A Welfare Analysis of the Central Bank Balance Sheet*

William Pagel[†‡]

March 2026

Abstract

During the period when interest rates were constrained at the effective lower bound, central banks relied on balance sheet expansions to set the stance of monetary policy. But in the long-run, when rates are away from their lower-bound, these balance sheet policies are not needed to provide stimulus. This paper asks: what is the socially optimal size for the central bank balance sheet in the long-run? I introduce a central bank into a model in which banks are liquidity mismatched and prone to sudden "bank runs". In this framework, the long-run provision of central bank reserves can reduce the likelihood of bank runs. However, central bank liquidity comes at the cost of crowding out more productive investments, owing to constraints on the set of securities the central bank can hold. I calibrate the model to conditions before the financial crisis, and demonstrate that the model can reproduce important non-linearities in the data on the demand curve for reserves, as reserve supply moves from scarce to abundant. I compute the socially optimal size of the central bank and find that it is optimal for the central bank to expand supply of reserves to the minimum point where reserve demand is satiated, but no further.

JEL Classification: E52, E44, E58, G21

*Any views expressed are solely those of the author and should not be taken to represent those of the Bank of England or of its policy committees. I would like to thank Özge Akinci, Fabrice Collard, Andrea Ferrero, Mike Joyce, Giovanni Lombardo, Pablo Ottonello, Ricardo Reis, Francesco Zanetti and participants at the Bank of England Agenda for Research Conference, the European Economic Association Congress 2026, and ASSA 2026, for their helpful comments on this work.

[†]University of Oxford

[‡]Bank of England

1 Introduction

During both the great financial crisis and the Covid-19 pandemic, central banks around the world purchased large volumes of assets from the private sector, financed by the creation of new central bank reserves. These actions drove reserve supply to levels which far exceeded the total amount demanded by the banking sector, referred to as a regime in which the supply of reserves is "abundant", and reserve demand "satiated" (Afonso et al. 2022). This expansion in central bank balance sheets was a necessary by-product of the unconventional monetary policy measures taken when interest rates were constrained by the effective lower bound. Since short-term rates began to rise in 2022, central banks have begun to shrink their balance sheets, and with them, the level of reserve supply (Figure 1). An open question now facing policy makers is where the process of balance sheet normalisation should stop.

Proponents of large central bank balance sheets argue that the expansive provision of central bank liquidity improves financial stability, alleviating an inefficiently low supply of ultra-liquid assets which only the central bank can provide (Greenwood, Hanson, and Stein (2018), Arce et al. (2020), Eren, Jackson, and Lombardo (2024)). These factors would suggest maintaining an abundant supply of reserves. On the other hand, critics argue that distortions created by a large central bank footprint in private markets can outweigh these benefits, at least for some levels of reserve supply (Borio (2023), Vissing-Jørgensen (2023), Bindseil (2016), Karadi, Thaler, and Tristani (2024)). This paper's key contribution is to develop a model that brings these perspectives into a unified welfare framework to characterise the socially optimal long-run level of reserve supply. The paper provides a tractable model in which the liquidity benefits of reserves emerge endogenously and match the key non-linearity in data on reserve demand. In addition, the model presents a micro-foundation of how central banks can drive capital misallocation as they expand their footprint in private markets. The optimal level of reserve supply trades off these two forces.

I begin with a model of a stylised banking sector that is exposed to endogenous liquidity panics ('bank runs'). In this framework, the central bank can improve welfare by supplying a perfectly liquid risk-free asset – reserves – which would not otherwise be available to the financial sector. In equilibrium, a larger level of long-run reserve

supply implies a positive externality: it lowers the risk of costly financial crises by ensuring the banking sector is better prepared to meet liquidity outflows. However, a larger central bank balance sheet comes at a cost: in supplying reserves through the purchase of assets from the private sector, the central bank allocates capital in a potentially sub-optimal manner, owing to constraints on the set of assets it is permitted to hold. To the extent that private markets fail to perfectly arbitrage this distortion, the central bank creates capital misallocation in the aggregate which pushes down output and weighs on welfare. This theory of the cost of the central bank represents a departure from the existing theoretical literature, which has treated these costs in a reduced form, or abstracted from the wider side effects of a large central bank footprint in financial markets.

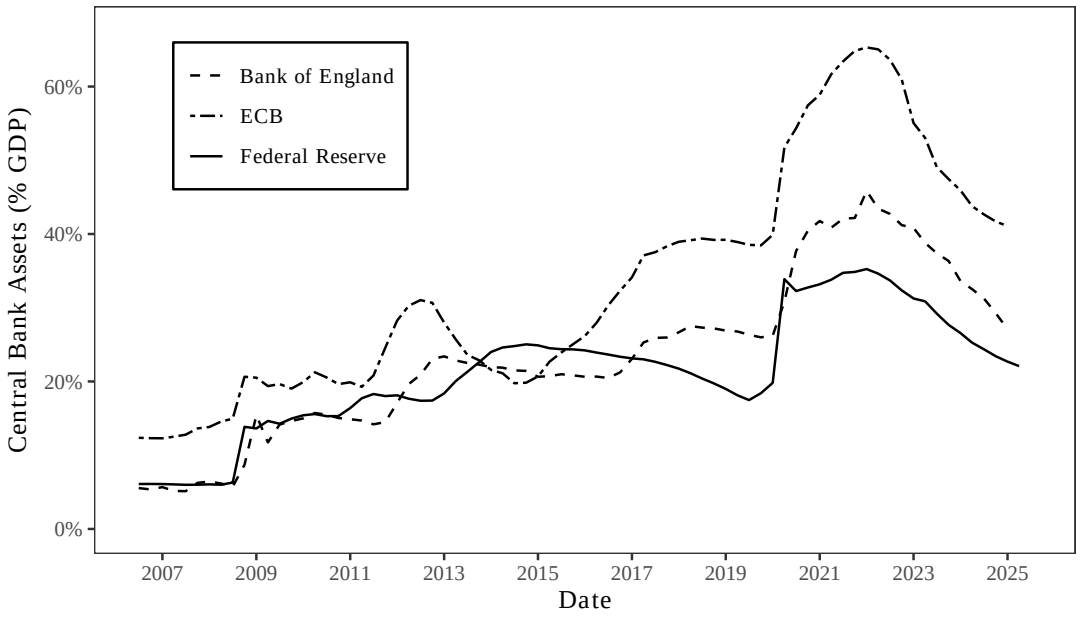


Figure 1: *Central Bank Assets as a Percentage of GDP.* Source: FRED; ECB; Bank of England.

I calibrate the model to match financial conditions in the US before the financial crisis and perform comparative statics. I evaluate the socially optimal level of reserve supply in steady state, allowing for aggregate uncertainty with respect to liquidity risk. Numerical analysis shows that it is optimal for the central bank to set reserve supply at a level high enough to satiate the demand for the reserves, but not higher. This is

known as an "ample" but not "abundant" reserve regime. Expansions in reserve supply beyond this point carry no further liquidity benefits but increase capital misallocation, decreasing welfare. I show that this result is robust to a range of plausible calibrations of the parameter governing the cost of the central bank.

Finally, I test the degree to which the model can fit empirical estimates of the reserve demand curve in the US between 2009 and 2025, based on the work of Lopez-Salido and Vissing-Jørgensen (2025). I confirm that the model provides a quantitatively good fit for the data, and explains the key non-linearity therein: a steep slope in the scarce-reserves region of the demand curve and a flat slope in the abundant-reserves region where demand is satiated.

This paper relates most closely to a new literature on the optimal size of the central bank balance sheet. Vissing-Jørgensen (2023) analyses the optimal balance sheet from the perspective of maximising the total surplus created by the availability of very liquid and safe assets ("convenience"), emphasising that the convenience provided by reserves should be offset against the convenience of the assets central banks hold to supply them. I build on this seminal study in two ways. First, I consider the general-equilibrium effects of expanding reserve supply, with a particular focus on their importance for reducing the risk of liquidity crises. Second, I explicitly model the social cost of expanding the central bank. I consider a more general cost mechanism from capital misallocation which emerges because of the strict mandates faced by central banks. This channel is consistent with Vissing-Jørgensen's argument that large central bank balance sheets can drive an inefficient scarcity in the net supply of the assets which it holds. However, in my framework this scarcity is not costly per se. Rather it pushes private markets to lend at sub-optimally low rates, distorting the economy away from the efficient allocation of capital. This treatment of capital misallocation resonates with empirical work by Acharya et al. (2025), who document how central bank asset purchases drove capital misallocation in the US by artificially subsidising bond financing.

On the theoretical side, Eren, Jackson, and Lombardo (2024) study the size and composition of the central bank balance sheet that allocates duration risk optimally in a model with canonical financial frictions, but abstract from the details of reserve

demand. They find that if the central bank faces no costs related to the size or adjustment of its balance sheet, it would optimally hold all eligible assets. Arce et al. (2020) study how reserve supply can alleviate frictions in the interbank market and compare the welfare and stabilisation properties of large balance sheet regimes against pre-crisis lean balance sheet regimes in a New Keynesian framework. They find that steady-state welfare is monotonically increasing in the size of the central bank even when the central bank must fund itself with distortionary taxes. Relative to these studies, my paper focusses on a novel cost channel, arising from capital misallocation. In my framework, the marginal benefits of reserves are endogenously decreasing, while the cost arising from capital misallocation is increasing in the size of the central bank. As a result, the optimal size of central bank balance sheet is an interior solution: at the optimum, the central bank holds a non-zero fraction of assets, but does not hold the maximum eligible amount. I show that this conclusion is robust to a wide range of assumptions on the extent to which larger central bank balance sheets generate capital misallocation.

My theoretical results match the well-established empirical finding of a strong non-linearity in the demand for reserves. Beyond some endogenous level of reserve supply, the demand for reserves is said to be "satiated": the price of reserves becomes insensitive to shocks in reserve supply as the marginal benefit to market participants of additional central bank liquidity tends to zero (Afonso et al. 2022; Borio 2023). The model produces a quantitatively accurate reserve demand function in which this non-linearity emerges endogenously. I demonstrate that this non-linearity at the point where reserve demand is satiated plays a crucial role in the determination of optimal long-run policy.

The modelling strategy of this paper is based on recent research combining the incentive compatibility-based approach to financial frictions (Kiyotaki and Moore 1997; Gertler and Kiyotaki 2010; Gertler and Karadi 2011) with the possibility of financial panics, led by Gertler and Kiyotaki (2015) and extended in Gertler, Kiyotaki and Prestipino (2016, 2020).[1] I contribute to this field by developing a method to derive a steady-state welfare function for this class of model, in the case where a

[1]For a complementary account of endogenous financial crises, emphasising the fragility of financial markets (rather than intermediaries) see Boissay et al. (2021).

bank run equilibrium exists as a sunspot, implying that the steady state allows for aggregate uncertainty with respect to liquidity risk. I show that expected steady state utility depends on the endogenous probability of runs arising, and the severity of the outcomes which occur in those runs. The social welfare function must therefore account for multiple possible equilibrium paths, including the possibility of recurrent runs, even in the deterministic steady state.

Finally, I note two areas that have received significant attention already, and which are beyond the scope of my paper. First, I do not consider the effect of supplying reserves on the implementation of monetary policy, and overnight interest rate volatility.[2] Equally, my framework abstracts from the political economy implications of central bank balance sheet risk and its implications for central bank independence from the fiscal authority.[3]

Section 2 outlines the key agents in the model, and the risk of financial panics. This micro-founds the convenience benefits of reserves and defines the demand curve for reserves. Section 3 develops a simple model of the production side of the economy in order to show how the central bank can drive capital misallocation in the aggregate. This provides a micro-foundation for the social cost of expanding the central bank balance sheet. Section 4 derives an expression for social welfare in the case where the steady state admits aggregate uncertainty with respect to the risk of financial panics. Section 5 presents the results of quantitative analysis of the model, and its performance against the data. Section 6 discusses policy implications of the findings, and Section 7 concludes.

2 Households, banks and the central bank

I begin with a general equilibrium model of a stylised banking sector that is endogenously exposed to bank runs, drawing on the key friction in Gertler and Kiyotaki (2015) (GK15). In this model, the risk of financial panics arises from a liquidity

[2]For a recent summary of this topic and its interaction with different balance sheet regimes, see Borio (2023) and Schnabel (2023).

[3]Hall and Reis (2015) and Del Negro and Sims (2015) consider the risk that central banks become insolvent as a result of losses on their expanded asset portfolios, and the implications of such risks for inflation, monetary policy, and the relationship between the central bank and the fiscal authority.

mismatch on the banking sector balance sheet, driven by the fact that banks cannot liquidate their assets at their full market value in the event of a run. When banks are sufficiently liquidity mismatched, this model will have two rational expectations equilibria: a 'normal times' equilibrium and a costly 'bank run' equilibrium. The endogenous liquidity risk (i.e. the risk of 'runs') exacerbates standard financial frictions, contracting lending, output and consumption even in the 'normal times' equilibrium. In essence, an under-supply of liquid assets drives a negative externality. As a result, this environment provides an ideal laboratory in which to study the welfare properties of central bank reserve supply. In this section I set out the key agents in the model, and the endogenous risk of financial panics.

There are three goods determined by the market allocation: a non-durable consumption good, capital, and central bank reserves.

