Monetary Policy with Reserves and CBDC: Optimality, Equivalence, and Politics*

Dirk Niepelt†

August 5, 2021

Abstract

We analyze the introduction of retail central bank digital currency (CBDC) into a two-tiered monetary system. Deposits, reserves, and CBDC differ in terms of operating costs and liquidity. We identify the optimal monetary system and characterize first- and second-best policies that account for externalities and bank market power. Optimal spreads satisfy modified Friedman rules; deposits are taxed or subsidized; interest rates on reserves and CBDC differ; and the second-best policy targets the composition of real balances. With commensurate resource costs of single- and two-tiered payment systems the central bank can neutralize all macro effects of CBDC. But political constraints might prevent this and a two-tiered system requires higher taxes. The model implies implicit subsidies to U.S. banks of up to 1.5 percent of GDP.

JEL codes: E42, E43, E51, E52, G21

Keywords: Central bank digital currency, reserves, two-tiered system, monetary policy, Friedman rule, equivalence, money creation

*For valuable comments or discussions, we are grateful to David Andolfatto, Raphael Auer, Marco Bassetto, Javier Bianchi, Florian Böser, Antonio Fatás, Martín Gonzalez-Eiras, Timothy Jackson, Antoine Martin, Cyril Monnet, Thomas Moser, Juan Pablo Nicolini, Chris Phelan, Antti Ripatti, Jean-Charles Rochet, Remo Taudien, Harald Uhlig, Stephen Williamson, Shengxing Zhang, as well as conference or seminar participants at the Barcelona GSE Summer Forum, Bundesbank, CB&DC Virtual Seminar, European Central Bank, Federal Reserve Bank of Minneapolis, Joint Online Macro/Finance Seminar (Academia Sinica, Hong Kong Baptist University, Monash University, National Taiwan University, New York University-Shanghai, Yonsei University, Renmin University of China), National Bank of Ukraine, Swiss National Bank, SNB-CIF Conference, TU Berlin, Universities of Copenhagen, Frankfurt, Liverpool, and Zurich (joint with Federal Institute of Technology), Virtual Finance Theory Seminar, and the 6th International Macroeconomics and Finance Conference (Keio/CEPR). We thank Luca Benati, Han Gao, and Juan Pablo Nicolini for help with the data and Pablo Kurlat for sharing his data.

†Study Center Gerzensee; University of Bern; CEPR; CESifo. `dirk.niepelt@vwi.unibe.ch`, `www.niepelt.ch`.

1 Introduction

Fast technological change in the payment sector as well as the prospect of retail central bank digital currency (CBDC) or "Reserves for All" have rekindled interest in fundamental questions about the monetary system. In modern economies this system exhibits two tiers: Nonfinancial institutions in the private sector use deposit liabilities of banks or claims on such deposits as payment instruments, while banks themselves pay each other with reserves that are issued by the central bank. New financial service providers build on these payment rails and exploit synergies but they do not undermine the monetary architecture. The introduction of CBDC, in contrast, which grants direct access to digital central bank money also to the general public, would call the two-tiered system into question.[1]

Independently, change in the monetary system is already underway. Since the financial crisis erupted in 2007 balance sheets of central banks have lengthened and banks have responded to "quantitative easing" and tightened liquidity regulation by backing a larger share of their deposits with reserves held at the central bank. As a consequence money multipliers have collapsed. While proposals to outlaw fractional reserve banking, such as the "Chicago Plan" from the 1930s or the more recent Swiss "Vollgeld" initiative, have not been enacted the payment system has evolved into an architecture with a more limited role for private money creation.[2]

This paper analyzes the contemporary two-tiered monetary system as well as the likely future, integrated system with deposits, reserves, and CBDC. We address normative questions concerning the optimal monetary system and the characteristics of optimal policy within a given system. We explore under what conditions the introduction of CBDC does or does not alter equilibrium outcomes. And we examine the fiscal burden that different monetary systems impose on taxpayers and possible repercussions for political support. As a byproduct we quantify the funding cost reduction for banks in a two-tiered system due to their liquidity provision backed by lender-of-last-resort guarantees of the central bank. We find that for the U.S. this was on the order of 0.5 to 1.5 percent of GDP just before the financial crisis and similarly large more recently, but negligible or even negative in other periods.

