allows banks to be included in the welfare analysis since the obvious alternative is that, after a systemic run, some banks’ wealth drops to zero, leading to a minus infinity value for utility. A common solution is to assign an arbitrary utility value in these situations, e.g., assuming that bankers die and their continuation payoff is zero. This is not suitable with an endogenous arrival intensity of the death event, since the arbitrary value influences the preference ranking between situations with different exposures to such a event. Alternatively, one can focus on households for welfare analysis and disregard the transfer policy, i.e., set $\mathcal{T} = 0.$

Equilibrium selection mechanism. The dependence of the selection mechanism on the aggregate average loss of deposits relates the fundamental that determines the existence of a run equilibrium to the likelihood such an equilibrium would be realized. This follows the spirit of global games which application would deliver the probability of a run as a function of the relevant fundamental without the necessity of a sunspot. Importantly, for a systemic run to be self-fulfilling, it needs to generate a drop in asset prices, which cannot be influenced by individual decisions. This implies that the relevant fundamental needs to be an aggregate variable. The approach follows work by GK. Welfare results depend on this assumption, but qualitative results about joint dynamics of risk do not.

3 Recursive equilibrium solution

A. Aggregate state

As mentioned before, the shock structure and the linear production function eliminate capital $k$ as a state variable (scale invariant) for each individual household and bank. Therefore, the only individual state variable of an agent’s problem is his net worth. So, in general, the aggregate state of this economy will be the distribution of net worth. Fortunately, the state space of this economy can be simplified to a single aggregate state variable. First, all agents’ decisions will be linear in their net worth, which makes the

---

12 In this case, the off-equilibrium cost of underperforming would need to diverge toward infinity. 
13 A direct application of global games in this model is not pursued due to technical complications.

---

that followed from staying out of Internet and technology stocks. Julian H. Robertson, manager and co-founder, correctly anticipated the collapse to come but could not stay in business long enough to profit from his prediction.
net worth heterogeneity within sectors irrelevant. In other words, it reduces the set of aggregate states to banks’ aggregate net worth $N_t$ and households’ aggregate net worth $N_t$. Second, I can instead use total capital in the economy $K_t$ and the net worth share of banks

$$\eta_t \equiv \frac{N_t}{N_t + N_t}$$

as states of the economy. The reason is that capital is the only asset in positive net supply, and therefore aggregate net worth is equal to the total value of the capital stock. Third, I can dispense with $K_t$ as an aggregate state because allocations (scaled by capital) do not depend on the level of the capital stock. This result follows from the linear production technology and the linearity of decisions in net worth. To summarize, the only aggregate state is the banking sector’s wealth share $\eta_t$.

From now on, I switch to recursive notation, i.e., time subindexes are suppressed. All equilibrium objects are functions of the aggregate state of the economy $\eta_t$, but this dependence is left implicit. Also, I denote by $\tilde{z}$ the value that variable $z$ would take if the sunspot were realized when the aggregate state is $\eta$, i.e., $\tilde{z} = z(\tilde{\eta}(\eta))$, where $\tilde{\eta}(\eta)$ is banks’ wealth share just after a systemic run is realized.

### B. Risks

The economy is subject to two different types of aggregate risk. The first one corresponds to the exogenous real shock and its propagation mechanism. In this case, the risk measure associated with a variable of interest is its sensitivity to that shock, which I refer to as total diffusion risk or diffusion volatility. This risk can be decomposed into an exogenous component and an endogenous one. The former represents the direct effect of the shock absent any change in the behavior of agents, while the latter is the additional sensitivity generated by agents’ endogenous responses, which generates amplification. For concreteness, I refer to these components as exogenous risk and amplification risk, respectively. Note that these risks are defined with respect to a variable of interest, e.g., returns on capital or output growth.

The second type of aggregate risk corresponds to vulnerabilities that arise due to the system’s internal dynamics without any real exogenous impulse, e.g., systemic runs. This type of risk can materialize whenever the economy is sufficiently fragile. In the case of runs, this fragility is captured by the losses depositors would experience in case a run is realized, and it manifests through the arrival rate of runs.\(^{14}\) I refer to the latter as run risk, and it constitutes a risk measure not associated with a particular variable.

\(^{14}\)In fact, losses for depositors and the arrival rate of runs are positively related through the equilibrium selection mechanism chosen. As discussed before, this assumption is natural since it links the vulnerability’s determinant with the likelihood of the risk being realized.
of interest. Of course, the sensitivity of a variable with respect to the realization of the run is also informative about the risk a run represents, but I focus the discussion on the the arrival rate of runs.

Importantly, the two types of risks are qualitatively different. While exogenous real shocks are small and frequent, systemic runs are rare events that occur suddenly and have a discrete impact on real and financial variables. Continuous time allows the model to sharply distinguish between these two types of risk by associating real shocks to a diffusion process and systemic runs to a jump process.

C. Capital price and returns

**Poisson process.** To alleviate notation, from now on I work with Poisson process $J_t$ that captures the realization of systemic runs. Runs require a panic wake-up call $d\hat{J}_t = 1$ and a vulnerable economy $\mathcal{L}^d > 0$, i.e. $dJ_t = 1_{\{\mathcal{L}^d > 0\}} d\hat{J}_t$. The associated arrival rate is $p = 1_{\{\mathcal{L}^d > 0\}} \hat{p}$. Note that the selection mechanism considered implies $p_t = \hat{p}_t = \Gamma(\mathcal{L}^d) = 0$.

**Physical capital.** Return on capital managed by banks is

$$dr^k = \left(\frac{(a - i) k}{q^k}\right) dt + \frac{d(q^k)}{q^k},$$  \hspace{1cm} (6)

where the first term represents capital’s dividend yield net of investment and the second term is the capital gains rate, which includes price and capital quantity fluctuations. Returns on capital managed by households, $dr^k$, are defined analogously using their lower productivity $a$ and corresponding investment rate $\hat{\ell}$.

Given the uncertainty driving the economy, I conjecture the following dynamics for the capital price process

$$\frac{dq}{q} = \mu^q dt + \sigma^q dZ_t - \ell^q dJ_t,$$  \hspace{1cm} (7)

where $\mu^q_t$ is the capital price growth conditional on no aggregate shocks, $\sigma^q$ is the sensitivity of the capital price growth to the aggregate capital quality shock $dZ_t$, and $\ell^q$ is the percentage loss on capital value when a bank run is realized. $\{\mu^q, \sigma^q, \ell^q\}$ are endogenous objects, i.e., functions of the aggregate state. Liquidation value corresponds to $\bar{q} = q(\bar{\eta}) = q(\eta) (1 - \ell^q(\eta))$.

Then, using capital evolution (1), capital return for banks can be written as

$$dr^k = \mu^R dt + \left(\sigma + \frac{\sigma^q}{\ell^q}\right) dZ_t - \frac{\ell^q}{\ell^k} dJ_t$$  \hspace{1cm} (8)
where
\[
\mu^R \equiv \frac{a - \ell}{q} + \Phi(\ell) - \delta + \mu^q + \sigma \sigma^q,
\]
The expected return conditional on no aggregate shocks, \(\mu^R\), includes a dividend component (first term) and deterministic capital gains. The latter are associated with deterministic quantity gains \(\Phi(\ell) - \delta\), deterministic price gains \(\mu^q\), and the interaction between unexpected quantity and price gains (Ito’s term, related to the nonlinearity of value = price \times quantity).

The dynamics of capital return show the different risks discussed above. In this economy, the exogenous aggregate shocks are the capital quality shocks, \(dZ_t\), and the endogenous aggregate shocks are the bank runs coordinated by Poisson shocks \(dJ_t\). Total diffusion risk of capital returns, \(|\sigma + \sigma^q|\), includes exogenous risk, \(\sigma\), and amplification risk, \(\sigma^q\). The former represents the sensitivity of capital quantity to the shock, while the latter represents the sensitivity of capital price. Given that \(\sigma > 0\), the price response amplifies the effect of the initial shock on returns whenever a price increase follows a positive quality shock, i.e., \(\sigma^q > 0\). If the economy experiences a systemic run, i.e. \(dJ_t = 1\), total capital value decreases proportionally to the percentage price drop \(\ell^q\) because there is no effect on quantity (no direct real effect).

The capital returns (in the absence of shocks) obtained by households are identical to those obtained by banks in terms of capital gains, but the dividend component is different due to a lower productivity and a potentially different investment rate, i.e.,
\[
\mu^R \equiv \frac{a - \ell}{q} + \Phi(\ell) - \delta + \mu^q + \sigma \sigma^q.
\]

### D. Consumption and real investment

**Consumption.** Optimal consumption for both agents is proportional to their wealth, i.e.,
\[
\frac{c}{n} = \frac{c}{n} = \rho. \tag{9}
\]
Logarithmic preferences imply that consumption decisions are independent of asset returns.

**Investment.** The return on capital for both agents is maximized by choosing the investment rate that solves
\[
\max_{\ell} q\Phi(\ell) - \ell.
\]

\[15\] The literature focuses on the equilibrium that exhibits amplification, i.e., \(\sigma^q > 0\), but there is another one where capital price movements hedge real shocks, i.e., \(\sigma^q < 0\). Mendo [2018] discusses hedging equilibrium in financial frictions models.
The first-order condition \( q \Phi'(\ell) = 1 \) (marginal Tobin’s Q) equates the marginal benefit of investment (extra capital \( \Phi'(\ell) \) times its price \( q \)) with its marginal cost (a unit of final goods). The concavity of adjustment cost \( \Phi(\cdot) \) implies that investment is an increasing function of the price of capital (if \( \ell > 0 \)), which can be written as

\[
\ell = \ell = \left[ \Phi'^{-1}\left(\frac{1}{q}\right) \right]^+ .
\]

(10)

The investment decision is a completely static problem (it only depends on the current capital price) because the investment process has no delays.

\[ D. \text{ Households’ portfolio decision} \]

\textit{Capital portfolio share.} The portfolio decisions, i.e., capital holdings \( k_t \) and deposits \( b_t \), can be summarized by the capital portfolio share \( x \equiv \frac{k}{n} \). Given \( x \) and individual state \( n \), capital holdings are \( k = xn/q \) and deposits \( b = (1 - x)n \). Also, note that \( \ell^d(k, b) \) presented in equation (2) can be written in terms of \( x \) alone as

\[
\ell^d(x) = \left[ 1 - \frac{q(\eta)}{q} \left( \frac{x}{x - 1} \right) \right]^+ .
\]

(11)

The optimal portfolio share of capital for households \( x \) satisfies

\[
\ell^R - r \leq \frac{x(\sigma + \sigma^d)}{\ell^R}, \quad \text{run risk compensation}
\]

\[
\text{with } \ell^R - r \leq \frac{x(\sigma + \sigma^d)}{\ell^R} \quad \text{run risk compensation}
\]

\[
\text{run risk compensation}
\]

where \( A \equiv \frac{c^{-1} = \rho n^{-1} \text{ denotes household’s stochastic discount factor. The LHS is the market excess return (conditional on no aggregate shocks) of capital over deposits and the RHS is the compensation for risk required by the households to hold capital. If } x > 0, \text{ the condition holds with equality and the excess return needs to compensate for diffusion risk and run risk.}}\)

Both compensations can be expressed as \textit{risk price times risk quantity. The prices of all}}

risks are associated with the agent’s stochastic discount factor. In the case of diffusion risk, the price is measured by the sensitivity of the agent’s net worth to this risk, i.e. \( x(\sigma + \sigma^d) \), while its quantity is the sensitivity of the asset return to diffusion risk, i.e., \( \sigma + \sigma^d \). In the case of run risk, the price is the arrival rate of the run \( p \) times a measure of the drop in net worth when a run is realized, i.e.,

\[
\frac{A}{\Lambda} = \frac{n}{n} = \left( 1 - \ell^d - x(\ell^d - \ell^d) - x \right)^{-1}
\]
while the quantity of risk is the excess loss of capital over deposits, i.e., \( \ell^t - \ell^d \). The following lemma provides a formal characterization of households’ portfolio decision.

**Lemma 1.** Household’s optimal capital portfolio share is

\[
x = \frac{1}{2} \left( x^{nr} + x^{th} - \sqrt{(x^{th} - x^{nr})^2 + 4p(\sigma + \sigma q)^2} \right)^+, \tag{13}
\]

where \( x^{nr} = \frac{\mu^R - r}{(\sigma + \sigma q)^2} \) is the optimal portfolio decision absent runs, i.e. \( p = 0 \), and \( x^{th} = \frac{1 - \ell^d - \ell^t}{\ell^t - \ell^d} \) is the threshold capital share above which the household’s net worth collapses to zero in case of a systemic run. Moreover, \( x \leq \min\{x^{nr}, x^{th}\} \). Households’ capital demand is decreasing in the run intensity \( p \), diffusion volatility \( |\sigma + \sigma q| \), and capital loss \( \ell^d \).

**Pricing of deposits.** Recall that the recovery value of a bank’s deposits depends on the composition of its portfolio. In equilibrium, households have access to a unique type of deposit because all banks choose the same portfolio; however, banks need to know how households would price deposits if they were to choose a different portfolio. The asset-pricing condition for a bank’s portfolio with capital share \( x \) can be written as

\[
r(x) - r^f = p\frac{\bar{A}}{\Lambda} \ell^d(x), \tag{14}
\]

where \( r^f \equiv -\mathbb{E} \left[ \frac{d\bar{A}}{\bar{A}} \right] \) is the return households require for a risk-free asset. The spread of deposits over the risk-free rate needs to compensate households for the run risk and is increasing in the likelihood of the run \( p \), the net worth drop \( \bar{A}/\Lambda \), and the loss rate on deposits \( \ell^d(x) \). An intuitive derivation of this condition follows from including a risk-free asset with return \( r^f \) and a deposit with return \( r(x) \) and loss \( \ell^d(x) \) in households’ problem.