2.1 Capital

The stock of capital is durable, and does not depreciate. Households, banks and the central bank can all hold capital, and their total capital holdings equals the supply, normalised to 1:

$$1 = K_t^b + K_t^{cb} + K_t^h \tag{1}$$

where K_t^b, K_t^{cb} and K_t^h are the capital holdings of banks, the central bank and households, respectively.

When banks intermediate capital at period t, they retain their original capital (with no depreciation) and earn an income of $Z_{t+1}K_t^b$ in the following period, $t+1$, where Z_t is the aggregate rental rate of capital:

$$K_t^b \rightarrow \begin{cases} Z_{t+1}K_t^b & \text{output} \\ K_t^b & \text{capital} \end{cases} \tag{2}$$

Households are assumed to be less efficient than banks at intermediating capital. Formally, when households hold capital at time t, they pay an upfront 'management

cost' $f(K^h)$ in that period, before the return is realised at $t+1$:

$$\left.\begin{array}{ll} K_t^h & \text{capital} \\ f(K_t^h) & \text{goods} \end{array}\right\} \rightarrow \left\{\begin{array}{ll} Z_{t+1}K_t^h & \text{output} \\ K_t^h & \text{capital} \end{array}\right. \tag{3}$$

These management costs are intended to reflect the fact that commercial banks have special expertise, relative to households, in allocating and monitoring their investments. Consequently, they are able to earn higher returns when making investments. I assume that $f(.)$ is strictly increasing and convex in capital holdings:

$$f(K_t^h) = \frac{\alpha}{2}(K_t^h)^2 \tag{4}$$

2.2 Central bank

The central bank issues reserves, m_t, by purchasing capital. The central bank balance sheet is therefore given by the identity:

$$Q_t K_t^{cb} = m_t \tag{5}$$

where Q_t is the price of capital. The central earns a return on its capital holdings, and remunerates reserves at rate R_t^m.

Any profits or losses from the central bank are rebated to the household in the form of a lump-sum transfer, τ_t:

$$\tau_t = R_t^{k,cb} Q_{t-1} K_{t-1}^{cb} - R_t^m m_{t-1} \tag{6}$$

2.3 Households

Households choose consumption and a level of savings. Households save either by lending to the bank in the form of deposits, D_t, or holding capital directly, K_t^h. Deposits are one period bonds issued by the bank. If there is no bank run, they pay a non-contingent gross rate $\bar{R}_{t+1}$. If there is a bank run, banks are fully liquidated, and depositors all face an equal haircut on their claims on the bank: they receive the

return $x_{t+1}\bar{R}_{t+1}$ on their deposits, where $x_{t+1} \in [0,1]$ is the recovery rate in the event of bank run. The deposit rate households receive can be expressed:

$$R^d_{t+1} = \begin{cases} \bar{R}_{t+1} \text{ if no bank run} \\ x_{t+1}\bar{R}_{t+1} \text{ if bank run} \end{cases} \tag{7}$$

When $x_t < 1$, depositors are not paid the full value of their deposit contract if a bank run occurs.

As discussed in the following sections, households are the final owners of banks and producing firms.[4] The household earns Π^b_t dividends from bankers and profits earned by producing firms, Π^f_t. In addition, households receive an endowment of the consumption good, $A_t W^h$ (loosely representing labour income), which is proportional to the aggregate productivity shock, A_t, and the lump-sum transfer, τ_t, rebating profits or losses from the central bank. The household takes prices, its endowment, transfers and dividends as given.

Households share consumption equally amongst members, and maximise expected utility:

$$U_t = \mathbb{E}_t \sum_{i=0}^{\infty} \ln C_{t+i}$$

subject to the budget constraint

$$C_t + D_t + Q_t K^h_t + f_h(K^h_t) = A_t W^h + R^d_t D_{t-1} + (Q_t + Z_t)K^h_{t-1} + \Pi^b_t + \Pi^f_t + \tau_t \tag{8}$$

The household funds consumption and asset holdings from income from their asset portfolio, dividends, their endowment, and the lump-sum transfer from the central bank.

Let p_t denote the probability households assign to the bank run occurring at $t+1$, and use * to denote variables in the bank run state. The households' first order

[4] My treatment of bankers *as members of the household* mirrors Gertler and Karadi (2013). See the following sub-section for more details.

condition for deposits implies that:

$$\bar{R}_{t+1} = \left[(1-p_t)\mathbb{E}_t\Lambda_{t,t+1} + p_t\mathbb{E}_t(\Lambda^*_{t,t+1}x_{t+1})\right]^{-1} \tag{9}$$

where the stochastic discount factor takes different values in the normal and bank run states:

$$\Lambda_{t,t+1} = \beta\frac{C_t}{C_{t+1}} \quad \text{and} \quad \Lambda^*_{t,t+1} = \beta\frac{C_t}{C^*_{t+1}}$$

It is informative to compare $\bar{R}_{t+1}$ with the standard, unconditional risk-free rate R^f_{t+1}, given by:

$$R^f_{t+1} = \left[(1-p_t)\mathbb{E}_t\Lambda_{t,t+1} + p_t\mathbb{E}_t\Lambda^*_{t,t+1}\right]^{-1} \tag{10}$$

When the recovery rate equals 1, the deposit rate is identical to the risk-free rate. If the recovery rate falls below 1 ($x_t < 1$) and the probability of a bank run occurring is positive, $p_t > 0$, households demand a risk premium in the return on their deposits to compensate them for the risk of a bank run such that $\bar{R}_{t+1} > R^f_{t+1}$.

The return households earn on direct capital holdings will also depend on whether or not the economy enters the bank run equilibrium in the subsequent period. In the bank run equilibrium, the price of capital falls to its 'liquidation price', Q^*_{t+1}, as banks have to fire-sell the assets on their balance sheet to meet deposit withdrawals. I discuss how to determine the liquidation price of capital in Section 2.5.

Households' first-order condition for capital is:

$$1 = \mathbb{E}_t\left((1-p_t)\Lambda_{t,t+1}\underbrace{\frac{Q_{t+1}+Z_{t+1}}{Q_t+\alpha^h K^h_t}}_{R^{k,h}_{t+1}} + p_t\Lambda^*_{t,t+1}\underbrace{\frac{Q^*_{t+1}+Z_{t+1}}{Q_t+\alpha^h K^h_t}}_{R^{k,h*}_{t+1}}\right) \tag{11}$$

where $\alpha^h K^h_t$ captures the marginal cost households must pay for holding capital directly. Whenever households retain some direct holdings of capital, this firstorder condition will play a role in determining the market price of capital. This means that asset prices will be decreasing in the marginal cost of holding capital directly, $\alpha^h K^h_t$. As the following section will show, this mechanism will be crucial to driving a fall in

asset prices during bank runs, in which households must suddenly absorb all capital previously held in the banking sector.

2.4 Banks

Banks issue deposits to fund their investment in capital and reserves. Banks will face an endogenous leverage constraint limiting the amount of deposits they can issue. As is common in the literature on financial frictions, I assume banks are not infinitely lived. This prevents them from saving sufficient funds such that the leverage constraint never binds. Bankers stochastically exit the market with probability $1-\sigma$, such that on average a banker expects to operate their bank for $\frac{1}{1-\sigma}$ periods. As a result, $(1-\sigma)f$ bankers exit and pay their retained earnings back to the household. Exiting bankers are replaced one-for-one with new entering bankers such that the total number remains constant. The household provides new bankers with a one-off start-up fund, $\frac{w^b}{(1-\sigma)f}$, per new banker, amounting an aggregate value of w^b.

2.4.1 Bankers' optimisation problem

Bankers maximise the expected dividend paid to the household when they exit the market, equal to their net worth, n_t, upon exit. The bankers' value function can be written:

$$\begin{aligned} V_t =&(1-p_t)\mathbb{E}_t\left(\Lambda_{t,t+1}(1-\sigma)n_{t+1}+\sigma V_{t+1}\right)+ \\ &p_t\mathbb{E}_t\left(\Lambda^*_{t,t+1}(1-\sigma)n^*_{t+1}+\sigma V^*_{t+1}\right) \end{aligned} \tag{12}$$

In the event of a bank run, the bank's net worth falls to zero and the bank is forced to exit with a pay-off of zero. As a result, $n^*_{t+1}=V^*_{t+1}=0$ and (12) simplifies to:

$$V_t = (1-p_t)\mathbb{E}_t\left(\Lambda_{t,t+1}(1-\sigma)n_{t+1}+\sigma V_{t+1}\right)$$

Each period, banks issue deposit liabilities to households. They use these deposits and their net-worth to finance asset holdings in capital and central bank reserves. The bank's balance sheet is described by:

$$Q_t k^b_t + m_t = n_t + d_t \tag{13}$$

For surviving bankers, net-worth at t is pre-determined by their net interest income on their balance sheet from the previous period, given by the gross return on assets net of the interest paid on deposits. We can express n_t as a flow of funds constraint:

$$n_t = (Q_t + Z_t)k^b_{t-1} + R^m_t m_{t-1} - \bar{R}_t d_{t-1} \tag{14}$$

After issuing deposits and choosing their asset holdings, bankers can choose whether to operate honestly. If the banker operates honestly, they hold assets to their maturity at time $t+1$ and repay deposits with the promised interest rate. If the banker is dishonest, they can seize a fraction, θ, of their private sector assets and divert them to their own accounts. As is standard in the literature, I assume that the process of diverting funds is time-consuming, meaning that bankers must decide whether to operate honestly before any of the uncertainty about the following period is realised - the banker does not know the value of Z_{t+1}, or whether the households will run to the bank at $t+1$. In addition, when the bank operates dishonestly at t, the depositors can force it into bankruptcy (and therefore exit) at the beginning of period $t+1$.

Bankers operate honestly at period t if and only if the present discounted value of continuing to do so, V_t, exceeds the value of the funds it is possible to divert, $\theta Q_t K^b_t$. Depositors recognise this and only lend to the bank when they know the banker will prefer to operate honestly. As a result, the bank faces the incentive compatibility constraint:

$$\theta Q_t k^b_t \leq V_t \tag{15}$$

The banker chooses k^b_t, m_t and d_t to maximise the franchise value function in (12), subject to the incentive constraint, (15), the balance sheet condition (13), and the flow of funds constraint (14).

The bank's optimisation problem boils down to a choice of leverage, and a reserve holdings. Define the variables ϕ_t and γ_t as the ratio of bank capital holdings and reserves to net worth respectively:

$$\phi_t = \frac{Q_t k^b_t}{n_t} \tag{16}$$

$$\gamma_t = \frac{m_t}{n_t} \tag{17}$$

Combining the bank's balance sheet and flow of funds constraint, (13) and (14), the evolution of the bank's net-worth can be re-written as:

$$n_{t+1} = \left[(R^{k,b}_{t+1} - \bar{R}_t)\phi_t + (R^m_{t+1} - \bar{R}_{t+1})\gamma_t + \bar{R}_{t+1}\right] n_t \tag{18}$$

where

$$R^{k,b}_{t+1} = \frac{Q_{t+1} + Z_{t+1}}{Q_t} \tag{19}$$

It is useful to define ψ_t as the ratio of the bank's franchise value to net worth, $\frac{V_t}{n_t}$. Intuitively, ψ_t can be thought of as the shadow value of funds held within the bank per unit net worth, or equivalently, the 'Tobin's Q ratio' of the bank. Dividing equation (12) and (18) by n_t, the bank's optimisation problem can be expressed as a choice of ϕ_t and γ_t:

$$\begin{aligned} \max_{\phi_t,\gamma_t} \psi_t &= (1-p_t)\mathbb{E}_t\left\{\Lambda_{t,t+1}\left(1-\sigma+\sigma\psi_{t+1}\right)\frac{n_{t+1}}{n_t}\right\} \\ &= (1-p_t)\mathbb{E}_t\left\{\Omega_{t,t+1}[(R^{k,b}_{t+1} - \bar{R}_t)\phi_t - (R^m_{t+1} - \bar{R}_{t+1})\gamma_t + \bar{R}_{t+1}]\right\} \end{aligned} \tag{20}$$

subject to the incentive constraint which can be re-written:

$$\phi_t \leq \frac{\psi_t}{\theta} \tag{21}$$

and the deposit rate given in (9).

The banks' stochastic discount factor is

$$(1-p_t)\Omega_{t,t+1} \quad \text{where} \quad \Omega_{t,t+1} = \Lambda_{t,t+1}\left(1-\sigma+\sigma\psi_{t+1}\right) \tag{22}$$

$\Omega_{t,t+1}$ captures the present discounted shadow value of net-worth, weighted by the probability of exit, $(1-\sigma)$. In the event of a bank run, the bank defaults, with $V_t = n_t = 0$, and pays no dividend.[5] Consequently, the bank's value function only considers the non-bank run state, weighted by the probability of no run occurring at $t+1$, namely $(1-p_t)$.

[5] As I show in Section 2.5, the bank run equilibrium only exists when this is the case.

I consider a symmetric equilibrium in which all banks choose the same values of ϕ_t and γ_t. Banks will choose the maximum possible leverage subject to the incentive constraint, as long as the marginal return on leverage at the optimum is strictly positive, $\frac{\partial \psi_t}{\partial \phi_t} > 0$. In my numerical analysis, I restrict attention to a parametrisation where this condition holds, such that the bank chooses the maximum possible leverage given the incentive constraint, and equation (21) binds with equality at the optimum. This assumption simplifies computation of the non-linear solution to the model.