Our framework is an extended Sidrauski (1967) model that embeds banks, deposits, reserves, and CBDC into the baseline business cycle model. Households value goods, leisure, and the liquidity services provided by deposits and CBDC. Neoclassical firms

[1] Unbacked "cryptocurrency" bubbles such as Bitcoin, which are not the focus of this paper, are other payment instruments. CBDC would re-introduce rather than newly introduce noncash government issued payment instruments for retail customers since central banks offered accounts to nonbanks already in the past (BIS, 2018). Some countries such as Uruguay have introduced retail CBDCs, others such as China are in the process, and the leadership of the European Central Bank sends strong signals in favor, while Federal Reserve officials continue to voice skepticism via-à-vis a "digital dollar." For an overview of CBDC projects see for example Auer et al. (2020).

[2] After the financial crisis cash balances have risen in some monetary areas and declined in others. In any case the fall in money multipliers is predominantly a consequence of increased reserves holdings by banks. On the "Chicago Plan," see for example Knight et al. (1933) and Fisher (1935), and on the "Vollgeld" initiative, see vollgeld-initiative.ch.

produce using labor and physical capital. Banks invest in capital and reserves and fund themselves on the market for deposits where competition is limited. The central bank issues CBDC and reserves, which banks use to settle interbank payments. Deposits, CBDC, and reserves may differ both in terms of their liquidity services and with respect to the resources they require to operate the payment system. The main policy instruments include the interest rates on reserves and on CBDC as well as deposit subsidies which help to counteract markup distortions.

Our analysis starts with a transparent characterization result. The equilibrium conditions in a single-tiered, two-tiered, or mixed monetary system parallel those of the standard real business cycle (RBC) model, augmented by "pseudo wedges" which summarize the equilibrium effects of the policy instruments and the payment system characteristics. Policy adjustments as well as shocks and structural change in the payment system thus induce growth and business cycle dynamics in parallel to fiscal policy shocks in an RBC model with fiscal wedges.

In a two-tiered monetary system, the monetary transmission mechanism operates through banks. The equilibrium liquidity premium on deposits reflects bank market power, bank subsidies, the spread on reserves, as well as structural factors that determine the benefits for banks of holding reserves. We allow for two types of such benefits, internal and external. Banks internalize the former but not the latter, for example because they take into account that liquidity shortages can trigger costly asset liquidation while disregarding the costs of fire sales for other financial institutions.

In equilibrium, a higher spread on reserves as well as higher external costs of bank illiquidity increase the deposit spread. In contrast, higher internal costs may raise or lower the deposit spread because they induce banks to increase their reserves holdings and the externalities from reserves holdings amplify the resulting cost reductions. The same amplification mechanism affects the fiscal costs and benefits of interventions. To induce banks to raise the equilibrium deposit rate the central bank may increase interest on reserves or deposit subsidies. The former option always requires fewer fiscal resources because externalities from reserves holdings amplify the impact on banks' marginal costs.

In a single-tiered system the monetary transmission mechanism is much more direct. Since households only hold CBDC in such a system the spread on their means of payment is under the direct control of the central bank and there is no need to provide incentives for intermediaries. In a mixed system with both deposits and CBDC the transmission mechanism depends on the type of monetary policy instruments (e.g., the quantity of CBDC vs. its spread) and their numerical values.

Next, we analyze optimality. We start by characterizing the outcome chosen by a social planner that is only constrained by production and payment technologies. Not surprisingly, a modified Friedman (1969) rule emerges: Households are satiated with liquidity up to the point where the benefits equal the resource costs of liquidity management in the payment system. With CBDC these costs are the CBDC operating costs, which are borne by the central bank. With deposits they are the deposit operating costs, which are borne by commercial banks, as well as the costs for reserve operations in connection with interbank payments, which are borne by the central bank. The planner relies on the cost minimizing option, taking the liquidity benefits for households of deposits relative to

CBDC into account.