**E. Banks’ portfolio decision**

Capital portfolio share for banks is \( x \equiv \frac{4k}{n} \). Given \( x \) and the individual state \( n \), capital holdings are given by \( k = xn/q \) and debt (deposits) by \( b = (x - 1)n \). The optimal portfolio share for banks satisfies

\[
\mu^R - r(x) = \frac{x(\sigma + \sigma q)^2}{\text{diffusion risk compensation}} + \frac{(x - 1)r'(x)}{\text{increase in deposits’ cost}}. \tag{15}
\]

The LHS is the market excess return (conditional on no aggregate shocks) of capital over deposits and the RHS is the excess return required by households in order to hold capital. The logic behind the compensation for diffusion risk is analogous to that in
the case of households. Recall that banks’ portfolio decisions influence the deposit rate they need to pay households due to their effect on potential losses for depositors. The compensation for the increase in the cost of deposits is the marginal increase in the deposit rate, \( r'(x) > 0 \), times the portfolio share of deposits, \( (x - 1) \).

The first-order condition presented assumes that the cost of underperforming \( \phi \) is large enough, i.e., it satisfies condition (32) in Appendix A. When the economy is exposed to runs, the optimization with respect to capital’s portfolio share features two local maxima: a safe one that implies no failure after a systemic run, and a risky one that triggers bankruptcy if the systemic run is realized. In the absence of a cost \( \phi > 0 \) of underperforming the market, this feature can imply an asymmetric equilibrium among banks, i.e., one in which a fraction of banks choose the optimal risky portfolio and the rest choose the optimal safe one. Choosing the optimal safe portfolio implies underperforming the market and assuming the corresponding cost. The assumption regarding the cost \( \phi \) rules out the safe portfolio choice. The following lemma characterizes the optimal portfolio decision.

**Lemma 2.** For a sufficiently large underperforming cost \( \phi \), banks’ optimal portfolio weight on capital \( x \) satisfies

\[
x = \frac{1}{(\sigma + \sigma^q)^2} \left[ \mu^R - r^f - \frac{\lambda}{\lambda} \right].
\]

\[ (16) \]

banks’ capital demand is decreasing in the run intensity \( p \), capital loss \( \ell^0 \), and diffusion volatility \( |\sigma + \sigma^q| \). A sufficiently large underperforming cost is defined by equation (32) in Appendix A.

**F. Wealth dynamics**

**Evolution of aggregate state.** Agents’ decisions, i.e., on capital, investment, and consumption, are linear in their net worth. Therefore, aggregation is immediate. The dynamics of household and banking sector net worths are

\[
\frac{dN}{N} = \frac{dn}{n} - \frac{\lambda}{\lambda} + \frac{N}{N} \lambda
\]

and

\[
\frac{dN}{N} = \frac{dn}{n} - \lambda + \frac{N}{N} \Lambda
\]

respectively. The only difference between the evolution of a sector’s net worth and that of an individual agent within that sector comes from the Poisson type-switching
processes. The following lemma characterizes the evolution of the aggregate state in the economy.

**Lemma 3.** The banking sector’s wealth share dynamics are

\[
\frac{d\eta}{\eta} = \mu_\eta dt + \sigma_\eta dZ_t - \ell_\eta dJ_t,
\]

where

\[
\begin{align*}
\eta \mu_\eta &= \eta(1 - \eta) \left[ x \left( \mu^R - r \right) - \bar{x} \left( \mu^R - r \right) \right] - \eta \sigma_\eta (\sigma + \sigma_q) + \lambda(1 - \eta) - \lambda \eta \\
\eta \sigma_\eta &= \eta(1 - \eta) (x - \bar{x}) (\sigma + \sigma_q) \\
\eta \ell_\eta &= 1 - \bar{\eta}
\end{align*}
\]

(17) (18) (19)

and \( \bar{\eta}(\eta) \) represents the wealth share of banks after a systemic run and satisfies

\[
\bar{\eta}q(\bar{\eta}) = T
\]

(20)

whenever \( p > 0 \).

These dynamics determine the ergodic distribution of the share of wealth held by banks. Importantly, the dynamics are mainly driven by the portfolio decisions \( x \) and \( \bar{x} \). Absent the realization of any shocks, i.e., considering just the drift of \( \eta \), the wealth share of banks moves according to 1) the difference in portfolio expected returns, i.e., share times market excess return \( x \left( \mu^R - r \right) \), 2) an adjustment term due to the nonlinearity of wealth share as a function of the wealth levels of each sector, \( \eta \sigma_\eta (\sigma + \sigma_q) \), and 3) the type-switching processes.

The wealth share of banks is exposed to capital quality shocks as long as agents choose different portfolio shares. In particular, if banks choose a larger capital portfolio share than households, a positive capital quality shock increases their wealth share. The impact depends on the sensitivity of returns: it includes the direct quantity effect and the amplification effect on the price of capital. The condition that determines the wealth share after a systemic run (20) establishes that banks’ new wealth share \( \bar{\eta} \) is equal to the banking sector’s new wealth \( T K \) divided by the new value of total wealth \( q(\bar{\eta})K \).

\[ G. \text{ Markov Equilibrium} \]

**Market clearing.** For convenience, define \( \psi = x\eta \) as the capital share managed by banks. Also, since capital is the only asset in positive net supply in this economy, total wealth in the economy is \( qK = N + \bar{N} \).
Market clearing for goods (scaled by total capital) can be written as
\[ \rho q + \iota(q) = \psi a + (1 - \psi)q, \]  
(21)
where the LHS represents aggregate demand, and the RHS is aggregate supply. The first term on the left is aggregate consumption, which is a constant fraction \( \rho \) times total wealth \( qK \), and the second term represents aggregate investment \( \iota(q)K \). Total production per unit of capital is equal to agents’ productivities weighted by the share of capital that they manage.

Market clearing for capital (scaled by total wealth) is
\[ x\eta + x(1 - \eta) = 1, \]  
(22)
where \( xN = x\eta qK \) is total wealth invested in capital by banks, and \( xN = x(1 - \eta)qK \) is the corresponding quantity for households.

**Consistency.** Capital price dynamics need to be consistent with the dynamics of the aggregate state, i.e.
\[ q\mu_q = q\eta \mu_q + \frac{1}{2} q\eta \eta (\sigma_q \eta)^2 \]  
(23)
\[ q\sigma_q = q\eta \sigma_q \eta \]  
(24)
\[ \ell^q = 1 - \frac{q(\bar{\eta})}{q} \]  
(25)
These conditions follow from applying Ito’s lemma to function \( q(\eta) \). Also, the equilibrium loss on deposits in case of a run needs to be consistent with banks’ portfolio decision \( x \) and capital price dynamics, i.e., equation (11) has to hold. Finally, run intensity needs to be consistent with depositors’ average losses in case of a systemic run, i.e., equation (3) evaluated at \( \mathcal{L}^d = \ell^d \) or
\[ p = \Gamma(\ell^d). \]  
(26)

**Markov equilibrium.** Given an exogenous transfer policy \( T(\eta) \), an equilibrium is a set of functions of banks’ share of net worth \( \eta \) for prices \( \{q, \mu_q, \sigma_q, \ell^q, r, \ell^d\} \), allocations \( \{c/n, (c/q), \iota, \ell\} \), portfolio decisions \( \{x, \bar{x}\} \), run intensity \( \{p\} \), and the dynamics of banks’ wealth share \( \{\mu_q, \sigma_q, \ell_q, \bar{\eta}\} \) such that

- Agents optimize. Given prices, allocations and portfolio decisions solve the agents’ problem: (9), (10), (12), and (16).
Markets clear: (21) and (22).

The dynamics of the aggregate state are consistent with optimal decisions: (17) – (20).

Capital price dynamics are consistent with the dynamics of the aggregate state: (23) – (25).

Depositors’ losses are consistent with banks’ portfolio choices and capital price dynamics: (11).

Run intensity is consistent with depositors’ losses: (26).

Value functions. The Markov equilibrium definition does not include agents’ value functions but they are necessary for welfare analysis. The following lemma characterizes them.

Lemma 4. The value function for a bank with individual net worth \( n \) can be written as

\[
V(n; \eta) = v(\eta) + \frac{1}{\rho} \log(K) + \frac{1}{\rho} \log \left( \frac{n}{N} \right)
\]

while the corresponding value for a household with individual net worth \( \bar{n} \) is

\[
V(\bar{n}; \eta) = \bar{v}(\eta) + \frac{1}{\rho} \log(K) + \frac{1}{\rho} \log \left( \frac{\bar{n}}{\bar{N}} \right),
\]

where \( v(\eta) \) and \( \bar{v}(\eta) \) satisfy functional equations (35) and (34) in Appendix A, respectively.

The first term in (27) is the indirect utility (scaled by total capital) of a banker that owns the entire sector’s net worth, i.e., \( n = N \), and it is the relevant welfare measure discussed below. The second term is the adjustment for aggregate capital level and captures the non-stationary component of the value function since the economy is constantly growing. The third term represents a simple adjustment to consider the wealth share of the agent within the sector. If there were no switching identity shocks, i.e., \( \lambda = \bar{\lambda} = 0 \), the latter would be a constant. In general, it is time-varying but its dynamics are taken into account by \( v(\eta) \). The description of the value function for households is completely analogous.

H. Numerical approach

The Markov equilibrium defines a functional equation for \( q(\eta) \). Given the presence of jumps, this functional equation is not an ODE, as is the case in basically all continuous-time macroeconomic models in the literature. Besides \( \{q(\eta), q'(\eta), q''(\eta)\} \), the functional
equation also includes \( q(\tilde{\eta}(\eta)) \), the capital price if a systemic run were to arrive at state \( \eta \). Importantly, the post-run banks’ wealth share \( \tilde{\eta}(\eta) \) is endogenous to the system. There is no general approach to solving this class of functional equations.

The key to the two-step numerical approach I implement is that a subset of equilibrium conditions defines a standard ODE for capital price. In particular, it is possible find a map from \( \{\eta, q(\eta)\} \) to \( q'(\eta) \) that is independent from non-local changes \( \{\tilde{\eta}(\eta), q(\tilde{\eta}(\eta))\} \). The first step is to solve this ODE. Given function \( q(\eta) \), the second step is to solve for \( \tilde{\eta}(\eta) \). Finally, all other equilibrium objects are recovered from these two functions. Due to logarithmic preferences, the equilibrium definition does not include value functions, so the numerical procedure does not require iterations. Given the solution for the Markov equilibrium, I use an iterative approach to solve for the value functions. I also develop a numerical strategy to solve for the stationary distribution of the state variable without resorting to simulations. The details of the numerical implementation are discussed in Appendix B.

Importantly, the numerical procedures to solve for the Markov equilibrium, the value functions, and the stationary distribution of the state variable are sufficiently fast to be embedded into an optimization procedure. This makes it possible to implement a direct numerical approach to welfare analysis.

I. Baseline Parameters

The exposition of results combines theoretical insights and numerical illustrations, so I briefly discuss the baseline parametrization summarized in Table 1. I use a logarithmic adjustment cost function, \( \Phi(\iota) = \log(\kappa \iota + 1)/\kappa \), and a linear function for the sunspot arrival rate, \( \Gamma(\iota^{ed}) = \delta \iota^{ed} \). Time is measured in years. I use a standard value for the discount rate \( \rho \) (2 percent) and a slightly low value for the depreciation rate \( \delta \) (3 percent) given that capital quality shocks are similar to persistent productivity shocks with endogenous capital utilization, i.e., depreciation increases during booms since the model measures capital in efficiency units. The Poisson arrival rate for the transition from households to banks \( \lambda \) could be set to zero, but I choose a positive small number (1 percent) to prevent the economy from becoming trapped near \( \eta = 0 \) due to low volatility. The choice of a nonzero value for \( \lambda \) can be viewed as conservative in that it reduces the importance of bank runs for welfare. I also set the total transfer \( T \) to a small number (0.1 percent of total capital).

For the set of parameters \( \{a, a, \sigma, \kappa\} \), I match the set of moments presented in Table 1 at the stochastic steady state, i.e., at the point in the state space at which the economy converges absent any shocks. The moments in the data are calculated using U.S. data for the post-World War II period, except for the largest drop in asset prices, which I
set to 30 percent. The last two parameters require information about the dynamics in vulnerable regions and are therefore matched using the entire stationary distribution. The rate $\lambda$ at which banks become households targets a conservative leverage estimate of 5. Larger values of leverage translate into a larger fraction of time spent in vulnerable regions. Finally, the sensitivity of the arrival rate of runs to potential losses in deposit value $\gamma^e$ is set so that runs happen on average once every 40 years. This paper focuses on mechanisms, so I provide robustness exercises for numerical illustrations but leave more ambitious quantitative exercises to less stylized models.

Table 1: Baseline parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Moment</th>
<th>Target</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depreciation rate $\delta$</td>
<td>0.03</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Discount rate $\rho$</td>
<td>0.02</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Transition hh $\rightarrow$ banks $\lambda$</td>
<td>0.01</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Transfer after run $T$</td>
<td>0.001</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Stochastic steady state

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Moment</th>
<th>Target</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Productivity banks $a$</td>
<td>0.125</td>
<td>Investment/Capital</td>
<td>8.8%</td>
<td>8.8%</td>
</tr>
<tr>
<td>Productivity hh $a$</td>
<td>0.058</td>
<td>Max. price drop</td>
<td>30%</td>
<td>29%</td>
</tr>
<tr>
<td>Exogenous volatility $\sigma$</td>
<td>0.023</td>
<td>GDP growth volatility</td>
<td>2.31%</td>
<td>2.31%</td>
</tr>
<tr>
<td>Inv. adjustment costs $\kappa$</td>
<td>10.19</td>
<td>GDP growth</td>
<td>3.24%</td>
<td>3.24%</td>
</tr>
</tbody>
</table>

Including vulnerable regions

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Moment</th>
<th>Target</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transition banks $\rightarrow$ hh $\lambda$</td>
<td>0.047</td>
<td>Financial sector leverage</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Eq. selection mechanism $\gamma^e$</td>
<td>0.742</td>
<td>Run probability</td>
<td>2.5%</td>
<td>2.5%</td>
</tr>
</tbody>
</table>

4 Instabilities

The full dynamic solution allows me to address the following economic questions about the different forms of risk: 1) when is the economy more vulnerable to real shocks? 2) is this vulnerability affected by the possibility or realization of rare collapses such as systemic runs? 3) can the economy be exposed to systemic runs when real shocks are not amplified and volatility is limited? (iv) if so, how do we identify this hidden risk? and 5) is the economy prone to spending large periods of time exposed to systemic runs and/or amplification of real shocks?