Banks choose reserves to maximise ψ_t, noting the marginal effect of reserve holdings on their probability of experiencing a run ($\frac{\partial p_t}{\partial \gamma_t} = p'_\gamma$), and on deposit rates ($\frac{\partial \bar{R}_t}{\partial \gamma_t} = \bar{R}'_\gamma$) through the households' first order condition on deposits in (9).[6]

2.4.2 The convenience yield on reserves

Manipulating bankers' first order condition for reserve holdings ($\frac{\delta \psi_t}{\delta \gamma_t} = 0$), we can recover an expression for the spread between the reserve rate and banks' marginal funding cost, the deposit rate:

$$\mathbb{E}\Big[\Omega_{t,t+1}(R^m_{t+1} - \bar{R}_{t+1})\Big] = \frac{(1-p_t)\mathbb{E}\Big[\Omega_{t,t+1}\,\bar{R}'_\gamma(\phi_t + \gamma_t - 1)\Big] + p'_\gamma\,\mathbb{E}\Big[\Omega_{t,t+1}\left((R^k_{t+1} - \bar{R}_{t+1})\phi_t + \bar{R}_{t+1}\right)\Big]}{1 - \left(p_t + p'_\gamma \gamma_t\right)} \tag{23}$$

When (23) is negative, banks are willing to hold reserves at below their marginal funding cost. In this case, reserves earn a 'convenience yield'.[7] This yield disappears – i.e., the spread collapses to zero – when $p'_\gamma = \bar{R}'_\gamma = 0$: the bank is fully insured and indifferent to additional reserves. Reserve demand is then satiated, and further supply expansion has no marginal price effect. Micro-founding the convenience yield in this way reveals that it comprises two distinct components.

First, *Funding cost relief*: higher reserves holdings reduce a bank's funding cost by improving their perceived liquidity position. More liquid banks pay a lower risk premium to their creditors, captured in the first term in the numerator.

[6] We have not yet defined the functional form of p_t. However, in equilibrium, p_t will be a weakly decreasing function of the bank's expected recovery rate, and therefore, depend on the composition of the bank's assets (see Section 2.5).

[7] In this model R^m is a market determined price. In this sense, the spread in (23) is best thought of as the theoretical counterpart to the (negative) spread between the Effective Federal Funds Rate (the marginal cost of obtaining reserves) and the interest paid on reserves.

Second, *Liquidity insurance*: by reducing the probability that the bank experiences a run, higher reserve holdings also increase present discounted value of future returns on capital, captured by the second term in the numerator.

The total convenience yield is the sum of these terms, divided by the change in the banks' discount factor which results from the impact of reserve holdings on p_t. Together, funding cost relief and liquidity insurance explain why banks accept a negative return on reserves: these holdings serve as both a subsidy to external financing and a guarantee of operational continuity.

2.4.3 Aggregation of the financial sector

In a symmetric equilibrium, summing over all banks we have:

$$Q_t K_t^b = \phi_t N_t \tag{24}$$

and

$$m_t = \gamma_t N_t \tag{25}$$

Summing over both surviving banks' net interest income and the endowment of newly entering banks in the absence of bank runs, the aggregate quantity of net worth, N_t, evolves according to:

$$N_{t+1} = \sigma \left[R_{t+1}^{k,b} Q_t K_t^b + R_{t+1}^m m_t - \bar{R}_{t+1} D_t \right] + w^b \tag{26}$$

Exiting bankers pay the remaining fraction of the net earnings on assets as dividends:

$$\Pi_{t+1}^b = (1-\sigma) \left[R_{t+1}^{k,b} Q_t K_t^b + R_{t+1}^m m_t - \bar{R}_{t+1} D_t \right] \tag{27}$$

2.5 Bank Runs

I now turn to the case in which bank runs exist as a rational expectations equilibrium of this model. The treatment of runs mirrors that in GK15, with the caveat that the introduction of central bank reserves will impact the likelihood and severity of runs. I

consider only runs on the entire banking system, rather than individual banks.[8]

2.5.1 Existence Conditions for a Bank Run Equilibrium

Assume that at t, the bank is solvent and no bank run has occurred. A household who deposited at the bank at $t-1$ must decide at t whether to "roll over" their deposit and leave their funds in the bank, or "run" to the bank and withdraw it either for consumption or direct investment in capital.

A bank run equilibrium exists, as a 'sunspot', if and only if the household correctly believes that the bank would not be able to repay the full value of its obligation, $\bar{R}_t D_{t-1}$, if all other depositors were to run to the bank. Under these conditions, the depositor recognises that a run by the other depositors would leave the bank with a net worth of zero. The recovery rate for depositors in a run is given by the ratio of the *liquidation* value of the bank's assets to the value of its deposits:

$$x_t = \min\left[1, \frac{(Q_t^* + Z_t)K_{t-1}^b + R_t^m m_{t-1}}{\bar{R}_t D_{t-1}}\right] \tag{28}$$

where Q_t^* is the *liquidation* price of capital, defined in the following section. The recovery rate is bounded at 1: banks pay at most the promised value of the deposit contract. When the recovery rate is below 1, the net worth of a bank would be wiped out in the event of a run. Under this condition, it is rational for a depositor to run to the bank if and only if they expect all others to do the same. So, the bank run equilibrium exists as a sunspot whenever the following condition is satisfied:

$$x_t < 1 \tag{29}$$

The recovery rate is determined by endogenous variables, entailing that the possibility of bank runs depends on macro-economic conditions, and the reserve policy of the central bank.

[8] In the absence of any friction, individual bank runs would simply lead depositors to move their funds from one institution to another. This would have no aggregate effects in the representative agent framework above.

2.5.2 The liquidation price of capital

New banks cannot enter in the period of a run. This entails that if a bank run occurs, the banks pay their liabilities by selling off all their assets to households (rather than to new banks).[9] Consequently, households must purchase the entire stock of banks' capital in the event of a run, entailing that:

$$K_t^{h*} = 1 - K^{cb} \quad \& \quad K_t^{b*} = 0 \tag{30}$$

Using this fact, the liquidation price of capital is determined by the household Euler equation for capital, (11). Given that the no new banks enter, the probability of a run occurring in the subsequent period is zero. So, from equation (11), the liquidation price of capital is given by:

$$Q_t^* = \mathbb{E}_t \left(\Lambda_{t,t+1}^* (Q_{t+1} + Z_{t+1}) \right) - \alpha^h (1 - K^{cb}) \tag{31}$$

The liquidation price is strictly lower than the market value of capital, whenever banks hold capital. This drives the liquidity mismatch in bank's balance sheets: bank assets cannot be liquidated at their full market price. The liquidation price (and its gap relative to Q_t) is therefore an index of liquidity conditions in the market.

The liquidation price is decreasing in the size of the marginal household management cost, $\alpha^h K^{h*}$: liquidity mismatch is worse when households are more inefficient, or when they have to absorb a large quantity of capital in a run. From (30), the latter is a direct function of central bank balance sheet policy. When the central bank balance sheet is larger, a greater proportion of capital assets are intermediated by the central bank and do not need to be absorbed by households during a bank run. So, the liquidation price which clears the market for capital during a bank run is higher (that is, banks are more liquid) when the central bank is larger.

In contrast to capital, central bank reserves are perfectly liquid. Reserves can be liquidated by the bank at full value and transferred directly to the household as 1

[9] This assumption is necessary for the run to have an aggregate impact. The qualitative results of the model are robust to longer delays. But, bank runs become more costly as the time before new banks may enter increases, since households must continue to pay large management costs for a greater period of time.

period risk-free bonds in the event of a run. I assume households hold these risk-free bonds until banks re-enter the market, when they return to deposits (whose returns weakly dominate those of reserves).

Following a run at t, aggregate bank net-worth follows

$$\begin{aligned} N_t^* &= 0 \\ N_{t+1} &= w^b + \sigma w^b \end{aligned} \tag{32}$$

After this period, aggregate banks net-worth continue to evolve according to the law of motion defined in equation (26).

2.5.3 Probability of a bank run

An additional assumption is required to endogenise the probability of runs occurring and close the model. Following GK15, I choose a simple functional form for p_t which links the probability of a bank run sunspot to the fundamental determinant of whether the bank run equilibrium exists, $\mathbb{E}_t(x_{t+1})$, and assigns a probability of zero whenever the bank run equilibrium does not exist:[10]

$$\begin{aligned} p_t &= g(\mathbb{E}_t(x_{t+1})) \\ &= 1 - \mathbb{E}_t(x_{t+1}) \end{aligned} \tag{33}$$

Linking probability of a bank run to the recovery rate will imply that the probability of runs is endogenously counter-cyclical, with the risk of runs increasing in bad times. This will amplify shocks by further tightening the financial frictions.

Note that the probability of runs will also depend on the central bank balance sheet. To the extent that increasing reserve holdings pushes up the recovery rate in the event of a run, a larger central bank balance sheet can push down p_t.

[10] This reduced form approach reproduces the qualitative result of a fully micro-founded approach in which the functional dependence of p_t is determined endogenously by introducing imperfect information in a global games approach as in Morris and Shin (1998). The piecewise linear approach is chosen to maximise transparency on the sources of non-linearity in the numerical results.

2.5.4 Determining the equilibrium

Even with a functional assumption for the probability of the sunspot bank run equilibrium, there is no criteria for setting out which equilibrium is realised. In this sense, the model remains indeterminate.

To assess the optimal steady state central bank balance sheet, I study how central bank policy affects the long-run expected utility in this model. An endogenous probability for the bank run state arising is sufficient to capture the effects of both the severity and the expected frequency of bank runs on expected utility.

As a result, I will not need to take a position on determining which equilibrium is realised at any point. I provide a full definition of the 'no bank run' equilibrium in Appendix A.

3 Production and Capital (Mis-)allocation

This section sets out the production block of the economy, intended to be the simplest possible environment which captures how the central bank can drive the misallocation of capital in the aggregate.

This block of the model is static so for the remainder of this section, I drop time subscripts to simplify notation.

There are S sectors, indexed by $i = \{1, 2, ...S\}$, each populated by a representative firm. Each firm's production technology exhibits decreasing returns to scale, such that distortions in capital allocation can drive dispersion in the marginal product of capital and reduce aggregate TFP (Hsieh and Klenow 2009). The firm in sector i uses loans, k_i, to produce output y_i:

$$y_i = Ak_i^{\alpha}$$

As a baseline, I assume all sectors have identical productivity, given by A_t (the aggregate shock introduced in the previous section) and capital share α, such that sector i can differ from sector $j \neq i$ only in scale. The total output of the economy is given by the sum of outputs across all sectors, such that outputs from different sectors are perfect substitutes.

Firms must pay back their loans with interest, z_i, making their total cost of capital $z_i k_i$. Firms choose k_i to minimise costs, taking z_i and A as given, implying the standard first order condition:

$$z_i = A\alpha k_i^{\alpha-1} \tag{34}$$

3.1 Capital allocation

The aggregate supply of capital across sectors will depend on the combined holdings of the private sector and the central bank.

3.1.1 Central bank asset portfolio

In practice, central banks face tight political and legal constraints on the assets they are permitted to hold (Vissing-Jørgensen 2023). For instance, outside times of crisis, the Fed cannot in general purchase securities which are not government guaranteed (Federal Reserve Act 13(3)). While these asset mandates differ across jurisdictions, they all typically limit central banks to holding a narrow range of securities, relative to the universe of assets held by the private sector.[11] For the purposes of this model the precise nature of these mandates will not matter: the crucial point is that central banks cannot purchase all possible securities.

To capture this constraint in a stylised manner, I assume that all the capital holdings of the central bank are allocated to a single sector, $i = 1$. In every other way, however, sector 1 is assumed to be identical to any other: there is no presumption that the central bank necessarily invests in less productive assets than other agents.

3.1.2 Private capital portfolio

Define private capital as the sum of household and bank capital holdings:

$$K^p = K^b + K^h$$

To capture the notion that demand across asset classes is not perfectly elastic (or equivalently, that private markets face some limits to arbitrage), I introduce a rep-

[11] Vissing-Jørgensen (2023) discusses these differing mandates and studies how they affect the level of reserves which maximises net liquidity across jurisdictions.

resentative "capital manager" which allocates the private capital across production sectors, subject to portfolio frictions. This agent is a reduced form modelling device to capture how the aggregate capital allocation responds to differences in returns across asset classes, in a flexible and tractable manner.

The fact that demand in financial markets is imperfectly elastic has been shown to be an important empirical factor in explaining variation in asset prices Gabaix and Koijen (2023), and has been studied at length in the literature on quantitative easing. Inelastic demand has been explained with reference to a wide-range of factors, including the presence of institutional investment mandates, regulatory requirements, risk aversion, market segmentation, balance sheet constraints, preference for diversification or other non-pecuniary properties of various assets.[12] To represent these factors, I simply assume that the capital manager faces a cost which is a CES aggregation of their asset portfolio. Formally, if K^p denotes the total private holdings of capital to be distributed across sectors, the capital manager faces the budget constraint:

$$K^p = \left(\sum_{i=1}^{S} k_i^{\eta} \right)^{\frac{1}{\eta}}$$

where $\frac{1}{\eta - 1}$ is the elasticity of substitution between assets.