The Ramsey government controls the allocation and liquidity provision only indirectly, by way of setting the policy instruments. Nevertheless, it implements the social planner outcome as we show. In a purely CBDC-based system this is simple: The central bank sets the spread on CBDC at a level that reflects the social costs of CBDC and it prices banks out of the market by making reserves holdings sufficiently expensive. In a two-tiered system the situation is more challenging since the government needs to correct two distortions—due to market power and externalities in the banking sector—but this is feasible because deposit subsidies and the spread on reserves offer two independent levers. Even in a two-tiered system the Ramsey policy therefore implements the first best.

The optimal subsidy on bank deposits may be positive or negative, depending on the importance of bank market power on the one hand and externalities on the other. When reserves holdings generate external benefits the Ramsey policy subsidizes them by raising the interest rate on reserves. This aligns bank incentives with societal tradeoffs along the reserves margin but it distorts bank incentives along another margin because it encourages a longer balance sheet and deposit funding. If this latter effect outweighs the opposing force due to market power then the optimal deposit subsidy is negative, i.e., deposits are taxed. When external benefits of reserves holdings are absent, in contrast, then the optimal spread on reserves equals the social costs of reserve operations and the Ramsey government subsidizes deposits.

Under standard functional form assumptions we derive intuitive optimal policy rules. In a single-tiered system this rule is straightforward, as mentioned before. In a two-tiered system the interest rate rule for reserves is strikingly simple as well: To provide banks with the appropriate amount of liquidity the spread on reserves should equal the central bank's marginal operating costs multiplied with a factor that equals banks' internal benefits of reserves holdings relative to the total benefits. The accompanying policy rule for deposit subsidies or taxes assures that banks' optimizing choices conform with the efficient transmission of liquidity to households.

When deposits are the resource-efficient retail means of payment but a deposit subsidy/tax is unavailable optimal liquidity provision requires an alternative instrument to complement interest on reserves. In some circumstances, namely when the Ramsey policy subsidizes deposits, a suitably chosen CBDC interest rate target constitutes such an alternative; it forces banks to raise the interest rate on deposits in order to stay in business. An interest rate target for reserves and another one for CBDC, which solely serves as a threat to banks, thus can implement the first best.[3]

In other circumstances, namely when the Ramsey policy taxes deposits (because strong externalities require a high interest rate on reserves), the lack of a deposit subsidy/tax instrument has graver consequences. In this case CBDC cannot replace the missing instrument. Intuitively, when equilibrium deposit rates are inefficiently high introducing a competing means of payment—CBDC—cannot force banks to lower them. We investigate a range of candidate second-best policies in such an environment and find that a target for the composition of real balances—CBDC vs. deposits—is the most effective complement

[3]See Andolfatto (2021) for a similar result.

to interest on reserves. Targeting the composition lets the central bank control retail liquidity provision conditional on the markup that banks choose. In contrast, targeting the quantity of CBDC does not afford this possibility.

A recurrent theme in our first- and second-best analyses is that the two central bank liabilities, reserves and CBDC, should pay different interest rates—independently of whether CBDC is actually used as a means of payment or just as a disciplining device. Intuitively, when a means of payment circulates then its spread should reflect operating costs and externalities—and these are generally different for reserves and CBDC. And when CBDC does not circulate but is only used to discipline banks then the interest rate on CBDC should reflect the central bank's target for the deposit rate which is only indirectly linked to the social costs and benefits of reserves.

Our next set of results concerns equivalence. We consider a scenario in which CBDC operations in the single-tiered system require the same resources per unit of liquidity services for households as deposit and reserve operations in the equilibrium of the two-tiered system. We then ask whether a substitution of CBDC for deposits would alter the equilibrium allocation and the price system. This substitution could be partial or, hypothetically, it could be complete as in a "Chicago Plan" or "Vollgeld" regime without private means of payment.[4] Extending results in Brunnermeier and Niepelt (2019) we establish that a suitably chosen central bank intervention can always neutralize the effects of the introduction of CBDC.