The economy is faced with two endogenous instabilities. First, the response to real shocks is larger than the exogenous fundamental impulse due to the presence of endogenous amplification risk. Second, the economy is exposed to systemic runs, which
trigger collapses in output and asset prices and propel the economy to an unstable regime characterized by large amplification risk and occasional runs.

These two risks vary in distinct ways according to the key endogenous state variable: the wealth share of the banking sector. Amplification risk, which operates through changes in misallocation, is present when this share is sufficiently low to prevent banks from holding all capital in the economy due to the large risk exposure it would involve. In contrast, run risk manifests in the intermediate region of banks’ wealth share and depends on potential losses for depositors. The economy is shielded from runs if banks’ wealth share is too high because their balance sheets are sound in these situations. Run risk is also absent for exceptionally low levels of the state variable because of low asset prices, i.e., small potential price drops.

The distinct dynamics of the two instabilities unveil a hidden risk regime in which volatility of macroeconomic and financial variables is low due to the absence of amplification, yet the economy is exposed to sudden collapses because of runs. In order to distinguish this regime from a truly safe low-volatility environment, it is necessary to consider a set of indicators that can flag the economy’s vulnerabilities, i.e., a risk topography. The model suggests the following as indicators of run risk: the joint presence of high asset prices and a highly (or, at least, moderately) levered banking sector, high excess returns per unit of measured volatility, and a low risk-free rate.

There is a sharp contrast between the feedback among the two forms of endogenous risk. The presence of amplification risk affects asset returns is also a macroeconomic risk, not just a financial one. In order to emphasize this point, I also discuss results in terms of output growth.

A. Financial and macroeconomic amplification risk

The exposition focuses on the effect of amplification risk on capital returns due to fluctuations in asset prices as presented in equation (8). In this model, the volatility of asset prices translates into volatility of output due to the financial accelerator mechanism discussed below. Therefore, the amplification that affects asset returns is also a macroeconomic risk, not just a financial one. In order to emphasize this point, I also discuss results in terms of output growth.
Let \( Y \equiv yK \equiv (\psi(a - a) + a)K \) represent total output in the economy. Then its dynamics can be written as

\[
\frac{dY}{Y} \equiv \mu^Y dt + (\underline{\sigma} + \overline{\sigma^y}) dZ_t - \ell^Y dJ_t,
\]

where \( \sigma^y \) is the amplification associated with changes in allocative efficiency. Changes in the fraction of capital allocated to banks \( \psi \) are directly related to the price of capital in equilibrium through (21). The following lemma characterizes the equilibrium relation between the amplification risk affecting capital returns and output growth.

**Lemma 5.** Amplification risk of output growth \( \sigma^y \) can be written as \( \sigma^y = \varepsilon_{y,q}\sigma^q \), where \( \varepsilon_{y,q} > 0 \) represents the elasticity of aggregate demand (per unit of capital), i.e., \( y^d \equiv \rho q + \nu(q) \), with respect to capital price.

### B. Amplification risk

Amplification risk for capital return \( dr^k \) corresponds to fluctuations in capital price \( \sigma^q \) generated by the real shock driving the economy. Given that shocks in this economy are to the quantity of capital and not to its productivity, a unit of capital always has the same productive capacity. Then, an efficient allocation would imply a constant capital price, i.e., any fluctuation in capital price in response to the real shock operates through the internal dynamics of the system. In particular, these changes follow from portfolio adjustments due to precautionary motives, which vary with the state of the system. This translates into state-dependent amplification risk despite a constant level of exogenous risk \( \sigma \).

The financial accelerator mechanism behind amplification was initially identified in KM and BGG, and the dynamics of this endogenous risk are discussed in BruSan using a similar model but without considering systemic runs. My main interest lies in the effect of run risk on the financial accelerator mechanism and amplification dynamics.

**Proposition 1.** Amplification risk \( \sigma^q \) is independent of the presence of run risk. That is, it is independent of the feasibility of a run equilibrium, the sunspot arrival rate function \( \Gamma(\cdot) \), and the potential losses associated with its realization \( \ell^q \). Moreover, the equilibrium capital allocation \( \psi \) and capital price \( q \) are also independent of run risk.
The amplification mechanism crucially depends on banks’ leverage $x$ (assets to net worth). Given that capital is the asset with the largest value drop in a systemic run, it should be expected that agents’ capital demand decreases when the economy is exposed to runs. Lemmas 1 and 2 show that this is indeed the case. However, the key determinant of amplification risk is the change in banks’ capital demand relative to households’, which turns out to be invariant with respect to run risk. Simple manipulation of optimal portfolio conditions (12), (15), and (14) delivers an illustration of the independence of capital allocation with respect to runs, i.e.,

$$x - \bar{x} \leq \left( \frac{a - a}{q} \right) \frac{1}{(\sigma + \sigma^q)^2}$$

with equality for $x > 0$. The excess capital demand of banks $x - \bar{x}$ depends on their excess return due to the productivity advantage $(a - a)/q$, and the risk associated with real shocks $\sigma + \sigma^q$, but not on run risk. Recall both agents require excess compensation due to run risk: households because capital losses exceed those on deposits, and banks due to the increase in the cost of deposits. The last equation implies that these excess compensations are the same.

There are two reasons to expect that banks’ capital demand decreases more than households due to run risk. First, runs affect banks more than households since they practically wipe out all their net worth and therefore they should be less willing to expose themselves. Second, it is precisely banks’ leverage decision that exposes the economy to systemic runs. However, neither line of reasoning is correct. Regarding the former, banks do not bear losses beyond bankruptcy due to limited liability, meaning marginal changes in banks’ portfolio choices do not affect their net worth after a run. Regarding the latter, systemic runs depend on aggregate leverage of the banking sector, so no agent internalizes this effect. In other words, there is a run externality: individual banks make their portfolio decisions disregarding the fact that, as a sector, they influence run risk. Externalities associated with banks’ leverage decision are discussed in section 5B.

(ii) Mechanism and dynamics

The irrelevance of runs for amplification risk implies that the mechanisms behind amplification risk and its dynamics will be similar to those in previous work. However, it is necessary to understand and characterize both in order to study the joint dynamics of the different forms of risk and the potential influence of amplification risk on run risk.

Endogenous amplification risk depends on 1) bank leverage $x$ and 2) the sensitivity of the capital price to changes in banks’ wealth share $q'(\eta)$. The latter captures the

\footnote{The description of the spiral behind amplification risk closely follows BruSan.}
sensitivity of banks’ capital demand to net worth, which in equilibrium is reflected by \( \psi'(\eta) \). Recall banks’ capital share \( \psi \) and the capital price \( q \) move together to preserve equilibrium in the goods market.

In equilibrium, banks finance capital holdings beyond their net worth by issuing short-term debt to households, i.e., \( x > 1 \), in order to exploit their higher efficiency in allocating resources. Then, a negative aggregate capital quality shock \( dZ_t < 0 \) directly reduces wealth share of banks \( \eta \). If banks’ capital demand is not linked to their net worth, i.e., if they just hold the entire capital stock, then there is no amplification. On the contrary, if banks’ capital demand is limited by their net worth, initial losses trigger a downward spiral. The reduction of relative capital demand by banks implies a drop in its price since households are less efficient at managing capital. The drop in asset prices generates a further reduction in \( \eta \) due to leverage, which translates again into lower capital demand by banks. Figure 1(a) illustrates the spiral.

Figure 1: Amplification risk

(a) Amplification spiral

(b) Amplification risk \( \sigma^q \)

The continuous-time framework allows a simple characterization of the spiral. Consider an exogenous shock that reduces capital quantity by 1 percent; the direct effect on capital price then is a

\[
\zeta \equiv \frac{q'(\eta)\eta}{q(\eta)} \frac{(x - 1)}{\varepsilon_{q,\eta}} \text{ direct drop in } \eta
\]

percent drop. The shock directly reduces the value of capital holdings by 1 percent, which translates in a reduction of \( (x - 1) \) percent of banks’ wealth share \( \eta \). The change in the state implies a price change governed by the elasticity of capital price with respect to the banks’ wealth share, \( \varepsilon_{q,\eta} \). The direct effect on the price of capital feeds back into \( \eta \) and its effect is identical to the effect of the initial quantity shock because price and quantity have the same impact on total value. Then, the second-round effect on capital
price is $\zeta^2$, and the total effect can be found by summing the corresponding geometric series. Therefore, the total effect of a unit shock to capital quantity on capital price, i.e., amplification risk, can be written as

$$\sigma^q = \frac{\zeta}{1 - \zeta} \cdot \sigma$$

The effect of the initial impulse, $dZ_t$, on capital quantity depends on exogenous risk $\sigma$, and the change in capital quantity triggers the spiral described above.

There are two different regimes relevant for amplification risk that are characterized by the strength of the banking sector. In normal times, banks are well-capitalized and are willing to absorb the risk of managing all capital in the economy, so $\psi = 1$ and $q'(\eta) = 0$. There is no amplification spiral, i.e., $\sigma^q = \zeta = 0$, since the equilibrium demand of banks does not depend on their net worth. However, when banks become poorly capitalized, the risk of managing all capital stock is too great and it is optimal to sell some to households despite their lower productivity, so in this regime $\psi < 1$ and $q'(\eta) > 0$. In these situations, banks’ capital demand depends on their net worth because, with larger net worth, their risk-bearing capacity increases, and therefore the spiral described is triggered after negative capital shocks. These dynamics are formalized by the following proposition and illustrated in Figure 1(b).

**Proposition 2.** There exists a threshold $\eta^\psi \in (0, 1]$ such that amplification risk is present, i.e., $\sigma^q > 0$, if and only if banks’ wealth share is below this threshold, i.e., $\eta < \eta^\psi$ (otherwise, $\sigma^q = 0$). Moreover, the positive amplification region is precisely the misallocation region, i.e. the region where $\psi < 1$.

The independence result implies that run risk can affect the manifestation of amplification risk only through its influence on the shape of the stationary distribution.\(^\text{18}\) There are two competing forces at play: 1) when the economy is exposed to run risk, runs cause banks’ wealth share to be almost completely erased, sending the economy into the amplification region, and 2) run risk increases the risk premium on capital, allowing levered banks to increase their wealth share faster. Under the baseline parameter specification, the former effect dominates, and the economy spends more time in the crisis regime when I allow for the possibility of bank runs.

### C. Run risk

Systemic runs represent a different class of risk faced by the economy. Runs are not driven by a real exogenous impulse; rather, they are the outcome of agents’ behavior and

\(^\text{18}\)The independence result shows that amplification risk function $\sigma^q(\eta)$ does not depend on run risk. Nevertheless, run risk will affect the time the economy spends at different $\eta$ values.
can arrive whenever the financial system is sufficiently fragile. The impetus for a run is a coordinating signal that arrives at a rate determined by the equilibrium selection mechanism. Hence, the possibility of a run is driven purely by the system’s internal dynamics and coordination among agents.

There are three relevant measures of the risk that systemic runs pose: 1) the region in the state space in which a run is possible, 2) the losses they impose in terms of variables of interest such as asset values or output, and 3) the frequency of runs. The following proposition shows that, out of these three measures, only the frequency of runs depends on the equilibrium selection device.

**Proposition 3.** Let \( S = \{ \eta \in [0, 1] : \ell^d > 0 \} \) be the set of aggregate states in which a run equilibrium is feasible, i.e., the run vulnerability region. Then, \( S \) is independent of the equilibrium selection device \( \Gamma(\cdot) \). Moreover, the capital value losses \( \ell^q \) and output growth drop \( \ell^y \) that a systemic run generates are independent of \( \Gamma(\cdot) \) as well.

A run may occur whenever the economy is fragile enough that a run on the banking system would imply losses for depositors. Under these circumstances, households’ decision not to roll over deposits yields a self-fulfilling equilibrium. Potential losses for depositors are determined by two factors: 1) bank leverage \( x \) and 2) the drop in asset prices \( q^\eta \) conditional on a run, as evidenced by the equation

\[
\ell^d = \left[ 1 - (1 - \ell^q) \left( \frac{x}{x - 1} \right) \right]^+, 
\]

where \( [\cdot]^+ = \max \{\cdot, 0\} \). Importantly, the preceding proposition demonstrates that the region in the state space in which \( \ell^d > 0 \) is independent of the coordination device chosen. The intuition underlying this result is that, as argued earlier, relative capital demand is independent of the equilibrium selection mechanism because households and banks require equal compensation to bear run risk. Therefore, bank leverage \( x \) is independent of the possibility of a run, so the price function \( q(\eta) \) (and hence the potential loss in the event of a run) is independent of the equilibrium selection mechanism as well.