When $\eta = 1$, this cost function becomes a linear sum: the capital manager has perfectly elastic demand. In this case, it allocates its entire portfolio to the sector paying the highest return. If $\eta > 1$, the cost function is strictly convex, converging to the max operator as $\eta \to +\infty$. In this case, the capital manager faces a cost increasing in the concentration of their portfolio, creating a motive for diversification. In this case, the agent's demand for each asset becomes inelastic with respect to returns. In the analysis that follows, the crucial assumption is that $\eta > 1$: private asset demand is not perfectly elastic.

The maximisation problem can be written:

$$\max_{\mathbf{k}} \sum_{i=1}^{S} z_i k_i - Z K^p \quad \text{subject to} \quad K^p = \left(\sum_{i=1}^{S} k_i^{\eta} \right)^{\frac{1}{\eta}}$$

[12] See for instance Krishnamurthy and Vissing-Jørgensen (2012), Chen, Cúrdia, and Ferrero (2012), Del Negro et al. (2017), Vayanos and Vila (2021), Gabaix and Koijen (2023).

where Z – the return earned by households and banks – ensures the capital manager earns zero profit. For $\eta > 1$, this implies optimal portfolio choices:

$$k_i^p = \frac{z_i^\rho}{\left(\sum_{j=1}^S z_j^{\rho\eta}\right)^{\frac{1}{\eta}}} K^p \quad \text{for } i = 1, ..., S \tag{35}$$

where $\rho = \frac{1}{\eta - 1}$.

3.1.3 Equilibrium

An equilibrium consists of vectors of capital allocations $(k_1, ..., k_S)$ and returns $(z_1, ..., z_S)$ such that:

1. The total supply of private capital satisfies:

$$K^p = K^b + K^h = \left(\sum_{i=1}^S k_i^\eta\right)^{\frac{1}{\eta}},$$

2. For each sector i, the capital market clears:

$$k_i = \begin{cases} k_1^p + K^{cb} & \text{if } i = 1 \\ k_i^p & \text{if } i \neq 1 \end{cases}$$

3. The first-order conditions for firm's problem (34) and for the capital managers's portfolio choice (35) are satisfied.

3.1.4 The cost of the central bank

The unique, first-best allocation of capital maximises aggregate output subject to the resource constraint. Efficiency requires that the marginal product of capital (MPK) is equalised across all sectors. Given symmetric productivities and technologies, this condition implies that all sectors will operate at the same scale. The feasible allocation

that equalises MPKs while exhausting the available capital supply is given by:

$$\begin{aligned} k_i = k^\star = \frac{1}{S^{\frac{1}{\eta}}} \quad \forall i = 1, ..., S \\ \text{with a corresponding vector of asset returns:} \\ z_i = z^\star = A_t \alpha (k^\star)^{\alpha - 1} \quad \forall i = 1, ..., S \end{aligned} \tag{36}$$

In light of the constraints on the securities the central bank can hold, it is then straightforward to show that:

Proposition 3.1. *When $\eta > 1$, the decentralised economy achieves the first best allocation, $\{\boldsymbol{k}^\star, \boldsymbol{z}^\star\}$ if and only if $K^{cb} = 0$.*

Sketch Proof: First show that if $K^{cb} = 0$ then the economy achieves the first best allocation. Guess and verify that (36) satisfies the equilibrium conditions, and equalises the marginal product of capital across sectors. Second, show that if the equilibrium allocation of capital equalises MPKs, then the capital manager's FOC implies $K^{cb} = 0$. (See Appendix B for details)

Proposition 3.1 shows under general conditions that expanding the central bank's footprint in private markets comes at a cost: when the private sector fails to fully arbitrage the distortion created by the central bank, a central bank with constraints on the set of securities it can hold will generate some degree of aggregate capital misallocation, consistent with empirical evidence from the US (Acharya et al. 2025).

3.1.5 Aggregation

Total output is the sum of each sector's output and the endowments for households and bankers:

$$Y_t = \sum_{i=1}^{S} y_{i,t} + A_t W^h + w^b \tag{37}$$

Total profits are:

$$\Pi_t^f = \sum_{i=1}^{S} (y_{i,t} - z_{i,t} k_{i,t}) \tag{38}$$

All output is used for management costs or consumption by households:

$$Y_t = C_t + f(K_t^h) \tag{39}$$

The aggregate return on capital sets the capital manager's profits to zero such that:

$$Z_t = \frac{\sum_{i=1}^{S} z_{i,t} k_{i,t}}{K_t^p} \tag{40}$$

4 Steady State Expected Utility

In this section, I characterise the steady state of the model in which there is aggregate uncertainty with respect to whether or not the economy will enter a 'bank run' state. I demonstrate that even in the absence of exogenous shocks, agents cannot rule out the possibility of endogenous dynamics arising as a result of bank run sunspots, and I derive a method to calculate expected steady state utility which accounts for this possibility. I apply this method in the numerical analysis in Section 5.

4.1 Defining a steady state with aggregate uncertainty

I use the term 'steady state' to mean the case where the only exogenous variable (A_t) is held constant at its long run expected value, A_{ss}, and the vector of all endogenous variables, denoted X_t, is at the fixed point of the system of equations defining the 'normal times' equilibrium, set out in Appendix A. Call this vector X_{ss}. If $p_{ss} > 0$, the possibility of a bank-run means that even in steady state, households account for the possibility of endogenous dynamics arising when forming their expectations. The steady state is ex-ante uncertain with the respect to which of the two possible equilibria will be realised.

To illustrate this point, consider the system in 'steady state' at the end of some period $t-1$: $X_{t-1} = X_{ss}$ with $p_{ss} > 0$. Assume a deterministic, and constant process for exogenous shocks given by $A_t = A_{ss}\ \forall t$.

With probability $1 - p_{ss}$, the economy remains in the no bank run steady state. In this case, the values of all endogenous variables remain constant, and period t is identical to the previous period, with the same aggregate uncertainty *vis-a-vis* $t+1$.

With the complementary probability p_{ss}, a bank run occurs. In this case, all endogenous state variables are "reset" to their run values. By the assumptions laid

out in Section 2.5, these values do not depend on the values of any other variables at the end of the previous period. Therefore any period in which a bank run occurs is identical (conditional on the contemporaneous value of A_t), and independent of the entire past. In this sense the bank run state can be said to be 'memory-less'.

In the context of the steady state, any period in which a bank run occurs from steady state can be denoted by the time subscript, τ, without any loss of generality: With the value of the endogenous state variable at τ known, we have enough information to determine the entire state of the economy: by:

Proposition 4.1. *The endogenous state variable, N_τ, in a period with a bank run is independent of all previous periods, and so the state of the economy can be determined fully using the facts that:*

The banks are bankrupt, such that: $D_\tau = 0 \quad \& \quad N_\tau = 0$

The households absorb all capital not held by the central bank: $K^h_\tau = 1 - K^{cb}$

The price of capital is the liquidation price: $Q_\tau = Q^*_\tau(A_{ss})$

These values determine the full vector of endogenous variables X_τ. Note, however, that X_τ is not a fixed point of the system of equilibrium conditions. So, in the period after a bank run, $\tau + 1$, all variables jump onto the saddle path back to the system's steady state. This path, $\{X_t\}_{t=\tau+1}^{\infty}$, is the solution to the set of equilibrium conditions of the model (in Appendix A). If no further bank runs occur, the system remains on this saddle path back to the steady state, X_{ss}.

However, along this saddle path, the probability of further bank runs may be positive: the probability of bank runs continues to be determined by the method in equation (33). If another run were to occur, the variables would return to the bank run state, X_τ. By the "memoryless" property of bank runs, each time a run reoccurs, the process repeats with identical dynamics.

4.2 Computing Steady State Expected Utility

The expected utility of the representative household in the steady state averages over the infinite set of possible paths the economy can take from X_{ss} holding $A_t = A_{ss} \forall t$,

weighted by their ex-ante probability.

To derive an expression for this value which can be calculated numerically, I rely on two pieces of information to simplify the state space. First, all bank runs from steady state, and the subsequent saddle paths to X_{ss} are identical (Proposition 4.1). Second, I guess and verify numerically that:

Proposition 4.2. *Conditional on no subsequent bank runs occurring and the deterministic path of the exogenous shocks, the saddle path back from the bank run state re-converges to the steady state, X_{ss}, in a finite number of periods, T.*

Let $\mathcal{L}$ denote the ex-ante[13] expected utility of the representative household at some period t where the system begins in steady state. This value can be expressed in the form of a recursive value function:

$$\mathcal{L} = (1 - p_{ss})V_{ss} + p_{ss}V_{run,\tau} \tag{41}$$

where V_{ss} expected utility for the household (or value) if no run occurs at t, and $V_{run,\tau}$ is the value if a run does occur at t, and the economy jumps to the run state.

The value of the no run state is given by:

$$V_{ss} = \ln(C_{ss}) + \beta(1 - p_{ss})V_{ss} + \beta p_{ss}V_{run,\tau} \tag{42}$$

If the economy stays in steady state, the value function is given by the flow utility from consumption at t plus the expected value at $t + 1$, which takes expected utility of the two possible states at $t + 1$ weighted by their probabilities.

The value of the run state is given by:

$$V_{run,\tau} = \ln(C_{\tau}) + \beta(1 - p_{\tau})V_{run,\tau+1} + \beta p_{\tau}V_{run,\tau} \tag{43}$$

where $V_{run,\tau+i}$ is the value i periods after the run occurs, conditional on no further runs occurring in that time. The value in the run state is the flow utility of household consumption, conditional on a run occurring, plus the expected value at $t+1$. $V_{run,\tau+1}$

[13] Ex-ante here denotes households do not yet know whether a run will occur at t

is the value in the case where no run occurs in the following period. $V_{run,\tau}$ is the expected value in the case where a subsequent run occurs. In effect, here the economy "returns" to state τ. Each is weighted by its respective probability.

Equations (42) and (43) define a system of 2 equations in 3 unknowns. To solve these, eliminate one unknown by noting that Proposition 4.2 implies that the value to the household T periods after a run, conditional on no subsequent runs occurring during that time, is the same as the value in steady state:

$$V_{run,\tau+T} = V_{ss} \tag{44}$$

Substituting this into (43), I then solve (43) backwards from period $\tau+T$. For instance at $\tau + T - 1$:

$$V_{run,\tau+T-1} = \ln(C_{\tau+T-1}) + \beta(1 - p_{\tau+T-1})V_{ss} + \beta p_\tau V_{run,\tau} \tag{45}$$

Iterating backwards, and rearranging, $V_{run,\tau}$ can be expressed as a linear function of V_{ss}:

$$V_{run,\tau} = \Gamma_0 + \Gamma_1 V_{ss} \tag{46}$$

where Γ_0 and Γ_1 are functions of the discount factor and the saddle paths for consumption and the probability of bank runs in the periods after a run occurs, $\{C_t\}_{t=\tau}^{\tau+T}$ and $\{p_t\}_{t=\tau}^{\tau+T}$:

$$\Gamma_0 = \left(1 - \sum_{i=1}^{T} \beta^i p_{\tau+T+i-(T+1)} \prod_{j=1}^{i-1}(1 - p_{\tau+T+j-(T+1)})\right)^{-1} \sum_{i=1}^{T} \beta^{i-1} \ln C_{\tau+T+i-(T+1)} \prod_{j=1}^{i-1}(1 - p_{\tau+T+j-(T+1)})$$

$$\Gamma_1 = \left(1 - \sum_{i=1}^{T} \beta^i p_{\tau+T+i-(T+1)} \prod_{j=1}^{i-1}(1 - p_{\tau+T+j-(T+1)})\right)^{-1} \left[\beta^T \prod_{i=1}^{T}(1 - p_{\tau+T-i})\right]$$

Plugging (46) into (42) provides an expression V_{ss} in terms of steady state consumption, steady state probability of a run, Γ_0, and Γ_1:

$$V_{ss} = \frac{\ln(C_{ss}) + \beta p_{ss}\Gamma_0}{1 - \beta + \beta p_{ss}(1 - \Gamma_1)} \tag{47}$$

Substituting this result into equation (41), I obtain an expression for long run expected

utility, $\mathcal{L}$:

$$\mathcal{L} = p_{ss}\Gamma_0 + (1 - p_{ss} + p_{ss}\Gamma_1)\frac{\ln(C_{ss}) + \beta p_{ss}\Gamma_0}{1 - \beta + \beta p_{ss}(1 - \Gamma_1)}$$

Partial derivatives confirm that $\mathcal{L}$ is increasing steady state consumption, Γ_0, and Γ_1, and decreasing with p_{ss}.[14]

These results are intuitive. Γ_0 is proportional to present discounted value of flow utility of consumption along the saddle path from the run equilibrium. So, the welfare is decreasing as the severity of the impact of runs on consumption increases. Γ_1 is proportional to the probability that the economy successfully returns to steady state with no further runs (given by the product on the numerator) implying that the welfare falls as the risk of recurrent runs increases. Finally, the welfare is decreasing in the probability of bank runs.

As $p_{ss} \to 0$ and all aggregate uncertainty is removed, V_{ss} and $\mathcal{L}$ both converge to the present discounted value of utility from an infinite consumption stream of C_{ss}. This implies that when $p_{ss} = 0$, central bank balance sheet policy only effects welfare via its effect on steady state consumption.