The run vulnerability region corresponds to intermediate levels of banks’ wealth share. At one extreme, when banks hold a large share of total wealth, the leverage required to hold the entire capital stock is low. Even though the potential fall in asset prices is large, bank equity is sufficient to absorb any losses in the event of a run. On the other hand, when banks hold a very small share of wealth, misallocation is severe and asset prices are depressed. Despite the fact that bank leverage is high, the economy is shielded from runs because the potential drop in asset prices is not large enough to wipe out bank equity. The next corollary formalizes this intuition.
Corollary 1. There exist wealth share levels for banks \( \{ \eta^L, \eta^H \} \in (0,1) \) with \( \eta^L \leq \eta^L \) such that the economy is not vulnerable to runs if banks’ wealth share is lower than \( \eta^L \) or larger than \( \eta^H \), i.e., \( \eta \notin \mathcal{S} \) if \( \eta < \eta^L \) or \( \eta > \eta^H \).

Figure 2 illustrates the state-dependence of depositors’ losses \( \ell^d \) as well as its determinants: the potential drop \( \ell^q \) in the price of capital and bank leverage \( x \). In the region where the economy is exposed to amplification risk, \( \ell^q \) increases in \( \eta \) as asset prices increase, but leverage declines in \( \eta \) as banks’ balance sheets become stronger. The increase in \( \ell^q \) dominates in all numerical exercises, so in this region the potential losses faced by depositors are increasing in \( \eta \). In the region where the economy is exposed only to run risk, \( \ell^d \) is decreasing as a function of \( \eta \) as asset prices are constant but leverage declines.

Within the region where runs are possible, the frequency of runs depends on the sunspot intensity function \( \Gamma(\cdot) \). In the baseline specification, \( \Gamma(\cdot) \) is a linear function of depositors’ losses, so the state-dependence of run frequency coincides with that of depositors’ potential losses.

The intuition behind capital price losses increasing in banks’ wealth share is the following. When a run occurs, banks’ wealth share jumps to \( \eta \approx 0 \), so households are forced to operate essentially the entire capital stock, and the price of capital falls to its minimum value. The drop in the price of capital is larger when the initial capital price is higher, i.e., when \( \eta \) is higher and misallocation is lower. Allocative efficiency also jumps to a minimum after a run, so the potential drop in output is increasing in \( \eta \) as well. This implies runs are more costly in the region where the economy is only exposed to runs. This logic is formalized below.

Corollary 2. For a non-contingent transfer policy, i.e., \( T(\eta) = \bar{T} \), potential losses in terms of output \( \ell^Y \) and capital value \( \ell^q \) are weakly increasing in \( \eta \) within the region vulnerable to runs.

The economy is not necessarily exposed to run risk in any region of the state space. Whether run risk emerges depends on the losses to which depositors are exposed, which in turn are a function of bank leverage and the potential fall in asset prices. When either the exposure of asset prices to runs or leverage is sufficiently low, depositors never rationally decide to run because the levered losses incurred by banks are too small to wipe out their equity. For instance, if banks’ productivity approaches that of households \( (a \to a) \), the potential drop in prices goes to zero, which itself is enough to rule out runs. More importantly, as will be shown below, large values of exogenous risk \( \sigma \) will endogenously restrain bank leverage to the point that households never have to fear a loss on their deposits. This result stands in stark contrast to Proposition 2, which establishes that amplification risk is always present in some region of the state space.
D. Joint dynamics and the “hidden risk” regime

The economy’s state space is comprised of three distinct risk regimes: 1) a highly unstable crisis regime, 2) an apparently stable hidden risk regime, and 3) a truly stable safe regime. The regimes are characterized by the economy’s exposure to endogenous risks (amplification risk and run risk) and correspond to distinct regions of the state space. As a consequence, the qualitative features of the economy’s dynamics and the distribution of variables of interest, e.g. capital returns or output growth, vary dramatically across these three regions.

Figure 3(a) presents the distributions of output growth conditional on the state $\eta$, and Figure 3(b) presents one example of a distribution per risk regime. These graphs illustrate how the different endogenous risks translate into outcomes for variables of interest. Amplification risk implies large variability, but it cannot generate a distribution of output growth with asymmetries or fat tails given the normality of real shocks. In contrast, run risk generates an output growth distribution with negative skewness and fat tails. Systemic runs are responsible for the second mode at large negative growth rates, so this mass in the left tail appears only when the economy is vulnerable to runs, i.e., at intermediate values of $\eta$. Figure 3(c) presents the stationary density for banks’ wealth share $\eta$ and distinguishes the three risk regimes.

The safe region, amplification risk and systemic runs are absent. In this region, banks’ share of wealth is large and they hold the entire capital stock without exposing depositors to losses in the event of a run. There is no misallocation, so the price of capital is constant. Banks’ wealth share is exposed only to the direct effect of real shocks, i.e., the change in the quantity of capital. Output growth inherits the distribution of the growth of the capital stock, which is normal with volatility $\sigma$. 

34
In the *hidden risk* region, the economy is exposed to systemic runs, yet insulated from amplification risk. Banks’ wealth share must lie in an intermediate range: it must be small enough for depositors to rationally fear losses in the event of a run, but large enough that banks do not wish to fire-sell assets to households. Apart from the economy’s response to the realization of systemic runs, the dynamics in this region are identical to those in the safe regime. In particular, in the absence of runs, the dynamics of output and asset prices are observationally equivalent in the two regimes. Banks hold the entire capital stock, so capital price is constant and banks’ portfolios are exposed only to exogenous risk and systemic runs. The distribution of output growth is similar to that in the safe regime, except for a second mode at large negative growth rates resulting from the economy’s exposure to runs.

In the *crisis* region, banks’ share of aggregate wealth is low enough that the economy...
is exposed to amplification risk. Negative real shocks force banks to sell assets, which increases misallocation, lowers asset prices, and sets off a loss spiral. As a result, the volatility of macroeconomic and financial variables is high in this region. Systemic runs may also be possible in this region due to the weakness of bank balance sheets, but their incidence is not guaranteed because asset prices in this region may be so low that the potential loss during a mass liquidation is insufficient to wipe out bank equity. The distributions of output growth and returns on capital have high variance and are bimodal when runs are possible.

Relative to earlier work such as BruSan, the main qualitative difference brought about by the introduction of systemic runs lies in the dynamics of the hidden risk regime. Unlike the safe region and the crisis region, the hidden risk region may not always exist. Indeed, as argued earlier, there are some parameter combinations for which the economy is never exposed to runs. Therefore, the relevant question is whether a hidden risk regime is always present, conditional on the economy being exposed to systemic runs, i.e., $S \neq \emptyset$. The following corollary describes a necessary and sufficient condition for the existence of the hidden risk regime: positive losses for depositors (conditional on a run) when the state is at the boundary $\eta^\psi$ of the crisis regime.

**Corollary 3.** Given a constant transfer policy $\mathcal{T}$, there exists a hidden risk regime if and only if a run equilibrium is feasible at the threshold banks’ wealth share level $\eta^\psi$, i.e., $\ell^d(\eta^\psi) > 0$.

The condition mentioned in corollary 3 is satisfied for all parameter combinations for which the economy is exposed to run risk. Moreover, in each such numerical exercise, depositors’ losses $\ell^d$ reach a maximum at $\eta^\psi$. Absent a formal proof, this illustrates that the hidden risk regime is a robust feature of the model.

Continuous-time methods allow me to analytically characterize the stochastic evolution of the variables of interest at any point in the state space and, therefore, to calculate any relevant conditional moment. In order to illustrate the relative contributions of each risk to variability, Figure 3(d) presents the decomposition of the variance of output growth\(^{19}\) into components associated with exogenous risk, amplification risk, and run risk.

For baseline parameters, the contribution of exogenous risk to the conditional variance of output is dwarfed by the contributions of endogenous risks. The variability associated with run risk is at least as important as that due to amplification risk in the crisis regime. Furthermore, it is clearly the dominant source of risk in the hidden risk regime, where the potential drops in output and asset prices are largest.

\(^{19}\)See Appendix C for details about the decomposition.
E. Ergodic instabilities

The stationary distribution of the state of the economy has several interesting properties. To begin with, the economy is highly prone to instability under the baseline parametrization, as evidenced by the stationary distribution in Figure 3(c). In fact, the economy spends more time in the hidden risk region than it does in the safe region. More subtly, the stationary distribution exhibits a counterintuitive feature: despite the fact that, more often than not, the state of the economy is in a region where endogenous risks are present, it is also true that most of the time there is no observable amplification. This can be seen in Figure 3(c) by noting that the economy spends a considerably larger span of time in the hidden risk region than in the crisis region.

Another striking property of the stationary distribution is that there is a significant amount of mass in the extreme left tail near $\eta = 0$. The following lemma shows that this mass is entirely due to runs.

Lemma 6. Consider the model without runs, i.e., $\Gamma(\cdot) \equiv 0$. If there is a positive flow of agents from the household to the banking sector $\Lambda > 0$, then the stationary distribution has density zero at $\eta = 0$.

This lemma shows that exogenous risk and amplification risk together do not generate enough volatility to keep banks’ wealth near zero when there is a stabilizing force on their wealth. The implication is that systemic runs on the financial system are wholly responsible for the most severe crises. Further, the size of the mass near $\eta = 0$ in the stationary distribution indicates that runs are not only severe but subject to a slow recovery.

There is a sharp contrast between the shape of the ergodic distribution in the crisis regime and outside it. The reason is that, in the crisis regime, excess returns on capital include compensation for amplification risk, which favors wealth accumulation in the banking sector. Therefore, the endogenous dynamics imply a relatively fast though highly volatile transition out of the crisis regime. This explains the relatively low mass in this region. Once amplification disappears, banks’ accumulation of wealth is significantly slower, as reflected by the large increase in the density at threshold banks’ wealth share $\eta^\psi$.

The shape of the $\eta$-density within the crisis regime follows from the dynamics of endogenous risks: when risks are small or absent, excess returns on capital are also limited and banks accumulate wealth slowly relative to households. Therefore, recovery is slowest immediately after a systemic run, i.e., close to $\eta = 0$, since there is no run risk and amplification risk is small (because for low levels of $\eta$, even a large percentage movement implies a small absolute change, which determines the extent of the loss spiral). As $\eta$
increases in the crisis regime, amplification and run risk increase, so the transition out of this highly unstable regime speeds up as long as no run is realized. The downward slope of the density in the crisis regime follows from this logic.

F. Risk topography

The hidden risk regime and the safe regime are observationally equivalent in several respects. For one, the level of asset prices, output per unit of capital, and real investment per unit of capital are all identical across the two regimes. Moreover, in each of the two regimes the volatility of output is equal to the exogenous volatility of the quantity of capital, and the volatility of asset prices is equal to zero.

In order to distinguish between these two regimes, it is necessary to develop a risk topography to describe the joint behavior of macroeconomic variables, financial indicators, asset prices, and returns. In this environment, three important quantities suggest presence of hidden run risk. First, the simultaneous incidence of high asset prices and high bank leverage directly implies that depositors will take large losses in case of a run, since these are precisely what drive depositors’ potential losses \( \ell^d \). Second, high excess returns on capital per unit of observed volatility can be indicative of run risk as well. Even though run risk does not affect relative capital demand, it does affect the risk premium required by both types of agents. Hence, risk premia relative to asset price volatility will be high in the hidden risk regime because observable volatility in that regime is low. Third, a low risk-free rate indicates the possibility of a run. Although there is no risk-free asset in the model, an asset-pricing exercise shows that the risk-free rate tends to be low in the hidden risk regime. This is because the possibility of a collapse due to a run drags down agents’ expected consumption growth.

5 Risk to control risk

The following examines how the model’s endogenous instabilities change as the amount of exogenous risk \( \sigma \) varies. The interactions between the different types of risks are key to understanding the main insight of the model: stability breeds instability and its contrapositive, risk controls risk. Economies with the lowest level of exogenous risk will be most exposed to systemic runs in the sense that the region in which they can occur is largest for low \( \sigma \). Importantly, this result will hold independently of the equilibrium selection device. On the other hand, total volatility \( \sigma + \sigma^q \) will decrease alongside exogenous risk.

These results lead to a non-monotonic welfare ranking with respect to exogenous risk. For high levels of exogenous risk, the economy is shielded from runs and agents dislike
increases in exogenous risk since they translate into a larger probability of entering the crisis regime and worse crises, i.e., higher misallocation and amplification. However, for low levels of exogenous volatility, the economy is exposed to runs and agents favor increases in exogenous volatility because they restrain risk-taking by banks and prevent runs. The welfare analysis assumes the equilibrium selection mechanism is independent of exogenous volatility.

A. Stability breeds instability

A brief examination of the mechanisms at play illustrates why low exogenous volatility leads to lower total volatility but greater run risk. When exogenous volatility is low, the capital demand of both types of agents increases, but banks’ demand increases relative to households’ because of their greater exposure to the exogenous risk. Therefore, banks’ equilibrium leverage increases. The increase in leverage is dampened (but not completely reversed) by the corresponding increase in amplification risk. Higher leverage reduces misallocation and increases asset prices. However, higher leverage and asset prices also increase the economy’s exposure to a run. Unlike the increase in amplification risk, the increase in run risk does not feed back into banks’ equilibrium leverage because of 1) limited liability, which makes the banks’ fate after a run – bankruptcy – insensitive to their risk-taking at the margin, and 2) banks’ failure to internalize their contribution to aggregate run risk. Figure 4 illustrates this logic.

Next I describe in detail how each type of endogenous vulnerability responds to lower exogenous risk and provide formal results.

Amplification risk. Lower exogenous risk has two opposing effects on the loss spiral that generates amplification risk. First, it decreases the impact of a given exogenous impulse on the quantity of capital. Second, it increases the sensitivity of banks’ relative capital demand to changes in net worth, i.e., $\psi'(\eta)$ and $q'(\eta)$ are higher in the crisis
regime. As a result, the crisis regime shrinks, but amplification risk is higher in the new crisis regime. Nevertheless, the increase in amplification is never enough to undo the initial reduction in exogenous volatility, so total risk associated with real shocks decreases. Panels (a) and (b) in Figure 5 illustrate this effect and the following proposition formalizes it.

**Proposition 4.** Total diffusion risk \( \sigma + \sigma^q \) is weakly increasing in exogenous risk \( \sigma \) for every \( \eta \in [0,1] \). Also, capital price \( q \) and banks’ capital share \( \psi \) (or leverage \( x \)) are weakly decreasing in \( \sigma \) for every \( \eta \in [0,1] \).

**Run risk.** When exogenous risk is reduced, total risk associated with real shocks decreases as well. Banks then increase their leverage, resulting in less misallocation and higher capital prices. The greater fragility of bank balance sheets and higher capital prices both contribute to a larger potential loss for depositors and hence a greater exposure to systemic runs (holding the liquidation price fixed). The reduction in exogenous volatility does not have a large effect on the liquidation price because the transfer to reboot the banking sector is small enough that the liquidation price is close to \( q(0) \), which is independent of exogenous volatility. Nevertheless, in general the small increase in the liquidation price due to lower exogenous risk dampens the increase in run risk. This effect is illustrated in Figure 5(c) and made precise by the following proposition.

**Proposition 5.** For a constant liquidation price, the size of the run vulnerability region, i.e., the Lebesgue measure of \( S \), is weakly decreasing in \( \sigma \). A constant liquidation corresponds, for example, to the case where there are no transfers to banks after their collapse, i.e., \( T \to 0 \).

In this model, both endogenous risks remain as exogenous volatility vanishes. For amplification risk this effect is simply the “volatility paradox” of BruSan: as the exogenous
impulse vanishes ($\sigma \to 0$), the spiral effect grows unboundedly ($\zeta/(1 - \zeta) \to \infty$) in a way that endogenous amplification risk persists ($\sigma^q \to \delta^q \in \mathbb{R}$). The persistence of run risk as exogenous risk vanishes in this model is actually more powerful than the volatility paradox. First, given the previous result, when exogenous risk vanishes run risk is actually maximized. Second, if exogenous risk is initially high enough that the economy is not exposed to runs, a reduction of exogenous risk may actually introduce the possibility of runs to the economy. This intuition is partially formalized below.

**Corollary 4.** Amplification risk does not vanish as exogenous risk disappears, i.e., threshold level for banks’ wealth share $\eta^b \to \eta^+ > 0$ as $\sigma \to 0$.

**B. Externalities and welfare measures**

In order to understand how exogenous risk affects welfare in this economy, it is first necessary to examine the externalities arising from banks’ portfolio decisions and discuss the relevant welfare measures.

*Externalities.* Absent runs, there are two pecuniary externalities related to banks’ portfolio decisions (or capital demand). First, there is a static pecuniary externality associated with the level of capital price. Banks do not take into account that a decrease in their leverage has a negative impact on the price of capital, which reduces the risk-bearing capacity (or net worth) of other banks, forcing them to decrease their capital positions and increase misallocation. In this sense, banks’ leverage is inefficiently low. Second, there is a dynamic pecuniary externality associated with the volatility of capital price. Banks disregard that, by increasing their leverage, they increase the amplification of future real shocks in the economy. In this other sense, banks’ leverage is inefficiently high. Therefore, these externalities bias leverage in different directions.

Bank runs introduce two additional externalities in banks’ portfolio decisions. First, banks do not internalize that their choice of leverage affects the economy’s exposure to systemic runs. Larger leverage increases potential losses for depositors directly and through asset prices: higher leverage increases the price of capital and therefore the size of the drop to its liquidation price. Agents fail to internalize this effect because individual decisions have zero impact on aggregate variables and run risk depends on aggregate losses of depositors if the bank run is realized. Second, banks disregard the cost of transfers after their collapse. These run externalities lead to excessive leverage. Importantly, even though their presence does not depend on the equilibrium selection mechanism (as long as the economy is exposed to runs), their magnitude does.

*Welfare measures.* The solution to the model provides a detailed characterization of agents’ well-being. In particular, it delivers a solution for the value function (scaled
by total capital) conditional on each aggregate state, i.e., $v(\eta)$ for banks and $\underline{v}(\eta)$ for households, and also describes the frequency with which the economy visits each of these states, i.e., the stationary distribution. Then, a welfare comparison across economies needs to select a relevant summary statistic. For a given agent, the most intuitive statistic is the expected indirect utility where the expectation is taken with respect to the stationary distribution, i.e., $E[v(\eta)]$ and $E[\underline{v}(\eta)]$. For the entire economy, a Pareto-weighted average of these statistics is a natural measure. The non-monotonic welfare ranking with respect to exogenous risk (emphasized later in this section) will hold not only for any Pareto weights but will be robust to alternative welfare measures such as the indirect utility conditional on a particular value of $\eta$.

C. Exogenous risk and welfare

The welfare consequences of changing the level of exogenous risk in the economy can be conceptually decomposed into three components: 1) a fundamental effect, 2) a redistributive effect, and 3) the impact on externalities.

**Fundamental effect.** The fundamental effect of exogenous risk on welfare is the impact it would have in a frictionless economy. In this model, the frictionless benchmark corresponds to a version that allows for equity issuance. In this case, the model collapses to a growth model with a representative agent that has access to the most productive technology. All financing happens through equity contracts, so there is no short-term debt and no runs. Larger exogenous risk decreases the utility of the representative agent due to risk aversion. The following lemma formalizes this intuition.

**Lemma 7.** With complete markets, i.e., allowing for equity issuances, the indirect utility function of the representative agent is decreasing in exogenous volatility $\sigma$.

**Redistributive effect.** The redistributive effect corresponds to shifts in the wealth distribution. Without externalities, any wealth redistribution has a positive welfare impact for the agent whose net worth increases and a negative one for the agent whose wealth decreases. In comparisons across economies, this effect is related to shifts of stationary distribution. For example, a switch toward a stationary distribution that spends more time at states where banks own a small fraction of wealth would represent a redistribution toward households. This is precisely the case when considering a reduction of exogenous risk.

Lower $\sigma$ increases the exposure to systemic runs – see Proposition 5 – so the economy is more frequently propel to situations where households own a large fraction of wealth. It also decreases excess capital returns in the safe regime, making the economy less likely to visit this region of low wealth share for households. Therefore, if reducing exogenous
risk has a negative welfare effect for households, it will be in spite of the redistributive impact.

Impact on externalities. A reduction of exogenous risk increases banks’ leverage – see Proposition 4 below – so the impact of $\sigma$ on externalities in this model depends on whether banks’ leverage is inefficiently high or inefficiently low. If leverage is inefficiently low, the main effect of a marginal reduction of $\sigma$ is to increase allocative efficiency (reduce the static pecuniary externality), and therefore it benefits both agents. If leverage is inefficiently high, the principal effect of a marginal decrease of $\sigma$ is the increase of endogenous instabilities (larger dynamic pecuniary externality and run externalities), so it is detrimental for both agents.

Whether leverage is inefficiently low or high depends on the relative strengths of externalities (and potentially on the welfare measure adopted). When exogenous risk is high enough to shield the economy against runs, leverage tends to be too low, so reducing exogenous volatility reduces inefficiencies in the economy. Moreover, as discussed before, the larger endogenous amplification risk does not overcome the fundamental effect of lower $\sigma$, so total risk associated with real shocks decreases.

In contrast, when exogenous volatility is low enough to expose the economy to systemic runs, equilibrium leverage tends to be too high and a marginal decrease in $\sigma$ is costly for agents. Intuitively, the inefficiencies caused by systemic runs tend to be stronger because there is no feedback from larger run risk to portfolio decisions (as is the case when considering risk associated with real shocks).

Non-monotonic welfare ranking and run risk. The overall effect of exogenous risk $\sigma$ on each agents’ welfare is illustrated in Figure 6(a). It shows the expected indirect utility of agents, i.e., $E[v(\eta)]$ and $E[u(\eta)]$, and the annual probability of systemic runs $E[p(\eta)]$ for economies with different levels of exogenous risk. The non-monotonicity of the welfare measure for both agents is evident. Moreover, the change in the relationship between the welfare measures and exogenous risk coincides with the presence of systemic runs. This feature is robust to alternative parameter configurations (see Appendix D). It is important to mention that results presented in this section assume that the equilibrium selection mechanism $\Gamma(\cdot)$ is invariant to exogenous volatility $\sigma$.

The reversal of the relationship between welfare and exogenous risk in the presence of systemic runs indicates that the main driving force is the effect of $\sigma$ on externalities. The directions of the other two effects discussed do not depend on the presence of run risk: lower $\sigma$ implies a positive fundamental effect and a redistribution toward households. In contrast, leverage switches from being inefficiently low to inefficiently high in the presence of run risk, which implies that lower $\sigma$ exacerbates externalities.

The non-monotonicity result and its dependence on the presence of runs not only hold
on average, i.e., considering the expectation operator over the stationary distribution, but also hold conditioning on any particular wealth distribution, i.e., taking $\eta$ as given. Figure 6(b) shows the expansion of the Pareto frontier of the economy as $\sigma$ decreases when the economy is not exposed to runs (high exogenous risk), while Figure 6(c) illustrates the exact opposite effect when the economy is exposed to runs (low exogenous risk). The frontier represents combinations $(v(\eta), v(\eta))$ for each $\eta \in [0, 1]$. To further illustrate the dependence on run risk, Appendix D shows the welfare ranking assuming away run risk, i.e., $\Gamma(\cdot) \equiv 0$, and verifies that welfare is monotonic for households.

6 Policy

Given the financial friction, i.e., the restriction not to issue equity, short-term debt fulfills the key role of allowing banks to finance larger capital holdings in order to improve productive efficiency in the economy. However, this comes at the cost of systemic instability due to amplification and run risk. As discussed before, banks do not appropriately internalize all the costs or benefits of their portfolio decisions, which opens the possibility for a welfare-improving policy intervention.

First, I consider a simple and intuitive macroprudential policy: a constant leverage cap $\tilde{L} \geq 1$. The analysis unveils important shortcomings of this policy as a tool to limit endogenous risks. Surprisingly, a leverage constraint increases amplification risk whenever it is not binding and exacerbates exposure to systemic runs within the hidden risk regime. Despite these limitations, a leverage cap can improve welfare by reducing

---

20 The expansion (contraction) of the frontier does not necessarily imply that both agents are better (worse) off at each aggregate state $\eta$ but that is indeed the case for Pareto frontiers shifts discussed (for details and robustness exercises, see Appendix D).

21 Absent runs, the welfare ranking might be non-monotonic in exogenous risk for banks because an increase of $\sigma$ redistributes wealth toward them.
overall run risk.

Second, I study a leverage cap contingent on the wealth share of the banking sector: \( \bar{L}(\eta) \geq 1 \). The optimal state-contingent leverage cap restrains banks’ risk-taking when the economy would be exposed to runs, i.e., in the intermediate range of banks’ wealth share, but allows for large leverage when banks’ wealth share is particularly low, i.e., during downturns.

Third, I study economies in which exogenous risk varies along with the wealth share of the banking sector and explore which of these relations, \( \sigma(\eta) \), benefits agents the most. Results indicate that agents would prefer to live in an economy in which exogenous risk increases in regions of \( \eta \) where the economy would be exposed to systemic runs. Beyond the model, this suggest that exposure to systemic runs can be an unintended effect of policies that stabilize real fluctuations.

6.1 Constant leverage constraint

The aim of implementing a leverage cap, despite the increase in misallocation it implies, is to limit the endogenous risks that leverage fosters, i.e., amplification and run risk. However, there are important shortcomings in using a constant leverage cap as a policy tool to control risk.

A. Misallocation and risk control

\textit{Misallocation.} The main cost of implementing a leverage constraint is to decrease allocative efficiency, i.e., to reduce the capital share of banks. This drop is the largest during downturns when banks are poorly capitalized and prices signal they should use substantial leverage. The drop in equilibrium leverage is not limited to situations in which the constraint is binding (the low \( \eta \) region) but extends to other regions of the state space because banks reduce their capital demand in anticipation of a deeper recession. Since capital price is directly related with allocative efficiency, the leverage constraint also reduces it. The following proposition formalizes the increase in misallocation.

\textbf{Proposition 6.} The share of capital \( \psi \) held by banks weakly decreases with the implementation of a leverage cap, i.e., \( \psi(\eta) \geq \psi(\eta; \bar{L}) \) for \( \eta \in [0,1] \). The same applies to the capital price function \( q \).

\textit{Amplification risk.} The leverage cap controls amplification risk whenever it is binding. This is the case during downturns, i.e., the region where banks’ wealth share is low. However, this leverage constraint actually increases amplification risk whenever both: 1) the economy exhibits misallocation and 2) the constraint is not binding. This increase
in amplification risk is necessary to ensure both agents still want to hold capital. The intuition is the following. A lower capital price implies a larger gap between market returns for banks and households \((a - q)/q\) while a lower capital share for banks implies their relative capital demand \((x - x)\) is lower, so the only way that both agents could still be willing to hold capital is if total diffusion risk \(\sigma + \sigma^q\) increases. This logic is illustrated by

\[
\frac{a - q}{q} = \frac{(x - x)(\sigma + \sigma^q)^2}{q}.
\]

An alternative interpretation is that the only way banks will reduce their relative capital demand (with respect to households’) given the increase in the return gap is if diffusion risk increases. Moreover, the leverage constraint also enlarges the region of positive amplification risk. The following proposition formalizes the limitations of a constant leverage cap to contain amplification risk, and Figure 7(a) shows how amplification risk varies with the inclusion of a leverage cap.