Welfare is an increasing function of steady state consumption, given by:

$$C_{ss} = Y_{ss} - \frac{\alpha^h}{2}(K^h_{ss})^2$$

Here, the cost of a larger central bank acts directly on welfare: to the extent the central bank drives capital misallocation, it pushes down steady-state consumption.

5 Quantitative Analysis

I find the non-linear solution to the model in a setting where agents have perfect foresight with respect to the exogenous state variables (A_t, m_t). In line with the results from the previous section, I do not assume agents have perfect foresight over

[14] Parameter signs are inferred from numerical results, and therefore these statements apply in the neighbourhood of the parametrisation used in the numerical section.

whether or not runs will occur. I implement a numerical procedure which finds a full non-linear solution, needed to capture the quantitatively important non-linearities when runs occur.[15]

I outline the key steps in this procedure in Appendix C.

5.1 Calibration

I set the quarterly discount factor to 0.99, and the capital share for producers to 0.35. I set the number of producing sectors S to 2.[16]

I set $\eta = 1.02$ to match the estimated persistent "stock effect" of the Federal Reserve's LSAP1 programme in D'Amico and King (2013). They estimate that the central bank asset purchases purchases (worth roughly 8% of the US treasury market at the time) persistently pushed down long-term yields by as much as 50bps. However, the authors note that this is an upper-bound, and indeed, their estimates lie at the top of the wide range of estimates in the literature (See Table 1 in Chen, Cúrdia, and Ferrero (2012)). Smaller estimates of the yield impact would imply smaller values of η. I show that the results are robust to a wide range of η values both lower and higher than my central estimate in Section 5.4.

To calibrate the remaining parameters of the model, I match financial conditions in the US before the financial crisis, when the total central bank balance sheet was roughly 9% of banking sector assets. I estimate the parameter vector $(\sigma, \theta, \alpha^h, W^h, w^b, A)$ which minimises the Euclidean distance between a set of target moments and the steady state moments of the model, and normalises the price of capital to 1. Philippon (2015) reports that before the financial crisis, commercial banks operated with a leverage ratio of 8, and a return on bank assets of 200bps. However, institutions like shadow banks operated with considerably higher leverage ratios and lower interest rate margins (Gertler, Kiyotaki, and Prestipino 2016). Allowing for this, I target leverage ratios between 8 and 10, and asset returns of between 150 and 200bps.

[15] The derivative terms in (23), for which there are not closed forms, are approximated using finite differences.

[16] A larger S would imply that the central bank's impact on the oversupply of funds to sector 1 is relatively larger, as the efficient allocation to any given sector would be smaller in absolute terms. For simplicity, I fix $S = 2$ and consider a wide range of η values, including large values which also imply a higher social cost of the central bank, in Section 5.4.

I target a long-run probability of default equal to 4% in steady state, to match the average frequency of financial crises in Reinhart and Rogoff (2009). I target a convenience yield of reserves of 25bps, based on the average spread between EFFR and the 3-month T-bill rate, from 2003 to 2008.[17] I target an increase in credit spreads during a run of 275bps, consistent with the movement in the (option-adjusted) excess bond premium from the financial crisis (Gilchrist and Zakrajšek 2012). Finally, in order to ensure households find it profitable to directly hold capital in the normal times equilibrium, I target a steady state household capital of 0.33.

Table 1 shows the calibrated parameter values, and Table 2 shows the model moments compared to the target moments. The results yield parameter values consistent those used in existing literature. Given the relatively simple environment of the model, it hits the calibration targets accurately. All model moments are at or very close to their target range, with the exception of the spread increase during a run which falls within 10% of its target.

Table 1: Parameter Values

Parameter	Value	Description
β	0.99	Discount rate
σ	0.93	Bankers survival probability
θ	0.21	Seizure rate
α^h	0.0111	Household managerial cost
W^h	0.0305	Household endowment
w^b	0.0013	Bankers start-up funds
A	0.0262	Steady state productivity
α	0.35	Output Elasticity of Capital
η	1.02	Investor CES Elasticity

[17] Before the crisis, the Fed relied a very different operating framework in which reserves were not remunerated but reserve balances were actively managed. Consequently, there is no direct empirical counterpart for this period in the form an EFFR-IOR spread

Table 2: Model vs. Target Moments

Variable	Model	Target	Source
Bank Leverage, ϕ	8.42	8-10	Philippon (2015)
Bank Spread, $R^{k,b}$ (bp)	149	150-200	Philippon (2015)
Run Probability, p (pp)	4.13	4	Reinhart & Rogoff (2009)
Convenience yield on reserves (bp)	25	25	Average EFFR - T-bill spread
HH Capital Holdings, K^h	0.33	0.33	GK15
Price of Capital, Q	1.00	1	
Spread Increase in a run (bp)	254	275	Gilchrist & Zakrajšek (2012)

5.2 Optimal Reserve Supply

Having parametrised and solved the model, I carry out comparative statics with the respect to the size of the central bank balance sheet. In steady state, expected utility, and the values of all endogenous state variables depend on the entire saddle path for all endogenous variables in the event of a run. Consequently, statics must be computed numerically. Figure 2 shows ex-ante expected utility, $\mathcal{L}$, evaluated at different levels of reserve supply, and normalised to consumption equivalent terms (pp) relative to the optimal level in the baseline model.

In addition to the results from the baseline model (shown in the solid black line), I show results from two additional cases stripping out key frictions, in order to illustrate the roles of the novel elements of my theory, namely the risk of bank runs as motivation for reserve supply, and the explicit treatment of the cost of the central bank.

No runs. In the case with 'naive households' (shown as the red hollow dotted in line in Figure 2), depositors are assumed to always perceive the ex-ante probability of bank runs to be zero, meaning runs have no effect on agents' expected utility. This effectively switches off the liquidity component of this model, and the model becomes isomorphic to a standard model of financial frictions with an endogenous leverage constraint, and a central bank which creates capital misallocation. In this case, the optimal size of the central bank balance sheet is zero – the size which minimises the

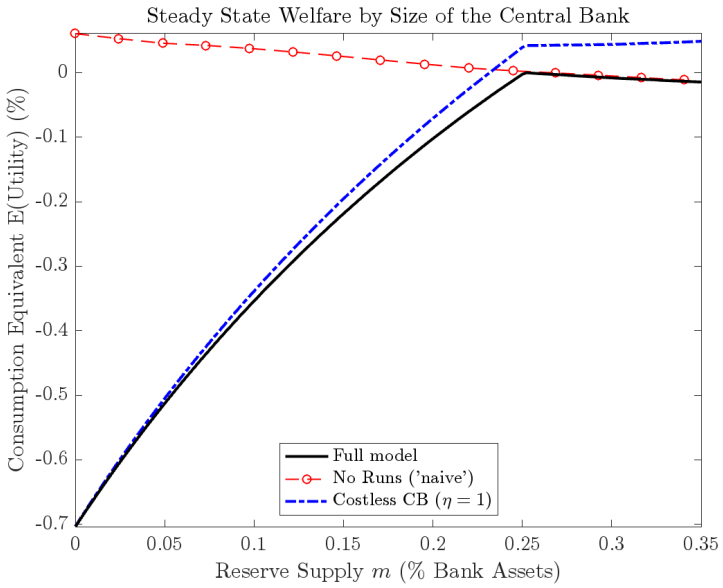


Figure 2: *Comparative Statics: Ex-Ante Expected Utility.* Notes: The figure plots the expected utility in 3 cases: the solid black line shows the full model. The hollow-dotted line shows results for the case where households assign zero probability to the bank run state, regardless of the banks' financial conditions (are 'naive'). The grey-dashed line shows the case where $\eta = 1$.

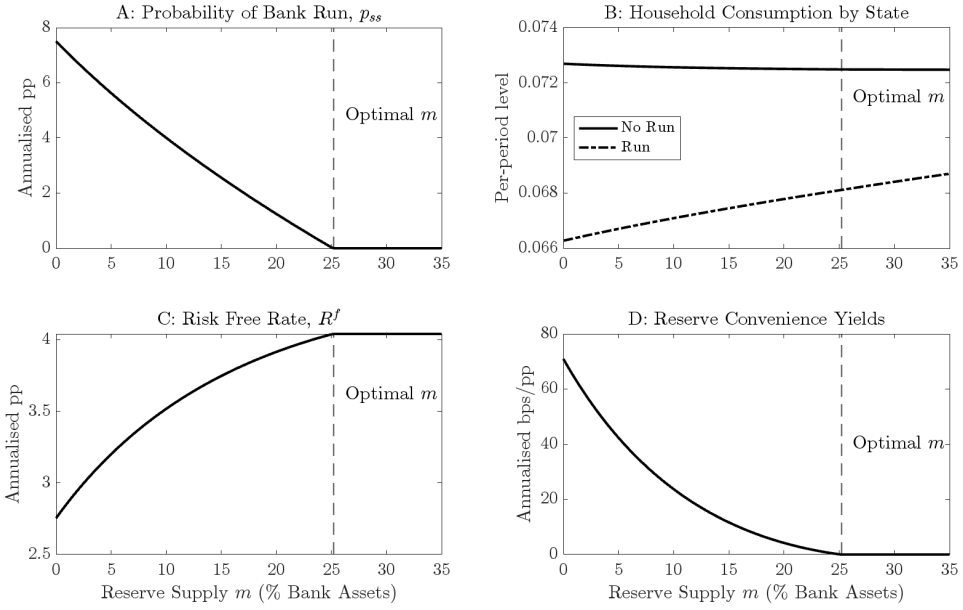


Figure 3: *Comparative Statics: Endogenous Variables.* Notes: The optimal size of the central bank balance sheet is marked by a dotted line for reference.

capital misallocation as a result of the central bank. Intuitively, households do not recognise the liquidity mismatch on banks balance sheets or the fragility this entails, so the additional liquidity created by the provision of central bank reserves has no welfare benefit in expectation.[18] But, when the central bank supplies reserves by purchasing capital, it increases the total dead-weight loss by driving the aggregate capital allocation to inefficiently over-supply sector one asset class at the expense of another. At this parametrisation, welfare is therefore monotonically decreasing in the size of the central bank.

Full model. In the full model, households account for the probability of a bank run when forming expectations. The welfare results for this case are shown by the solid black line in Figure 2. Figure 3 shows comparative statics for key endogenous variables. Increasing the provision of reserves reduces liquidity mismatch on bank balance sheets, raising the recovery rate and lowering the probability of a bank run (Fig. 3, Panel A). This reduces the risk premium banks must pay their depositors to compensate them for the risk of runs, pushing down the risk premium on deposit rates, relaxing the banks' leverage constraint and lowering lending spreads. Financial intermediation by the banking sector increases, reducing the dead-weight losses incurred by household management costs. Consumption in the bank run state, C_τ, also increases with reserve supply as banks are better prepared for runs if they do occur (Fig.3, Panel B). Consequently, the initial expansions in reserve supply generate substantial improvements in expected utility for the household, as a result of the reduction in likelihood and severity of bank runs.

The reduction in aggregate uncertainty also pushes down demand for precautionary savings and the unconditional risk-free rate in (10) rises as reserve supply expands (Fig. 3, Panel C). This matches findings in Arce et al. (2020) that equilibrium short-term rates are increasing in the size of the central bank, albeit via a different mechanism, and echoes suggestions by policy-makers that expansions in the central bank balance sheet may drive up r* (Schnabel 2025). However, as reserve supply increases and the

[18]Note that if a run were to occur as an exogenous shock (assuming the conditions for a run were satisfied ex-post such that $x_t < 1$), then a larger central bank balance sheet would reduce the severity of the fall in consumption, because banks would have a greater stock of liquid assets with which to pay households. However, because households assign a probability of zero to this event, it has no effect on expected utility.

likelihood of runs falls, the marginal benefit of additional reserves is decreasing. This is, in effect, because they provide insurance against a form of liquidity risk which is becoming increasingly remote. As the marginal impact of reserves on aggregate risk falls, so too does their convenience yield (Fig. 3, Panel D): banks demand higher compensation to continue to increase their reserve holdings as the marginal benefits of doing so dwindle. In Section 5.3, I show that this is consistent with the empirical evidence on convenience yields.

When reserve supply is large enough to rule out the bank run sunspot equilibrium (i.e. $x_{ss} = 1$ and $p_{ss} = 0$), the marginal liquidity benefit of reserves is zero: there is no risk of bank runs in steady state and as a result, increases in reserve supply have no further liquidity benefits. The convenience yield falls to zero (Fig. 3, Panel D) and reserve demand is satiated. Beyond this point, banks' have a sufficient volume of liquid assets to remain solvent in a run, and will only hold additional reserves at marginal cost. At the point where reserve demand is satiated, the full model becomes isomorphic to the 'naive household' case - the black and red lines in Figure 2 coincide.

The optimal-sized balance sheet falls at the point where reserve demand is satiated: the lowest level of reserve supply needed to set the probability of a run equal to zero, and marginal liquidity benefit of reserves jumps to zero. When the risk of bank runs becomes sufficiently low, the marginal cost of increasing the size of the central bank outweighs the expected utility benefits of further reducing the likelihood of runs. Beyond this point, further increases in the size of the central bank reduce welfare. In the baseline calibration, consumers prefer to eliminate run risk entirely. Comparing the full model to the naive household case illustrates a key contribution of this paper: explicitly modelling liquidity panics provides a simple micro-foundation for the welfare derived from very reserves, in a representative agent framework where agents have no preference for liquidity in their utility function.