**Proposition 7.** Amplification risk weakly increases when including a leverage cap, i.e., \(\sigma^q(\eta) \leq \sigma^q(\eta; \bar{L})\), for the \(\eta\)-region in which there is misallocation in both environments and the leverage cap is not binding, i.e., \(\left\{\psi(\eta; \bar{L}), \psi(\eta)\right\} \in [0, 1)\) and \(x(\eta; \bar{L}) < \bar{L}\). Moreover, the region with positive amplification risk weakly increases, i.e., \(\eta^\psi \leq \eta^\psi(\bar{L})\).

**Run risk.** The leverage cap achieves lower equilibrium leverage, which dampens run risk. However, the inclusion of such a constraint decreases the liquidation price of capital, increasing potential drops in asset prices and the economy’s exposure to systemic runs. The leverage effect usually dominates when both forces are present, i.e., in the crisis regime, but only the second force is present in the hidden risk regime (where there is no misallocation, so leverage is the same with and without the constraint). Therefore, a leverage cap increases run risk in this regime. The following proposition formalizes the dimension in which a leverage cap increases run risk in the economy, and Figure 7(b) illustrates how this risk changes with the inclusion of a leverage cap.

**Proposition 8.** (a) The liquidation price weakly decreases with the inclusion of a leverage cap, i.e., \(q(\bar{\eta}(\eta)) \geq q(\eta(\eta; \bar{L}); \bar{L})\).

(b) For the \(\eta\)-region that belongs to the hidden risk regime in both solutions, i.e. \(\psi(\eta) = \psi(\eta; \bar{L}) = 1\) and \(\left\{p(\eta), p(\eta; \bar{L})\right\} \in \mathbb{R}^+\), the solution with the leverage constraint is weakly more prone to runs, i.e., \(p(\eta) \leq p(\eta; \bar{L})\).

**Resilience of endogenous risks to leverage caps.** While the only way to eliminate all amplification risk using a leverage cap is to ban any debt issuance, i.e., \(\bar{L} = 1\), it is possible that the economy is completely shielded from systemic runs with \(\bar{L} > 1\).
B. Leverage constraint and welfare

Despite the shortcomings of a constant leverage constraint in controlling endogenous risks, it can still improve agents’ welfare by reducing the economy’s exposure to runs. Figure 7(c) summarizes the comparison between economies with different leverage caps in terms of welfare and exposure to systemic runs. It shows that, as long the leverage cap helps reduce run risk ($L > 3.4$ for baseline parameters), both agents prefer a tighter leverage constraint. However, once the leverage cap is tight enough to shield the economy completely from runs, there is a disagreement among agents about reducing $L$ any further. Banks would prefer an even tighter leverage constraint while households prefer $L$ to be just small enough to avoid runs.

A leverage cap poses a trade-off between misallocation and endogenous instabilities: a tighter leverage cap increases misallocation but potentially reduces endogenous insta-
bilities. Results suggest it is worthwhile to impose the larger misallocation when the reduction of endogenous risk includes shielding the economy with respect to runs. For leverage caps low enough to prevent any chance of systemic runs, a tighter leverage constraint might benefit banks because of the redistributive effect: economies with lower $\bar{L}$ spend more time in regions where banks own a larger fraction of wealth. The reason is that risk-adjusted capital excess return increases for the region where the constraint is binding (lowest $\eta$ region). The leverage cap limits competition among banks, allowing larger risk-adjusted returns that translate into faster recoveries.

The same exercise assuming away run risk (not shown), i.e., $\Gamma(\cdot) \equiv 0$, further illustrates its role. Absent runs, households prefer not to have any leverage constraint (banks still like it due to the redistributive effect). This means that the misallocation cost tends to be stronger than the imperfect reduction in amplification risk, at least for the agent relatively less exposed to capital.

Given that welfare results weight opposing forces, they depend on the magnitude and sensitivity of the arrival rate of runs, i.e., they are sensitive to the equilibrium selection mechanism. Appendix D shows that the qualitative features are robust to alternative parameters. Also, an implicit assumption in this exercise is that the equilibrium selection mechanism is invariant to the leverage cap and exogenous risk.

6.2 State-contingent leverage cap

A state-contingent leverage cap $\bar{L}(\eta)$ can overcome some of the shortcomings highlighted in the case of the constant leverage constraint. In particular, it allows for controlling banks’ risk-taking when the economy would be exposed to runs, i.e., the middle region of the banking sector’s wealth share, without limiting leverage during downturns, i.e., when the aggregate state variable is particularly low. The latter avoids deepening recessions that follow a run, i.e., it prevents fostering run risk by depressing the liquidation price of assets. This intuition is illustrated by the following numerical exercise.

I search for the state-contingent leverage constraint that maximizes an equally Pareto-weighted average of expected indirect utilities. Figure 7(d) shows banks’ leverage, $x$, and the arrival rate of runs, $p$, for the laissez-faire equilibrium and the one with the optimal leverage constraint. The difference between leverage levels corresponds to the region where the optimal leverage constraint is binding. This optimal state-contingent leverage constraint unequivocally reduces run vulnerability, i.e., the $p$ decreases for

---

22In particular, I consider the family of functions described by $L(\eta) = \sum_{i=1}^{5} c_i x^i$ and use as a welfare measure $0.5\mathbb{E}(v(\eta)) + 0.5\mathbb{E}(\bar{v}(\eta))$, where the expectation is with respect to the stationary distribution of the state variable.
every value of the aggregate state variable because it reduces leverage and capital price without having an effect over its liquidation value.

Interestingly, the optimal leverage constraint reintroduces amplification risk in the hidden risk regime of the laissez-faire equilibrium in order to limit risk-taking by banks and shield the economy from run risk. This highlights the larger detrimental effects of systemic runs, especially the ones that occur when the economy is seemingly stable since these generate the largest collapses. This point is also illustrated by the fact that the optimal state-contingent leverage constraint allows for run risk only for relatively low levels of banks’ wealth share, i.e., it eliminates completely run risk for levels of the aggregate state variable associated with the hidden risk regime in the laissez-faire equilibrium.

A parallel exercise in an environment without runs (not shown), i.e., \( \Gamma(\cdot) \equiv 0 \), shows that the optimal state-contingent leverage constraint only slightly limits risk-taking around the banks’ wealth share level that divides the crisis regime and the safe one, \( \eta^C \), around which amplification risk is the highest. This indicates that only for extreme amplification risk it is worthwhile to increase misallocation by using a leverage constraint to control this instability.

### 6.3 Stabilization of real shocks

Previous results showed that stability can breed instability, and they illustrated that, given this trade-off, agents prefer an economy with strictly positive exogenous risk. This suggested that a full stabilization of real shocks has a cost in terms of financial fragility and might not be an appropriate policy objective. In order to further illustrate this point, I assume a benevolent policymakers are able to influence fluctuations due to real shocks directly and at no cost, and evaluate how they would stabilize real fluctuations depending on the aggregate state variable of the economy: the wealth share of the banking sector.

The exercise searches for the exogenous risk function \( \sigma(\eta) \) that maximizes an equally Pareto-weighted average of expected indirect utilities. Results reinforce the message that full stabilization is not optimal in the presence of run risk: it is optimal to allow relatively large fluctuations due to real shocks whenever the economy is exposed to runs, i.e., in the crisis and hidden risk regimes for baseline parameters, to control exposure to runs. Nevertheless, it is optimal to stabilize, i.e., insure against real fluctuations in the economy, when there is no exposure to runs, such as in the safe regime. Figure 8(a) illustrates this exercise (the colors indicate the regimes in the economy without the policy) and Appendix D shows robustness to alternative parameter configurations.

---

\(^{23}\)The optimization is performed over fifth-degree polynomial functions.
A similar exercise conducted in the version of the economy without runs delivers the result that a low non-contingent level of exogenous risk is optimal, as shown in Figure 8(b).\textsuperscript{24} This is aligned with the common stabilization prescription and indicates that the misallocation cost and the negative fundamental effect of larger real fluctuations are stronger than the benefit of reducing amplification risk alone. However, the trade-off changes once real fluctuations also reduce the likelihood of large collapses. It is important to mention that, absent financial frictions, full stabilization is always the optimal policy prescription with respect to real fluctuations.

7 Conclusions

This paper presents a tractable macroeconomic framework that includes two different endogenous risks related to the financial sector: amplification of real shocks and systemic runs. The model demonstrates that economies with smaller exposure to real shocks are more prone to systemic runs, i.e., \textit{stability breeds instability}. This interaction can be exploited by using \textit{risk to control risk}, i.e., by eschewing the full stabilization prescription and permitting small disruptions in order to prevent excessive risk-taking. Welfare analysis suggests that such an approach can benefit all agents in the economy.

The analysis also illustrates important challenges in the identification of systemic risks. In particular, the threat of large collapses due to systemic runs can be present even when volatility of real and financial variables is low. In other words, systemic risk may be \textit{hidden}. The model provides guidance on how to differentiate such a regime from safe, low-volatility environments.

\textsuperscript{24}It is not optimal to set $\sigma \to 0$ because I include banks in the welfare measure and such a policy redistributes wealth toward households. If Pareto-weight on banks is zero, then it is indeed optimal to fully stabilize the economy, i.e., set $\sigma \to 0$.\textsuperscript{50}
Finally, the study of a simple and intuitive macro-prudential policy such as a leverage cap exemplifies the difficulties in controlling two different dimensions of risk at the same time, even in an economy in which the financial sector invests in a single asset. This suggests that policies that discourage agents’ risk-taking by influencing (the appropriate type of) risk in the economy rather than controlling portfolio decisions may be more successful at promoting systemic stability.
References


Appendix A: Analytical results.

**Proof Lemma 1.** Households’ problem. The HJB equation for households’ problem is

\[ \rho V(n, \eta) = \max_{c, \xi > 0} \log(c) + \mathbb{E} \left[ \frac{dV(n, \eta)}{dt} \right] + \Delta (V(n, \eta) - V(n, \eta)) \]

s.t.

\[ \frac{dn}{n} = x \Delta^k + (1 - \xi) r - \xi dt - \left[ (x - 1) \ell^d + \tau \right] dJ_t, \]

where the last term in the HJB equation is associated with idiosyncratic Poisson shock that turns the household into a bank and \( V \) represents the value function of banks. The dynamics of aggregate state \( \eta \) are taken as given by the household.

**Solution.** Take as given \( V(n, \eta) = \rho^{-1} \log(n) + \xi(\eta) \), which is proved below. Conjecture \( V(n, \eta) = \rho^{-1} \log(n) + \xi(\eta) \) with \( d\xi = \xi \Delta^d dt + \sigma^\xi \xi dZ_t + (\xi(\eta) - \xi) dJ_t \). These dynamics will need to satisfy consistency conditions with the aggregate state’s law of motion, detailed below. Given the conjecture and using Ito’s lemma, the evolution of \( V \) can be written as

\[ dV = V_n \left( \mu^n ndt + \sigma^n n dZ_t \right) + V_\xi \left( \xi \Delta^d dt + \sigma^\xi \xi dZ_t \right) + \frac{1}{2} V_{nn} (\sigma^n n)^2 dt + \frac{1}{2} V_{\xi\xi} (\sigma^\xi \xi)^2 dt \ldots \]

\[ \ldots + V_{\xi n} \sigma^n \sigma^\xi \xi n dt + [V(n, \eta) - V(n, \eta)] dJ_t \]

Then, the HJB reduces to

\[ \log(n) + \rho \xi(\eta) = \max_{c, \xi > 0} \log(c) + \frac{1}{\rho} \mu n - \frac{1}{2\rho} \sigma^2 + p \left[ \frac{1}{\rho} \log \left( \frac{\tilde{n}}{n} \right) \right] \ldots \]

\[ \ldots + p \left( \xi(\tilde{\eta}) - \xi(\eta) \right) + \Delta \left[ \xi(\eta) - \xi(\eta) \right] + \mu \xi \]

where the tilde notation refers to the value of a variable just after a systemic run. Note that decisions only affect the first four terms on the RHS so, if the conjecture is correct, the maximization is equivalent to optimize over these terms. First-order conditions deliver \( \xi = \rho n, \ell = [\Phi(1/q)]^+, \) and

\[ \mu^R - r \leq \rho (\sigma + \sigma q)^2 + p \frac{(\ell q - \ell^d)}{1 - \ell^d - \bar{x}(\ell q - \ell^d) - x}, \quad (29) \]

which holds with equality for \( x > 0 \). Solving the latter for \( x \) delivers two roots but only the smaller one, explicitly displayed in equation (13) in the main text, is a solution to the maximization problem (the other violates \( n \geq 0 \) after a run). Second-order conditions are satisfied and ensure a unique global maximum.
Replacing these optimal decision on the HJB equation,
\[ \rho \xi = \log(\rho) + \frac{1}{\rho} \left[ x \left( \mu^R - r \right) + r - \rho \right] - \frac{1}{2\rho} \sigma^2 (\sigma + \sigma^q)^2 \]
\[ \ldots + \frac{1}{\rho} \log \left( 1 - \epsilon^d - \xi(\eta) \right) + \xi(\eta) - \xi(\eta) \right] + \lambda \left[ \xi(\eta) - \xi(\eta) \right] + \mu_{\xi} \xi, \]  
(30)
where \( x \) is given by equation (13) and \( \mu_{\xi} \) satisfies
\[ \mu_{\xi} \xi = \xi(\mu^\eta \eta) + \frac{1}{2} \xi^2 (\sigma^\eta \eta)^2 \]
which follows from Ito’s lemma. Since equation (30) does not depend on individual state \( n \), function \( \xi(\eta) \) can be chosen to ensure it is always satisfied. This verifies the conjecture.

I have characterized the solution of the HJB equation but I still need to verify that process
\[ G_t \equiv \int_0^t e^{-\rho s} \log(\xi_s) ds + e^{-\rho t} V(n_t, \eta_t) \]
is martingale under the optimal policy and a supermartingale under any other admissible policy. I omit this technical discussion.

The relation of \( x \) with respect to arrival rate of a run \( p \) and total diffusion volatility \( \sigma + \sigma^q \) follow from (29). Note that the RHS is strictly increasing in \( x \) (for solutions that imply positive net worth after the run \( \tilde{n} \)), run intensity \( p \), diffusion volatility \( |\sigma + \sigma^q| \), and capital loss \( \ell^q \). So \( x \) is decreasing in the mentioned variables.

**Proof Lemma 2.** Banks’ problem. The HJB equation for the banks’ problem is
\[ \rho V(n, \eta) = \max_{\epsilon, t \geq 0} \log(\epsilon) + \mathbb{E} \left[ \frac{dV(n, \eta)}{dt} \right] + \lambda (V(n, \eta) - V(n, \eta)) \]
s.t.
\[ \frac{dn}{n} = xdr^k - (x - 1) r(x) dt - \frac{e}{n} dt + \mathbb{1}_{\{\ell^d(x) > 0\}} (x\ell^d - 1 + \tau) dJ_t - \mathbb{1}_{\{\ell^d > 0 \& \ell^d(x) < 0\}} \phi dt \]
where the last term in the HJB equation is associated with idiosyncratic Poisson shock that turns the bank into a household and \( V \) represents the value function of households. The law of motion of aggregate state \( \eta \) is taken as given by the banker.

**Solution.** Take as given \( V(n, \eta) = \rho^{-1} \log(n) + \xi(\eta) \). Conjecture \( V(n, \eta) = \rho^{-1} \log(n) + \xi(\eta) \) with \( d\xi = \mu\xi dt + \sigma\xi dZ_t + (\xi(\eta) - \xi) dJ_t \). These dynamics will need to satisfy consistency conditions with the aggregate state’s law of motion, detailed below. Given the conjecture and using Ito’s
lemma, the evolution of $V$ can be written as
\[
\begin{align*}
\ dV &= V_n \left( \mu^n ndt + \sigma^n n dZ_t \right) + V_\xi \left( \mu^\xi \xi dt + \sigma^\xi \xi dZ_t \right) + \frac{1}{2} V_{nn} \left( \sigma^n n \right)^2 dt + \frac{1}{2} V_{\xi \xi} \left( \sigma^\xi \xi \right)^2 dt \\
& \quad + V_{n \xi} \sigma^n \sigma^\xi n \xi dt + \left[ V(\tilde{n}, \tilde{\eta}) - V(n, \eta) \right] dJ_t 
\end{align*}
\]

Then, the HJB reduces to
\[
\log(n) + \rho \xi(\eta) = \max_{c, t \geq 0, x \geq 0} \left[ \log(c) + \frac{1}{\rho} \mu_n - \frac{1}{2\rho} \sigma_n^2 + p \left[ \frac{1}{\rho} \log \left( \frac{\tilde{n}}{n} \right) \right] \right. \\
& \quad \left. + p \left[ \xi(\tilde{\eta}) - \xi(\eta) \right] + \lambda \left[ \xi(\eta) - \xi(\eta) \right] + \mu \xi \right]
\]

Note that decisions only affect the first four terms on the RHS. So, if the conjecture is correct, the maximization is equivalent to optimize over these terms. First-order conditions for consumption and investment are symmetric to households’ i.e., $c = \rho n$, and $\nu = \Phi(1/q)|^\dagger$. There are two local maxima with respect to $x$, one safe that satisfies $x^s \leq (\ell^q)^{-1}$ and another risky that satisfies $x^r > (\ell^q)^{-1}$. The optimal condition for $x^s$ is
\[
\mu^R - r(x^s) \leq x^s (\sigma + \sigma_q)^2 + p \frac{\ell^q}{1 - x^s \ell^q} ,
\]
which holds with equality for $x^s > 0$. The optimal condition for $x^r$ is
\[
\mu^R - r(x^r) \leq x^r (\sigma + \sigma_q)^2 + (x^r - 1) r' (x^r) ,
\]
which holds with equality for $x^r > (\ell^q)^{-1}$. The risky solution is optimal if
\[
(x^r - x^s)(\mu^R - r^f) - (x^r - 1)(r(x^r) - r^f) + \phi \geq \frac{1}{2} \left( (x^r)^2 - (x^s)^2 \right) (\sigma + \sigma_q)^2 + p \left( \log (1 - x^s \ell^q) - \log(\tau) \right) .
\]

The assumption is that the cost of underperforming $\phi$ is sufficiently large for this condition to be satisfied and therefore ensure that equilibrium is symmetric (avoid that a fraction of banks choose the safe leverage while the others choose the risky one). The optimal risky leverage condition can be rearranged to render equation (16) using the pricing condition for deposits (14).

Replacing these optimal decisions on the HJB equation,
\[
\rho \xi = \log(\rho) + \frac{1}{\rho} \left[ x \left( \mu^R - r \right) + r - \rho \right] - \frac{1}{2\rho} x^2 (\sigma + \sigma_q)^2 + p \left[ \frac{1}{\rho} \log(\tau) + \xi(\tilde{\eta}) - \xi(\eta) \right] \\
& \quad + \lambda \left[ \xi(\eta) - \xi(\eta) \right] + \mu \xi ,
\]

---

Note: The text contains mathematical expressions and equations that are typical in economics and finance contexts, specifically dealing with dynamic optimization and hedging strategies. The context suggests a discussion on optimal control theory, possibly in the context of financial markets or economic decision-making. The equations and conditions provided are typical in advanced economic models where optimal decisions are analyzed with respect to various state variables and constraints.
where \( x \) is given by equation (16) and \( \mu_x \) satisfies

\[
\mu_x = \xi_{\eta} (\mu^\eta \eta) + \frac{1}{2} \xi_{\eta \eta} (\sigma^\eta \eta)^2,
\]

which follows from Ito’s lemma. Since equation (33) does not depend on individual state \( n \), function \( \xi(\eta) \) can be chosen to ensure it is always satisfied. This verifies the conjecture. I omit the technical discussion about the verification argument.

**Proof Lemma 3.** Aggregate state dynamics. The proof follows directly from applying Ito’s formula for processes with jumps to \( \eta (N, \overline{N}) = N / (N + \overline{N}) \) and the law of motions of \( N \) and \( \overline{N} \).

**Proof Lemma 4.** Scaled value functions. Households. The value function of a household with net worth \( n \) can be written as

\[
V(n, \eta) = \rho^{-1} \log(n) + \xi(\eta)
= \rho^{-1} \log \left( \frac{n}{N} (1 - \eta) q K \right) + \xi(\eta)
= \rho^{-1} \left[ \log \left( \frac{n}{N} \right) + \log(K) \right] + \frac{1}{\rho} \log \left( \frac{1 - \eta q}{1 - \eta} \right)
\]

where the first line follows from Lemma 1, the second from \( N = (1 - \eta) q K \), and the third just illustrates how \( \psi(\eta) \) is defined in terms of equilibrium objects already derived. Using this definition to replace \( \xi(\eta) \) with \( \psi(\eta) \) and market clearing conditions, equation (30) becomes

\[
\rho \psi(\eta) = \log(\rho(1 - \eta) q) + \frac{1}{\rho} \left[ \Phi(\epsilon) - \delta - \frac{1}{2} \sigma^2 + \lambda \left( \frac{\eta}{1 - \eta} \right) \right] + \psi_{\eta},
\]

\[
... + p(\eta) [\psi(\overline{\eta}) - \psi(\eta)] + \lambda \left[ \frac{1}{\rho} \log \left( \frac{\eta}{1 - \eta} \right) + v(\eta) - \psi(\eta) \right]
\]

where

\[
\mu_{\psi} = \psi_{\eta} (\mu^\eta \eta) + \frac{1}{2} \psi_{\eta \eta} (\sigma^\eta \eta)^2
\]

Alternatively, this expression can be derived directly by conjecturing directly (28) and replacing this together with optimal decisions in the HJB equation. In this approach, it is necessary to take into account that the share of individual wealth to sector’s wealth has non-trivial dynamics due to the idiosyncratic shocks that switch agents’ type. I verified my results using both procedures.

**Banks.** Derivation is symmetric using value function in Lemma 2 and \( N = \eta K q \). In this case, \( \psi(\eta) \equiv \rho^{-1} \log(\eta q) + \xi(\eta) \). Using this definition to replace \( \xi(\eta) \) with \( \psi(\eta) \) and market clearing
conditions, equation (33) becomes

\[ \rho v(\eta) = \log(\rho \eta q) + \frac{1}{\rho} \left[ \Phi(\lambda) - \delta - \frac{1}{2} \sigma^2 + \lambda - \frac{1}{\eta} \left( \frac{1 - \eta}{\eta} \right) \right] + v\mu_v \ldots \tag{35} \]

\[ \ldots + p(\eta) [v(\eta) - v(\eta)] + \lambda \left[ \frac{1}{\rho} \log \left( \frac{\eta}{1 - \eta} \right) + v(\eta) - v(\eta) \right] \]

where

\[ \mu_v v = v_\eta (\mu^\eta) + \frac{1}{2} v_{\eta\eta} (\sigma^\eta)^2 \]

The alternative procedure described for value function of households also applies for banks.

**Proof Lemma 5. Output growth and capital returns.** Aggregate demand per unit of capital is

\[ y^d(q) = \frac{C + C + \iota K^b + \iota K}{K} \]

\[ = \rho q + \iota(q) \]

where the second line follows from optimal consumption and investment decisions. Then, the result follows directly from applying Ito’s lemma to function \( y^d(q) \).

**Proof Proposition 1. Invariance to run risk.** I characterize functions \( \{q(\eta), \psi(\eta), \sigma^q(\eta)\} \) independently from selection mechanism \( \Gamma(\cdot) \).

First, consider the case in which households hold some capital, \( \psi < 1 \). In this case, expression (12) holds with equality. Subtract the latter from (15) and use equation (14) to find the explicit expression for \( r'(x) \). This yields

\[ (x - x) (\sigma + \sigma^q)^2 = \frac{a - \bar{a}}{q}. \tag{36} \]

Equations (24) and (18) render

\[ \sigma + \sigma^q = \frac{\sigma}{1 - \frac{\psi_\eta}{q} \eta(1 - \eta)(x - x)}. \tag{37} \]

Replacing \( x = \psi/\eta \), and \( x = (1 - \psi)/(1 - \eta) \) into these two conditions and using goods market clearing (21), I find an ODE for \( q(\eta) \), i.e.,

\[ \left( \frac{\psi(q) - \eta}{\eta(1 - \eta)} \left( 1 - \frac{\psi_\eta}{q} (\psi(q) - \eta) \right) \right)^2 = \frac{a - \bar{a}}{q} \tag{38} \]

where \( \psi(q) \equiv [q (\iota(q) + \rho) - \bar{a}] / (a - \bar{a}) \). The boundary condition for this first-order ODE follows from goods market clearing condition evaluated at \( \eta = 0 \), i.e.,

\[ q(0) [\iota(q(0)) + \rho] = \bar{a}. \]

57
The latter equation has a unique solution due to the concavity of $\Phi(\cdot)$. This ODE describes the solution for $q(\eta)$ as long as respects the non-negativity constraint for households’ capital holdings, i.e., $x \geq 0$.

Second, consider the case in which households hold no capital, i.e., $\psi = 1$. In this case, goods market clearing condition implies $q(\eta) = \bar{q}$ which is defined by

$$\bar{q}[\psi(\bar{q}) + \rho] = a.$$ 

Then, $q_0 = 0$ and equation (37) implies $\sigma^q = 0$. This is a solution as long as households do not desire to hold any capital, i.e.,

$$x\sigma^2 < \frac{a - a}{q}.$$ 

Note that in both cases the equations that characterize $q(\eta)$ are independent from selection mechanism $\Gamma(\cdot)$. The same applies for $\{\psi, \sigma^q\}$ which are related to $q$ through static conditions presented above. Also, the following endogenous objects do not appear on the ODE or boundary condition: arrival rate of runs, $p$, losses conditional on runs, $\ell^q$ and $\ell^d$, and the state of the economy after the run $\tilde{\eta}$.

**Proof Proposition 2.** Amplification region. The first step of the proof is to show that $q(\eta)$ is a weakly increasing function. The ODE (38) can be written as

$$q_\eta = \frac{q}{\psi(q) - \eta} \left( 1 \pm \sigma \sqrt{\frac{\psi(q) - \eta}{\eta(1 - \eta)(a - \bar{a})}} \right).$$

Equation (36) implies that $\psi > \eta$. If both agents are holding capital, they price it the same. For this to be possible banks who earn a larger return on capital need to have a larger exposure. Then, if we consider the solution with the plus sign, which is associated with $\sigma + \sigma^q < 0$, we have that $q_\eta > 0$ immediately. However, this paper (and the literature) does not focus on this solution. For a discussion about the type of equilibria that includes this solution, see Mendo [2018].

For the solution with the minus sign, which is characterized by $\sigma + \sigma^q > 0$, a sufficient condition for $q_\eta > 0$ is the presence of amplification, i.e. $\sigma^q > 0$. This is evident from writing the ODE as

$$q_\eta = \frac{q}{\psi(q) - \eta} \left( 1 - \frac{\sigma}{\sigma + \sigma^q} \right).$$

Since, in this paper, I only consider solutions with $\sigma^q \geq 0$, we have that $q_\eta \geq 0$. In fact, the inequality must be strict, i.e. $\sigma^q > 0$, because the original ODE (38) is not satisfied by a constant $q$.

Then, I have proved that $q_\eta > 0$ if $\psi < 1$ (for the equilibria of interest). For the region in which $\psi = 1$, we know that $q = \bar{q}$ and $q_\eta = 0$. See proof of proposition 1.