In consumption terms, the effect of reserve supply on the level of welfare is small, but not negligible. The representative household would sacrifice approximately 0.4pp percentage points of their per-period consumption to move from the pre-crisis steady state, targeted in calibration, to the optimal balance sheet. In the US in 2025Q2, this corresponds to roughly $20 per month on average, comparable to the cost of a mobile

phone contract.[19]

Costless central bank. Finally, to illustrate the importance of explicitly modelling the negative externalities implied by a central bank footprint in private markets, the grey-dashed line in Figure 2 considers the edge case where the central bank creates no capital misallocation, either because asset markets are perfectly elastic ($\eta = 1$) or because the central bank can allocate its assets to any sector.[20] In this case, larger central bank balance sheets monotonically increase welfare. More reserves initially increase welfare through the mechanisms described above. But now, when the probability of bank runs falls to zero, additional reserves do not come with any cost in the form of capital misallocation. Instead, they continue to increase welfare by relaxing the financial frictions in the model, swapping risky bank assets which tighten the leverage constraint with reserves which do not (as studied by Eren, Jackson, and Lombardo (2024)). While this mechanism is present in the full model and the naive household case, its effect on welfare is dominated by the negative effect on consumption from the greater capital misallocation.

A figure showing statics for other endogenous variables is included in Appendix D.

5.3 Empirical validity

To evaluate the quantitative performance of the model, I compare the properties of the model implied reserve demand curve, which drives the key results of the welfare analysis above, to empirical estimates of the reserve demand curve.

I replicate the estimation exercise in Lopez-Salido and Vissing-Jørgensen (2025), to construct a reserve demand curve for the US, based on data from January 2009 to 2025. The authors estimate a relation linking the convenience yield on reserves (the spread between the effective federal funds rate (EFFR) and the interest rate paid on reserves (IOR)), to the aggregate supply of reserves and the stock of liquid deposits in the banking sector.[21] When this spread is positive (net of borrowers'

[19] Personal consumption expenditures per capita (A794RC0Q052SBEA) in 2025. Source: U.S. Bureau of Economic Analysis (2024)

[20] For values of $K^{cb} > k^{\star}$, perfectly elastic markets would fail to fully undo the distortion created by the central bank because they do not have sufficient capital holdings to do so.

[21] The authors argue the supply of reserves plus ONRRP is plausibly exogenous in the post-2008

marginal balance sheet costs), banks are willing to pay a premium in order to obtain liquidity in the form of reserves. As such, it measures the extent to which reserves are scarce. In practice, the EFFR-IOR spread is bounded from below by rate paid by the Federal Reserve's ONRRP facility. When reserve supply is sufficiently high relative to demand, the ONRRP allows the Fed to reduce reserve supply, whilst retaining control of short-term interest rates by allowing a broad range of counter-parties who do not have reserve accounts of their own to lend to the Fed overnight using repo agreements.[22] To account for this lower bound, they propose estimating the following equation by non-linear least squares:

$$EFFR - IOR = \max \begin{bmatrix} \mathbf{B} + \mathbf{C}\log(\text{Reserves}) + \mathbf{D}\log(\text{Deposits}), \\ (r^{ONRRP} - IOR) + \mathbf{A} \end{bmatrix} + U \qquad (48)$$

where r^{ONRRP} is the rate paid on the ONRRP facility, and "Reserves" denotes total reserve supply plus ONRRP balances. $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ estimate the parameters of the reserve demand curve, in the absence of the ONRRP-induced lower bound. $\mathbf{A}$ estimates the marginal cost of posting collateral in repo contracts. I estimate this equation using monthly data and obtain estimates very close to those in Lopez-Salido and Vissing-Jørgensen (2025). The estimates are presented in Appendix E, and the sources of all data used are given in Appendix F. The model fits the data well with a pseudo-R^2 of almost 90%.

This empirical strategy has three key features which make it suitable for testing the quantitative performance of my model. First, it accurately captures the quantitative relationship between the convenience of reserves and the level of supply. This convenience yield provides the natural empirical counterpart to the convenience yield on reserves implied in the model, given by the negative of equation (23), $\bar{R} - R^m$. Second, by including the ONRRP rate as a lower bound, it allows for the possibility that the Fed can supply liquidity in excess of demand, leading to a satiation point where the marginal convenience of reserves is zero. Third, it models the convenience

period when it was driven by quantitative easing and tightening, rather than shocks to reserve demand.

[22] When these counter-parties use the ONRRP facility, they liquidate bank deposits and lend directly to the Fed, mechanically reducing the reserves balances of their bank.

of reserves using only contemporaneous variables. As such, the estimates capture relations which should be present in steady state, rather than dynamic responses to shocks. This facilitates comparison with the static results in this paper.

In order to generate comparable estimates, I calculate the reserve demand elasticity (RDE), measured as the basis point change in the convenience yield on reserves for a given shock to reserve supply. This captures the *slope* of the demand curve for reserves. Following Afonso et al. (2022), I consider a reserve shock equal to 1pp of bank assets. This measure captures both the sensitivity of the convenience of reserves to changes in supply, and the level of reserves which satiates demand (or equivalently sets the RDE to 0), offering two dimensions against which to test the predictions of my model.

In the model, the RDE can be calculated by dividing the difference in convenience yield on reserves, associated with two levels of reserve supply, by the change in reserve supply, and multiplying the result by 1% of total bank assets:

$$\hat{RDE}^{\text{model}} = \frac{\Delta CY}{\Delta m} \times 0.01 \times (QK^b + m)$$

where CY is the annualised spread between bank's marginal cost – the deposit rate, $\bar{R}$ – and the rate on reserves, R^m.

To obtain the empirical counterpart, take the derivative of equation (48) with respect to reserves, noting the non-linearity when the lower bound bites. Evaluate this derivative at every point in the sample, and multiply the estimates by 1% of bank assets.[23]:

$$\hat{RDE}^{\text{data}} = \begin{cases} \dfrac{\hat{C}}{\text{Reserves}} \times 0.01 \times (\text{Bank Assets}) & \text{if } EFFR - IOR > r^{ONRRP} - IOR + \hat{A} \\ 0 & \text{otherwise} \end{cases}$$

Figure 4 plots both the model implied and empirical estimates of the reserve demand elasticity against the level of reserve supply (normalised by total bank assets). The model generates quantitatively realistic estimates of the elasticity of reserve demand, and reproduces the key non-linearity in the reserve demand curve: the

[23]Retrieved from Board of Governors of the Federal Reserve System (US) (2024) in July 2025

(absolute) slope of the demand curve decreases as reserve supply increases, and at the point of satiation, jumps to zero. The model also estimates the level of reserve supply which satiates demand fairly accurately, albeit the model implied RDE tends to zero slightly slower than in the data. Nonetheless, in the regions of the data where reserve demand is abundant, the model implies an RDE of between 0 and -1 bp/pp, implying an extremely shallow demand curve for reserves in this region.

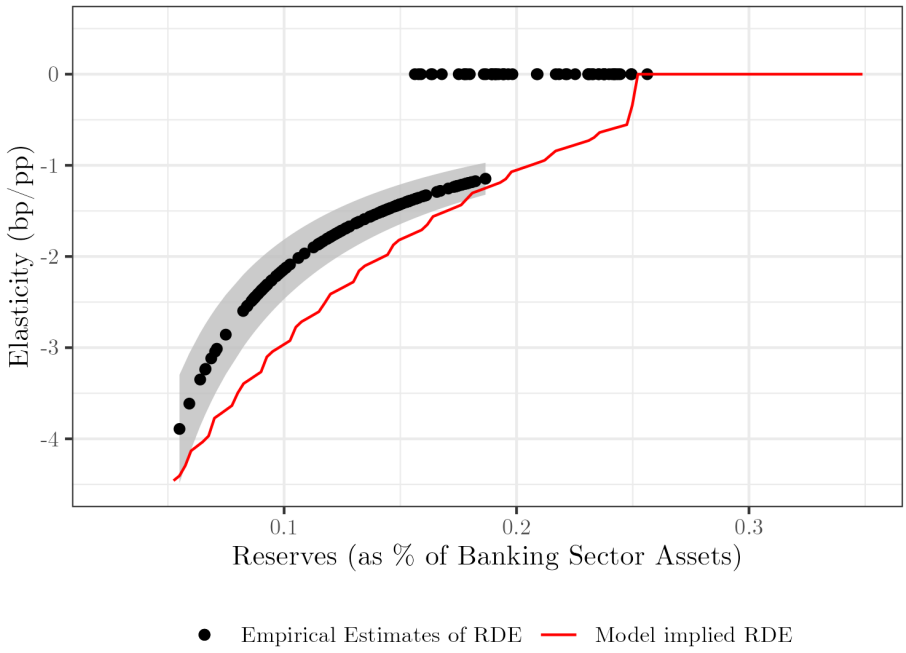


Figure 4: *Reserve Demand Elasticity: Model vs Data.* Notes: Empirical estimates are taken from the results reported in Appendix E. The swathe shows 95% confidence intervals from the estimation procedure, using robust standard errors.

5.4 Robustness to the strength of portfolio frictions

The strength of portfolio frictions governs the extent to which the central bank creates capital misallocation in the aggregate. As a result, the marginal social cost of increasing the size of the central bank depends directly on the calibration of η.

The baseline calibration of η was chosen to match a plausible upper bound on the persistent impact of central bank purchases (worth 8% of the US treasury market) on

long-term yields, as estimated by D'Amico and King (2013).[24] Below I demonstrate that the conclusions above are robust to a wide range of η values.

Figure 5 (Panel A) shows the optimal level of reserve supply as a function of η, holding all other parameters constant at the baseline calibration of the model. For context, the dotted blue line plots the minimum level needed to fully satiate reserve demand and set $p = 0$ in steady state. The optimal long-run level of reserve is not sensitive to this parameter. For very small values, it remains optimal for the central bank to operate an ample, but not abundant, reserve regime. Once reserve demand is satiated, additional reserve supply has no further liquidity benefits. Consequently, even a small social cost of further expansion is enough to make additional reserve supply sub-optimal.

When η is extremely close to 1 ($\eta < 1.0005$) the former channel dominates, and we return to the corner solution already discussed above in the "costless" case in 2. However, under this extreme assumption, even very large increases in central bank asset portfolios would have no effect on cross-sector yields, running counter to lengthy empirical evidence.[25]

When η is large ($\eta > 1.2$) and the cost of capital misallocation is higher, and the optimal size of the central bank falls slightly below the satiation point: the red line falls below the dotted blue line in panel A. In this case, the marginal cost of capital misallocation outweighs the liquidity benefits of additional reserves at slightly smaller levels of reserve supply, and the optimal size of the central bank is smaller than the satiation point. This result echoes the findings of Vissing-Jørgensen (2023): if the central bank purchases assets which themselves have a high marginal value to the private sector, then a smaller central bank balance sheet is optimal. However, this effect is not strong enough to warrant a return to scarce reserves: when $\eta = 1.3$ the convenience yield on reserves at the optimal size of central bank is still below 5bps (Fig.5, Panel B), suggesting reserves remain in ample supply relative to demand. This suggests that even if the costs of the central bank are significantly greater than those

[24] In the context of the model, I set η such that central bank holds 8% of total sector 1 holdings implies an annualised yield differential of 50bps.

[25] See Gulati and Smith (2022) for a comprehensive review of the evidence on the impact of central bank balance sheet expansions on yields.

estimated in the baseline calibration, it remains optimal to operate an ample, rather than scarce, reserve regime.

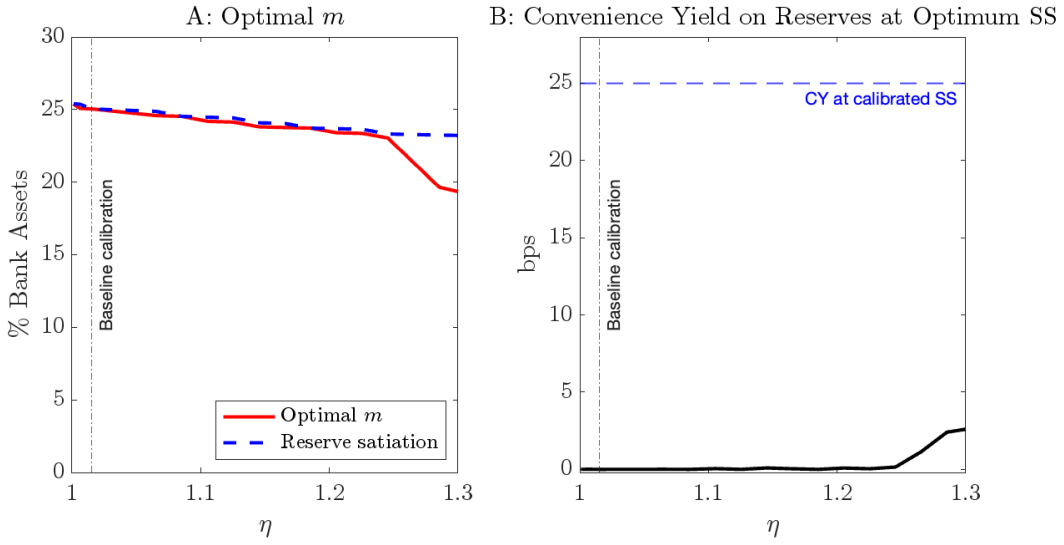


Figure 5: *Robustness of Optimal Central Bank Size to η.* Notes: The figure shows the optimal steady state level of m (Panel A), and the convenience yield on reserves at the optimal steady state (Panel B) for $\eta \in [1.0005, 1.3]$

6 Implications for Central Bank Balance Sheet Policy

The results imply a number of clear implications for policy-makers.

1. *Ample reserve regimes are optimal.*

The welfare-maximising size of the central bank balance sheet lies at the minimum level of reserve supply where reserve demand is satiated. This model therefore provides a parsimonious, welfare-based rationalisation for 'ample' reserve regimes. Such regimes supply enough reserves to satisfy aggregate demand for reserves, in recognition of the fact that prior to the crisis, the supply of reserves was inefficiently low. Yet, they do not surpass this point. For example, the Bank of England has stated in 2024 that it intends to continue to supply the "*minimum* level of reserves that satisfies commercial banks' aggregate demand, both to settle their everyday transactions and to hold cash

as a precaution against potential outflows in times of stress." (Bailey 2024, emphasis added), consistent with the optimal policy suggested by this model. Multiple methods exist to assess the level of reserve supply relative to demand in real-time (Afonso et al. 2022; Lopez-Salido and Vissing-Jørgensen 2025), making this policy relatively simple to implement.

2. *Minimising the distortions caused by the central bank balance sheet may not change optimal policy, but can affect welfare.*

Section 5.4 shows that even if asset demand is highly elastic and the extent capital misallocation caused by central bank is very small, households still do not prefer to expand the long-run size of the central bank to an abundant reserve regime. This implies that efforts to minimise the distortions caused by the central bank balance sheet (for example by supplying reserves against a broad range of collateral, or purchasing assets with highly elastic demand) are unlikely to change the conclusion that ample, not abundant, reserve regimes are optimal.

However, comparing the full model to the costless case does suggest that such efforts could make a non-negligible contribution to welfare and therefore should be a priority for policy-makers.

3. *Bigger central bank balance sheet can increase risk-free rates.*

The model suggests a positive relationship between long-run risk-free rates and the size of the central bank balance sheet, up to the point where reserve demand is satiated. This is a result of the fact that increasing steady state liquidity reduces aggregate liquidity risk, and with it, reduces the demand for precautionary savings. This echoes the findings in Aiyagari and McGrattan (1998), and more recently, Nuño (2025). The implications of this finding for the efficacy of central bank asset purchases, the optimal conduct of monetary policy, and the likelihood of hitting the effective lower bound are beyond the scope of my framework, but remain important topics for policy-makers, and further research.

Finally, policy-makers may ask to what extent standing 'lender of last resort' facilities may affect the optimal size of the central bank balance sheet. Clearly, a credible commitment to intervene during financial turmoil could reduce incentives for

depositors to run, and reduce the severity of runs when they do happen. As a result these policies may improve welfare, provided the costs of such interventions (such as fiscal costs of purchasing distressed assets, and moral hazard concerns of committing to intervene in stress) to do not outweigh the benefits. However, to the extent that runs remain possible in equilibrium, the results from this framework would suggest that expanding the central bank balance sheet up to an ample reserve supply would remain optimal, as agents would still prefer to insure themselves against the remaining run risk. As a result, lender of last resort policies would be unlikely alter the key policy prescription of this analysis.

7 Conclusion

This paper presents investigates the socially optimal long-run size of the central bank balance sheet. Welfare benefits arise in this model as a result of the liquidity of central bank reserves, by reducing the probability and severity of financial panics. The numerical results suggest that these benefits come cheaply, relative to the marginal cost of increasing the size of the central bank, making it optimal for central banks to provide reserves up to minimum point where their marginal benefit to liquidity conditions is zero, but not further. This is consistent with an ample supply of reserves in steady state. Beyond this point, additional reserves have no effect on liquidity conditions and further expansion of the central bank crowds out the private financial sector and reduces aggregate welfare. The conclusion that an ample reserve regime is optimal is robust to a wide range of assumptions about the strength of portfolio frictions, which govern the social cost of expanding the central bank.

For the purpose of tractability and exposition, I have made a number of simplifying assumptions. For instance, in studying a steady state where exogenous shocks are held constant, my framework is silent on whether other risks which drive demand for precautionary savings may create additional benefits of supplying reserves. Second, in my treatment of capital misallocation, I abstract from any dynamic effects of capital misallocation. Allowing for endogenous capital accumulation may help to account more fully for the long-term risks of over-supplying reserves. I leave these topics for further research.

8 References

Acharya, Viral V., Ryan Banerjee, Matteo Crosignani, Tim Eisert, and Renée Spigt. 2025. "Exorbitant Privilege? Quantitative Easing and the Bond Market Subsidy of Prospective Fallen Angels." *Journal of Financial Economics* 170 (August): 104084. https://doi.org/10.1016/j.jfineco.2025.104084.

Afonso, Gara, Domenico Giannone, Gabriele La Spada, and John C. Williams. 2022. "Scarce, Abundant, or Ample? A Time-Varying Model of the Reserve Demand Curve." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4121823.

Aiyagari, S. Rao, and Ellen R. McGrattan. 1998. "The Optimum Quantity of Debt." *Journal of Monetary Economics* 42 (3): 447–69. https://doi.org/10.1016/S0304-3932(98)00031-2.

Arce, Óscar, Galo Nuño, Dominik Thaler, and Carlos Thomas. 2020. "A Large Central Bank Balance Sheet? Floor Vs Corridor Systems in a New Keynesian Environment." *Journal of Monetary Economics* 114 (October): 350–67. https://doi.org/10.1016/j.jmoneco.2019.05.001.

Bailey, Andrew. 2024. "The Importance of Central Bank Reserves." https://www.bankofengland.co.uk/speech/2024/may/andrew-bailey-lecture-london-school-of-economics-charles-goodhart.

Bindseil, Ulrich. 2016. "Evaluating Monetary Policy Operational Frameworks." *Proceedings of the Jackson Hole Economic Symposium of the Federal Reserve Bank of Kansas City on "Designing Resilient Monetary Policy Frameworks for the Future"*, August.

Board of Governors of the Federal Reserve System (US). 2024. "Total Assets, All Commercial Banks." *FRED, Federal Reserve Bank of St. Louis*. https://fred.stlouisfed.org/series/TLAACBW027SBOG.

Boissay, Frederic, Fabrice Collard, Jordi Galí, and Cristina Manea. 2021. "Monetary Policy and Endogenous Financial Crises." *NBER Working Papers*, December. https://ideas.repec.org//p/nbr/nberwo/29602.html.

Borio, Claudio. 2023. "Getting up from the Floor." *BIS Working Paper*, no. 1100 (May).

Chen, Han, Vasco Cúrdia, and Andrea Ferrero. 2012. "The Macroeconomic Effects of Large-scale Asset Purchase Programmes." *The Economic Journal* 122 (564):

F289–315. https://doi.org/10.1111/j.1468-0297.2012.02549.x.

D'Amico, Stefania, and Thomas B. King. 2013. "Flow and Stock Effects of Large-Scale Treasury Purchases: Evidence on the Importance of Local Supply." *Journal of Financial Economics* 108 (2): 425–48. https://ideas.repec.org//a/eee/jfinec/v108y2013i2p425-448.html.

Del Negro, Marco, Gauti Eggertsson, Andrea Ferrero, and Nobuhiro Kiyotaki. 2017. "The Great Escape? A Quantitative Evaluation of the Fed's Liquidity Facilities." *American Economic Review* 107 (3): 824–57. https://doi.org/10.1257/aer.20121660.

Del Negro, Marco, and Christopher A. Sims. 2015. "When Does a Central Bank's Balance Sheet Require Fiscal Support?" *Journal of Monetary Economics*, Carnegie-Rochester-NYU Conference Series on Public Policy "Monetary Policy: An Unprecedented Predicament" held at the Tepper School of Business, Carnegie Mellon University, November 14-15, 2014, 73 (July): 1–19. https://doi.org/10.1016/j.jmoneco.2015.05.001.

Eren, Egemen, Timothy Jackson, and Giovanni Lombardo. 2024. "The Macroprudential Role of Central Bank Balance Sheets." *BIS Working Paper*, March.

Gabaix, Xavier, and Ralph S. J. Koijen. 2023. "In Search of the Origins of Financial Fluctuations: The Inelastic Markets Hypothesis." {SSRN} {Scholarly} {Paper}. Rochester, NY. https://doi.org/10.2139/ssrn.3686935.

Gertler, Mark, and Peter Karadi. 2011. "A Model of Unconventional Monetary Policy." *Journal of Monetary Economics*, Carnegie-Rochester Conference Series on Public Policy: The Future of Central Banking April 16-17, 2010, 58 (1): 17–34. https://doi.org/10.1016/j.jmoneco.2010.10.004.

———. 2013. "QE 1 Vs. 2 Vs. 3. . . : A Framework for Analyzing Large-Scale Asset Purchases as a Monetary Policy Tool." *International Journal of Central Banking* 9 (1): 5–53. https://ideas.repec.org//a/ijc/ijcjou/y2013q0a1.html.

Gertler, Mark, and Nobuhiro Kiyotaki. 2010. "Financial Intermediation and Credit Policy in Business Cycle Analysis." In *Handbook of Monetary Economics*, edited by Benjamin M. Friedman and Michael Woodford, 3:547–99. Elsevier. https://doi.org/10.1016/B978-0-444-53238-1.00011-9.

———. 2015. "Banking, Liquidity, and Bank Runs in an Infinite Horizon Economy."

The American Economic Review 105 (7): 2011–43. https://www.jstor.org/stable/43495449.

Gertler, Mark, Nobuhiro Kiyotaki, and Andrea Prestipino. 2016. "Wholesale Banking and Bank Runs in Macroeconomic Modeling of Financial Crises." In *Handbook of Macroeconomics*, edited by John B. Taylor and Harald Uhlig, 2:1345–425. Elsevier. https://www.sciencedirect.com/science/article/pii/S1574004816000100.

———. 2020. "A Macroeconomic Model with Financial Panics." *The Review of Economic Studies* 87 (1): 240–88. https://doi.org/10.1093/restud/rdz032.

Gilchrist, Simon, and Egon Zakrajšek. 2012. "Credit Spreads and Business Cycle Fluctuations." *American Economic Review* 102 (4): 1692–1720. https://doi.org/10.1257/aer.102.4.1692.

Greenwood, Robin, Samuel G. Hanson, and Jeremy C. Stein. 2018. "The Federal Reserve's Balance Sheet as a Financial-Stability Tool." In *World Scientific-Now Publishers Series in Business*, 15:63–124. World Scientific/NOW Publishers. https://www.worldscientific.com/doi/abs/10.1142/9789813236592_0003.

Gulati, Chaitri, and A. Lee Smith. 2022. "The Evolving Role of the Fed's Balance Sheet: Effects and Challenges." *The Federal Reserve Bank of Kansas City Economic Review*, November. https://doi.org/10.18651/ER/v107n4GulatiSmith.

Hall, Robert E., and Ricardo Reis. 2015. "Maintaining Central-Bank Financial Stability Under New-Style Central Banking." Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w21173.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India*." *The Quarterly Journal of Economics* 124 (4): 1403–48. https://doi.org/10.1162/qjec.2009.124.4.1403.

Karadi, Peter, Dominik Thaler, and Oreste Tristani. 2024. "The Central Bank's Balance Sheet in the Long Run: A Macro Perspective." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4975240.

Kiyotaki, Nobuhiro, and John Moore. 1997. "Credit Cycles." *Journal of Political Economy* 105 (2): 211–48. https://doi.org/10.1086/262072.

Koijen, Ralph S. J., and Motohiro Yogo. 2019. "A Demand System Approach to Asset Pricing." *Journal of Political Economy* 127 (4): 1475–515. https://doi.org/10.1086/701683.

Krishnamurthy, Arvind, and Annette Vissing-Jørgensen. 2012. "The Aggregate Demand for Treasury Debt." *Journal of Political Economy* 120 (2): 233–67. https://doi.org/10.1086/666526.

Lopez-Salido, David, and Annette Vissing-Jørgensen. 2025. "Reserve Demand, Interest Rate Control, and Quantitative Tightening." Rochester, NY. https://doi.org/10.2139/ssrn.4371999.

Morris, Stephen, and Hyun Song Shin. 1998. "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks." *The American Economic Review* 88 (3): 587–97. https://www.jstor.org/stable/116850.

Nuño, Galo. 2025. "Three Theories of Natural Rate Dynamics." *Working Paper*, May.

Philippon, Thomas. 2015. "Has the US Finance Industry Become Less Efficient? On the Theory and Measurement of Financial Intermediation." *American Economic Review* 105 (4): 1408–38. https://doi.org/10.1257/aer.20120578.

Reinhart, Carmen M., and Kenneth S. Rogoff. 2009. "This Time Is Different: Eight Centuries of Financial Folly." In *This Time Is Different*. Princeton University Press. https://doi.org/10.1515/9781400831722.