The second step is to show that $\exists \eta^\psi \in (0, 1]$ s.t. $\psi < 1$ if and only if $\eta < \eta^\psi$.
\( \psi'(\eta) \) has the same sign as \( q'(\eta) \) so \( \psi_\eta \geq 0 \). This follows from goods market clearing. We also know that \( \psi \leq 1 \) because no agent can short capital. Therefore, if \( \psi(\eta_{\alpha}) = 1 \), then \( \psi(\eta) = 1 \) for all \( \eta > \eta_{\alpha} \). I define the minimum \( \eta \) such that banks manage all capital as \( \eta^\psi \), i.e., \( \eta^\psi = \inf_{\eta \in \Psi} \eta \) where \( \Psi = \{ \eta \in [0, 1] \text{ s.t. } \psi(\eta) = 1 \} \). If \( \eta < \eta^\psi \), then the ODE above holds and \( q_\eta > 0 \). Note that \( \eta^\psi > 0 \) since without net worth banks cannot hold any capital \( \psi(0) = 0 \). The proof is completed by noting the following. First, \( q_\eta = 0 \) implies \( \sigma^q = 0 \). Second, \( q_\eta > 0 \) implies \( \sigma^q > 0 \).

**Proof Proposition 3.** Run vulnerability region. Proposition 1 establishes that \( \{q(\eta), \psi(\eta)\} \) are independent from selection mechanism \( \Gamma(\cdot) \). Losses for depositors \( \ell^d \) depend on banks’ leverage \( x = \psi/\eta \), capital price \( q(\eta) \), and liquidation value \( q(\tilde{\eta}) \). Then, it is only left to show that liquidation value does not depend on \( \Gamma(\cdot) \). This follows directly from equation (20), which determines \( \tilde{\eta} \) given \( q(\eta) \) and transfer policy \( T \). Note that \( \{\ell^d, \ell^v\} \) only depend on \( \{q(\eta), \psi(\eta), q(\tilde{\eta})\} \) so they are also independent from selection mechanism.

**Proof Corollary 1.** Propositions 1 shows that \( \{q, \psi\} \) are continuous functions of state \( \eta \). Then, equation (20) implies function \( \tilde{\eta} \) is also a continuous function if policy \( T(\eta) \) is continuous. Consider the uncapped potential losses \( \hat{\ell}^d = 1 - \frac{q(\tilde{\eta})}{q} \left( \frac{x}{x-1} \right) \) which are equivalent to \( \ell^d(\eta) \) except they can be negative. It follows that \( \hat{\ell}^d(\eta) \) is a continuous function.

Note that \( \lim_{\eta \to 0} \hat{\ell}^d(\eta) < 0 \) because there are no potential losses since \( q(\tilde{\eta}) \geq q(0) \) and \( \lim_{\eta \to 1} \hat{\ell}^d(\eta) < 0 \) because banks don’t use leverage anymore as \( \eta \to 1 \). Then, the result follows directly from the continuity of \( \hat{\ell}^d(\eta) \).

**Proof Corollary 2.** Given equation (20), a constant \( T \) implies a constant liquidation value \( q(\tilde{\eta}) \). Then, results follow from the fact that \( q(\eta) \) and \( \psi(\eta) \) are weakly increasing in \( \eta \) as shown in proposition 2.

**Proof Corollary 3.** Note that for \( \eta > \eta^\psi \), the expression for depositors’ losses is

\[
\ell^d = \left[ 1 - \frac{q(\tilde{\eta})}{q} \left( \frac{1}{1-\eta} \right) \right]^+
\]

which is strictly decreasing in \( \eta \) for a constant \( q(\tilde{\eta}) \). Therefore, if \( \ell^d(\eta^\psi) = 0 \), then \( \ell^d = 0 \) for all \( \eta > \eta^\psi \) and there is no hidden risk region. Conversely, if \( \ell^d(\eta^\psi) > 0 \), by the continuity of \( \ell^d \), there exists a \( \delta_{\alpha} > 0 \) s.t. \( \ell^d(\eta) > 0 \) for all \( |\eta - \eta^\psi| < \delta_{\alpha} \).

**Proof Lemma 6.** Consider the following geometric process (with constant coefficients)

\[
\frac{dw_t}{w_t} = \mu_w dt + \sigma_w dZ_t
\]

If the associated stationary distribution is not degenerate, it satisfies satisfies \( g(w) = c_w w^{2 \left( \frac{\mu_w}{\sigma_w} - 1 \right)} \) where \( c_w \) is a scaling constant. This follows directly from the Kolmogorov Forward Equation (KFE). For details, see for example BruSan.
I show that the dynamics of the aggregate state \( \eta \) converge to a geometric process as \( \eta \to 0 \) and that the associated drift and volatility satisfy \( \lim_{\eta \to 0} \mu_\eta > \lim_{\eta \to 0} \sigma_\eta \). This implies that stationary distribution \( g(\eta) = c_\eta \eta \left( \frac{\sigma_\eta^2}{\eta} - 1 \right) \) has mass zero at \( \eta = 0 \) since exponent is positive. I consider the model 1 without runs, i.e. \( \Gamma(\cdot) \equiv 0 \). The existence of a non degenerate stationary distribution follows from the idiosyncratic Poisson shocks that switch agents’ type.

Using lemma 3, market clearing and first order conditions, the geometric drift and diffusion of \( \eta \) can be written as

\[
\mu_\eta = \sigma_\eta^2 + (1 - \psi) \left( \frac{a - a}{q} \right) + \Lambda(1/\eta - 1) - \lambda
\]

\[
\sigma_\eta = \left( \frac{\psi}{\eta} - 1 \right) (\sigma + \sigma_\eta)
\]

I now work on the limits as \( \eta \to 0 \) of the necessary expressions. First, note that equation (37) and the fact that \( \lim_{\eta \to 0} \psi \to 0 \) implies that \( \lim_{\eta \to 0} \sigma^\eta \to 0 \). From goods market clearing, we have that \( \lim_{\eta \to 0} q = a \) where

\[
q \left( \nu(q) + \rho \right) = a
\]

It is left to define the limit of \( \psi/\eta \). Equation (36) can be written as

\[
\frac{a - a}{q} = \left( \frac{\psi}{\eta} - 1 \right) (\sigma + \sigma_\eta)^2
\]

so taking the limit as \( \eta \to 0 \), we have that

\[
\lim_{\eta \to 0} \frac{\psi}{\eta} = \frac{a - a}{q \sigma^2} + 1
\]

Therefore,

\[
\lim_{\eta \to 0} \sigma_\eta = \frac{a - a}{q \sigma}
\]

\[
\lim_{\eta \to 0} \mu_\eta = \left( \frac{a - a}{q \sigma} \right)^2 + \left( \frac{a - a}{q} \right) + \Lambda \left[ \lim_{\eta \to 0} \frac{1}{\eta} - 1 \right] - \lambda
\]

Clearly, if \( \Lambda > 0 \), we have that \( \lim_{\eta \to 0} \mu_\eta > \lim_{\eta \to 0} \sigma_\eta \) and the stationary distribution has mass zero at \( \eta = 0 \).

**Proof Proposition 4.** Exogenous risk and total diffusion risk. First, I show that \( q(\eta; \sigma) \) is weakly decreasing in \( \sigma \) for \( \eta \in [0, 1] \). I only consider the solution with \( \sigma + \sigma^\eta > 0 \), i.e. the one associated with the minus sign for ODE (39). Consider \( \sigma^H > \sigma^L \). Since for \( \eta = 0 \), the price of capital does not depend on \( \sigma \), we have \( q(0; \sigma^L) = q(0; \sigma^H) \). I prove that \( q(\eta; \sigma^L) \geq q(\eta; \sigma^H) \) by contradiction. Assume \( q(\hat{\eta}; \sigma^L) < q(\hat{\eta}; \sigma^H) \) for some \( \hat{\eta} \). Since both functions are continuous in \( \eta \), this means that
\[ \exists \tilde{\eta} \in (0, \tilde{\eta}) \text{ such that } q(\tilde{\eta}; \sigma^L) = q(\tilde{\eta}; \sigma^H) \text{ and } q_\eta(\tilde{\eta}; \sigma^L) < q_\eta(\tilde{\eta}; \sigma^H). \] This is a contradiction since \( q_\eta \) is decreasing in \( \sigma \) (and all other values are the same). Since \( \psi(\eta) \) is directly related to \( q(\eta) \) through market clearing condition, it is also weakly decreasing in \( \sigma \).

To prove that \( \sigma + \sigma^q \) is weakly increasing in \( \sigma \) for every \( \eta \), note that equation (36) can be written as

\[ (\sigma + \sigma^q)^2 = \frac{\eta}{\psi - \eta} \left( \frac{a - \bar{a}}{\bar{q}} \right). \]

Then, the result follows from \( q \) and \( \psi \) being weakly decreasing in \( \sigma \). Note this implies that the RHS of the latter equation weakly increasing in \( \sigma \).

**Proof Proposition 5.** Exogenous risk and run risk. It is sufficient to prove that depositors losses are weakly decreasing in \( \sigma \) for a given liquidation price, which I call \( q^{liq} \). Then, losses can be written as

\[ \ell^d = \left[ 1 - \frac{q^{liq}}{q} \left( \frac{1}{1 - \eta/\psi} \right) \right]^+. \]

The result follows from \( q \) and \( \psi \) being weakly decreasing in \( \sigma \), which is proved in proposition 4. Lower capital price and lower leverage shield the economy from runs.

**Proof Corollary 4.** The follows from noting that the ODE for \( q(\eta) \), i.e., equation (39), converges to

\[ q_\eta = \frac{q}{\psi(q) - \eta} \]

when \( \sigma \to 0 \). The relevant boundary conditions \( q(0) \) is the same as before. The solution to this differential equation reaches \( \psi = 1 \) at \( \eta^+ > 0 \) because \( q(0) < \bar{q} \) where \( \bar{q} \) is the capital prices associated with \( \psi = 1 \).

**Proof Lemma 7.** The equilibrium with complete markets is standard. I provide a brief overview of the equilibrium and focus on the welfare result. The economy collapses to a representative single agent model with production technology \( y_t = ak_t \), i.e., all capital is managed by banks and risk is perfectly shared among agents. Since all agents have the same preferences, perfect risk sharing implies symmetric positions (no deposits). Capital price \( \bar{q} \) solves \( \bar{q}(\epsilon(\bar{q}) + \rho) = a \).

The single agent problem can be written as

\[ \rho V(k) = \log(c) + \mathbb{E}[dV(k)] \]

s.t. (1). It is straightforward to guess and verify that the value function can be written as

\[ V(k) = \rho^{-1} \log(\bar{q}k) + \bar{v} \]

where

\[ \bar{v} = \rho \left( \log(\rho) + \Phi(\epsilon(\bar{q}) - \delta - \frac{1}{2}\sigma^2) \right) \]
Note that value for agents is decreasing in exogenous risk $\sigma$.

**Proof Proposition 6.** When the leverage constraint is not binding the ODE that characterizes the equilibrium is still (38) for $\psi < 1$. If the leverage constraint is binding, then $\psi = \bar{L}\eta$, and capital price $q(\eta)$ can be solved from goods market clearing condition,

$$q(\psi(q) + \rho) = \bar{L}\eta(a - a) + a,$$

which delivers a strictly increasing function of $\eta$. Then, amplification risk $\sigma^q$ solves (37).

I proceed to prove the result by contradiction. Consider $\eta_1$ as the minimum $\eta$ such that $\psi(\eta; \bar{L}) > \psi(\eta)$. Given that $\psi(0; \bar{L}) = \psi(0) = 0$, by continuity, $\exists \eta_2 \in [0, \eta_1)$ such that $\psi(\eta_2, \bar{L}) = \psi(\eta_2)$ and $\lim_{\eta \downarrow \eta_2} \psi(\eta; \bar{L}) > \lim_{\eta \downarrow \eta_2} \psi(\eta)$. Since the laissez-faire leverage $x = \psi(\eta)$ is a decreasing function of $\eta$, we know it is a feasible allocation for $\eta > \eta_2$. Therefore, the characterization of both solutions in this region is given by ODE (38) with boundary condition $q(\eta_2)$. This implies $\lim_{\eta \downarrow \eta_2} \psi(\eta; \bar{L}) = \lim_{\eta \downarrow \eta_2} \psi(\eta)$ which is a contradiction. Therefore, $\psi(\eta; \bar{L}) < \psi(\eta)$ for all $\eta$.

Given the direct relation between capital price $q$ and allocation $\psi$, it follows that $q(\eta; \bar{L}) \leq q(\eta)$.

**Proof Proposition 7.** Consider the $\eta$-region in which there is misallocation in both environments and the leverage cap is not binding, i.e., $\{\psi(\eta; \bar{L}), \psi(\eta)\} \in [0, 1)$ and $x(\eta; \bar{L}) < \bar{L}$. Then, both solutions need to satisfy

$$(\sigma + \sigma^q)^2 = \left(\frac{a - a}{q}\right) \frac{\eta(1 - \eta)}{\psi - \eta}.$$ 

Then, the result follows from $\psi$ and $q$ weekly decreasing when a leverage cap in introduced, see proposition 6.

**Proof Proposition 8.** (a) Given that $q(\eta)$ weakly decreases with the introduction of a leverage cap, the $\bar{\eta}$ that solves $T(\eta) = q(\bar{\eta})\bar{\eta}$ increases. Given that $T(\eta)$ is not affected by the inclusion of the leverage cap, the new liquidation price needs to be weakly lower, i.e. $q(\bar{\eta}(\eta; \bar{L}); \bar{L}) \leq q(\bar{\eta}(\eta))$.

(b) The result follows from the fact that $q(\eta)$ weakly decreases with the introduction of a leverage cap.

**Appendix B: Numerical strategy.**

- For an updated version of the appendix, please click here.