Schnabel, Isabel. 2023. "Back to Normal? Balance Sheet Size and Interest Rate Control." *ECB Speech*, March. https://www.ecb.europa.eu/press/key/date/2023/html/ecb.sp230327_1~fe4adb3e9b.en.html.

———. 2025. "No Longer Convenient? Safe Asset Abundance and r*." *Speech at Bank of England Agenda for Research Conference*, February. https://www.ecb.europa.eu/press/key/date/2025/html/ecb.sp250225~49a8dd911a.en.html.

U.S. Bureau of Economic Analysis. 2024. "Personal Consumption Expenditures Per Capita." *FRED, Federal Reserve Bank of St. Louis*. https://fred.stlouisfed.org/series/A794RC0Q052SBEA.

Vayanos, Dimitri, and Jean-Luc Vila. 2021. "A Preferred-Habitat Model of the Term Structure of Interest Rates." *Econometrica* 89 (1): 77–112. https://doi.org/10.3982/ECTA17440.

Vissing-Jørgensen, Annette. 2023. "Balance Sheet Policy Above the ELB." *Presented at ECB Forum on Central Banking*.

Appendix

A Equilibrium Conditions

Here I define the recursive competitive equilibrium with no bank run. It consists of aggregate quantities for:

$$(K_t^b, K_t^h, K_t^{cb}, C_t, N_t, Y_t, \tau_t, \Pi_t^b, \Pi_t^f)$$

prices:

$$\left(Q_t, \bar{R}_{t+1}, R_{t+1}^m, f_h(K_t^h), Z_t\right)$$

bank-related variables:

$$\left(\Omega_{t,t+1}, R_t^{k,b}, \psi_t, \phi_t, \gamma_t, p_t\right)$$

as a function of state variables:

$$\left(K_{t-1}^b, R_t D_{t-1}, R_t^m m_{t-1}, A_t, m_t\right)$$

which satisfy eqs. (1), (4) to (6), (9), (11), (16), (19), (20), (22) to (27), (33) and (37) to (40)

To evaluate household expectations, we also need counter-factual bank run values for consumption and prices which are used in households' expectations of the bank run equilibrium:

$$(C_{t+1}^*, Q_{t+1}^*)$$

These values must satisfy (39) evaluated in the bank run state, and (31). This yields a system of 22 equations and 22 unknowns. The household budget constraint is satisfied by Walras' Law.

B Proof of Proposition 3.1

1. If $K^{cb} = 0$, the decentralised economy achieves the first-best allocation:

Let $K^{cb} = 0$, such that all capital is allocated via the representative investor. The CES demand implies that if returns are equal across sectors, capital will be equally distributed. By symmetry, this allocation satisfies the firm FOCs in (34) and the investor's optimality conditions. This allocation equalises marginal products of capital and therefore implements the first-best.

2. If the decentralised economy achieves the first-best allocation, then $K^{cb} = 0$:

Suppose the decentralised allocation is efficient. Then, it equalises MPKs across sectors. Assume for contradiction that $K^{cb} > 0$.

The firm FOC implies:

$$A\alpha(K^{cb} + k_1^p)^{\alpha-1} = A\alpha(k_j^p)^{\alpha-1} \quad \forall j \neq 1 \quad \Rightarrow \quad K^{cb} + k_1^p = k_j^p. \tag{49}$$

The investor FOC (35) implies that when returns are equal and $\eta > 1$, capital allocations must also be equal:

$$k_1^p = k_j^p = k^\star \quad \forall j.$$

Substituting this result into (49), we get:

$$K^{cb} + k^\star = k^\star \Rightarrow K^{cb} = 0,$$

a contradiction. □

C Computation

The computational strategy relies on a time iteration algorithm based closely on that described in Gertler, Kiyotaki, and Prestipino (2016), Appendix E.

For computation, introduce a new variable denoting the ex-ante optimal value function for surviving bankers at time t:

$$\bar{\psi}_t = \Omega_{t,t+1} \frac{N_{t+1} - w^b}{\sigma N_t}. \tag{50}$$

Let the state of the economy if a run has **not** happened be denoted by

$$x = (N, A, m),$$

and the state in case a run **has** happened be

$$x^* = (0, A, m).$$

Policy functions are approximated as piecewise linear functions over the state space:

$$\{\, Q(x),\ C^h(x),\ \bar{\psi}(x),\ \Gamma(x) \,\}, \quad x \in [w^b, \bar{N}] \times [0.95A,\ 1.05A] \times [0, \bar{m}],$$

and

$$\{\, Q^*(x^*),\ C^{h*}(x^*),\ \ \Gamma^*(x^*) \,\}, \quad x^* \in \{0\} \times [0.95A,\ 1.05A] \times [0, \bar{m}],$$

where $\Gamma(\cdot)$ and $\Gamma^*(\cdot)$ are the laws determining the evolution of the state variables, as a function of the equilibrium conditions.

Computational algorithm (time iteration)

1. **Approximation space.** Choose a functional space for the equilibrium objects, and define state grids:

   $$G \subset [w^b, \bar{N}] \times [0.95A,\ 1.05A] \times [0, \bar{m}], \qquad G^* \subset \{0\} \times [0.95A,\ 1.05A] \times [0, \bar{m}].$$

2. **Initial guesses.** Set $j = 0$ and guess

   $$\text{NRPol}_{t,j} = \{\, Q_{t,j}(x),\ C^h_{t,j}(x),\ \ \overline{\psi}_{t,j}(x),\ \Gamma_{t,j}(x) \,\}_{x \in G},$$

$$\text{RPol}_{t,j} = \{\, Q^*_{t,j}(x^*),\ C^{h*}_{t,j}(x^*),\ \Gamma^*_{t,j}(x^*)\,\}_{x^*\in G^*}.$$

For $\Gamma_{t,j}(x)$, guess $\{p_{t,j}(x),\ \ N'_{t,j}(x),\ A'(A)\}$ and set

$$\Gamma_{t,j}(x) = \begin{cases} \Big(N'_{t,j}(x),\ A'(A),\ m\Big) & \text{w.p. } 1-p_{t,j}(x), \\ \Big(0,\ A'(A),\ m\Big) & \text{w.p. } p_{t,j}(x). \end{cases}$$

Define next-period states $x'^{NR}_{t,j}(x) = (N'_{t,j}, A'(A), m)$ and $x'^{R}_{t,j}(x) = (0, A'(A), m)$.

For $\Gamma^*_{t,j}(x^*)$, guess $\{A'(A)\}$ and set

$$\Gamma^*_{t,j}(x^*) = \Big((1+\sigma)w^b,\ A'(A),\ m\Big).$$

3. **Interpolate.** Given $\text{NRPol}_{t,i}$ and $\text{RPol}_{t,i}$ (for $j \le i < M$ where M is very large (e.g. 10000)), approximate functions for each variable (e.g., Q_i) that matches the policy functions at each point on the grid: $Q_i(x) = Q_{t,i}(x)$ for all $x \in G$. Do the same for the run case. I do so with linear interpolation, using CompEcon tool kit for Matlab.

4. **Update policy functions.** Compute $\text{NRPol}_{t,i+1}$ and $\text{RPol}_{t,i+1}$ by assuming equilibrium outcomes are determined by the functions calculated in step 3. Then use these values to compute the endogenous variables of the model using the equilibrium conditions in Appendix A, by replacing "t+1" variables with those implied by $\text{NRPol}_{t,i+1}$ and $\text{RPol}_{t,i+1}$. For instance, replace $C_{i,t+1}$ with:

$$\mathbf{C}_i(\Gamma_i(x)) = \begin{cases} C_i(N'_i(x), A'(A), m) & \text{w.p. } 1-p_{t,i+1}(x), \\ C^*_i(0, A'(A), m) & \text{w.p. } p_{t,i+1}(x). \end{cases}$$

5. **Convergence and damping.** With the values implied for the endogenous variables,compute NRPol_t and RPol_t and the maximum distance between the new and old policy functions

$$d_{NR} = \max_{x\in G}\Big\|\text{NRPol}_t - \text{NRPol}_{t,i}\Big\|, \qquad d_R = \max_{x^*\in G^*}\Big\|\text{RPol}_t - \text{RPol}_{t,i}\Big\|.$$

If both are below tolerance (e.g., 10^{-6}), accept; otherwise apply damping

$$\mathrm{NRPol}_{t,i+1} = \alpha\,\mathrm{NRPol}_{t,i} + (1-\alpha)\,\mathrm{NRPol}_t, \qquad \mathrm{RPol}_{t,i+1} = \alpha\,\mathrm{RPol}_{t,i} + (1-\alpha)\,\mathrm{RPol}_t, \quad \alpha \in (0,1),$$

and iterate.

D Additional Comparative Statics

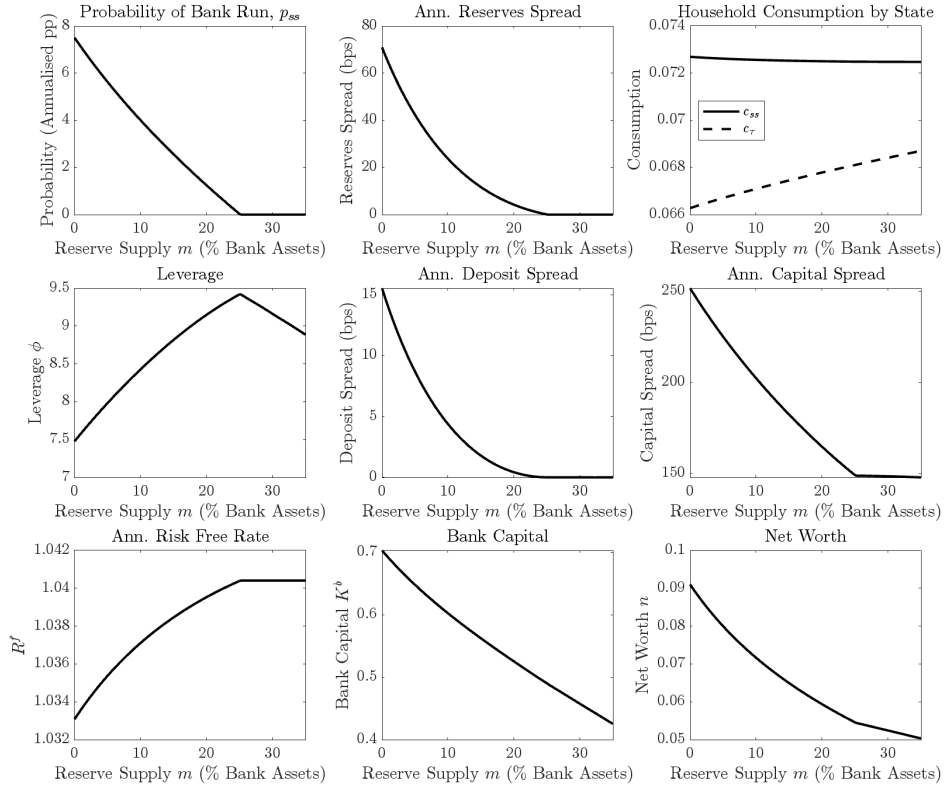


Figure 6: Additional Comparative Statics.

E Replication of Lopez-Salido and Vissing-Jørgensen (2025)

The variables are defined as in Lopez-Salido and Vissing-Jørgensen (2025). "Reserves" are the sum of total reserve supply and ONRRP balances. The ONRRP rate is the rate paid on the ONRRP facility. "Deposits" are sum of demand deposits and other liquid deposits (retrieved directly from FRED for data after 2020M5) and constructed as the sum of savings deposits and other checkable deposits prior to that period.

Table 3: Estimation of Equation (48) by non-linear least squares

	Dependent variable: EFFR - IOR
B	−123.942***
	(4.603)
C	−21.400***
	(0.733)
D	30.293***
	(0.902)
A	3.440***
	(0.255)
Num.Obs.	194
R2	0.877

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001
Robust standard errors are Newey-West HAC (lag = 12).

F Data

US

- Banking Sector Assets - Federal Reserve Economic Data, TLAACBW027SBOG.
- Demand Deposits - Federal Reserve Economic Data, WDDNS.
- Effective Federal Funds Rate - Federal Reserve Economic Data, EFFR.
- Federal Reserve Total Assets - Federal Reserve Economic Data, WALCL.
- GDP - Federal Reserve Economic Data, GDP.

- Interest on Reserves (2008 - 2021) - Federal Reserve Economic Data, IORR.
- Interest on Reserve Balances (2021 - 2025) - Federal Reserve Economic Data, IORB.
- ONRRP rate - Federal Reserve Economic Data, RRPONTSYAWARD.
- ONRRP - Federal Reserve Economic Data, RRPONTSYD.
- Other Checkable Deposits - Federal Reserve Economic Data, OCDNS.
- Other Liquid Deposits - Federal Reserve Economic Data, MDLM.
- Reserves - Federal Reserve Economic Data, WRESBAL.
- Savings Deposits - Federal Reserve Economic Data, SAVINGS.

UK

- Central Bank Assets - Sum over all line items listed under Assets in the Bank of England Weekly Report Table B1.1.3
- GDP - Federal Reserve Economic Data, UKNGDP.

Euro-system

- Central Bank Assets - Federal Reserve Economic Data, ECBASSETSW.
- GDP - Federal Reserve Economic Data, EUNNGDP.