In the equivalent equilibrium with CBDC the length of households' balance sheets does not change since households swap assets (deposits, CBDC, and capital). In contrast, banks shorten their balance sheets because they hold fewer reserves while maintaining the exposure to physical capital; that is, banks continue to extend credit to main street as in the initial two-tiered equilibrium. The funding that compensates for banks' reduced deposit liabilities net of reserves holdings stems from the central bank: The monetary authority extends loans to banks which in turn are financed by the newly issued CBDC net of banks' reduced reserves holdings.

In order to induce noncompetitive banks to go along with these equivalent balance sheet positions the central bank's loan funding schedule must replicate the deposit funding schedule of households. In particular, the interest rate on central bank loans to banks must correspond to banks' costs of sourcing deposit funding for capital investments. Since these costs reflect the interest rate on deposits, deposit operating costs, and the fact that banks hold reserves, the equivalent central bank loan interest rate depends on the interest rates on deposits and reserves, operating costs, bank subsidies, and the equilibrium reserves-to-deposits ratio. The formula for the equivalent central bank loan interest rate that we derive holds independently of specific functional form assumptions and much more widely than in the context of our specific model.

Our equivalence result cautions against claims that the introduction of CBDC would imply bank disintermediation, higher bank funding costs, a reduction of credit, or have other potentially harmful macroeconomic consequences. None of this is the case as long as the central bank chooses to refinance banks. Of course, the central bank also has other

[4] A complete substitution is unrealistic since "money" is most difficult to define in a coherent, legally binding way and its creation by banks or other actors even harder to suppress.

options and could implement other equilibria, better or worse; that is, CBDC enlarges the set of implementable equilibria while preserving the option to maintain the status quo.

While the equivalence logic far transcends the specific framework of this paper it does require certain model features, which we examine in detail. Most importantly, equivalence requires that deposit funding is not "special" in ways that central bank funding cannot be. We discuss the plausibility of this restriction and point to arguments in the banking literature that might lead one to question it. We also discuss the fact that, if deposits in the initial equilibrium are not collateralized, equivalence requires the central bank loan to be uncollateralized as well. We argue that it may be inconsistent to reject uncollateralized loans in an equivalent CBDC arrangement while simultaneously tolerating implicit lender-of-last-resort guarantees in the contemporaneous two-tiered system.

As mentioned before, banks' marginal costs of sourcing deposit funding for capital investments are given by the equivalent central bank loan interest rate. Using U.S. data for the period 1999 to 2021 we compute this rate and find that it fell by roughly 300 basis points over the sample period. It exhibited frequent fluctuations, reflecting a huge increase in banks' reserves holdings at the onset of the financial crisis, time-varying interest rate compression, and other factors.

We also quantify the funding cost reduction for banks due to their ability to issue liquid liabilities. This funding cost reduction, which can be interpreted as an implicit subsidy for banks, is the product of two terms: The quantity of net funding that banks generate by issuing deposits, namely the deposit base net of reserves holdings; and the liquidity premium on that funding net of operating costs. We find that the funding cost reduction for U.S. banks amounted to 0.5 to 1.5 percent of GDP just before the financial crisis and also more recently. In other years, banks did not benefit from cost reductions, or they even bore additional funding costs. These numbers compare with NIPA financial sector profits on the order of 3 percent of GDP prior to the financial crisis, negative profits during the crisis, and 2 to 3 percent after the financial crisis.

Finally, we consider political economy implications. We argue that the introduction of CBDC would expose banks and the economy to political risks: While the central bank could in principle ensure that CBDC has minor or no macroeconomic implications, doing so would require transparent refinancing of banks at terms that differ from market prices. In some periods the refinancing would have to be subsidized to guarantee equivalence rendering the equivalent arrangement politically unpalatable and non-implementable.

We also show that optimal monetary policy in a two-tiered system requires taxpayer support—even in the absence of crises—while this is not the case in a single-tiered system with CBDC. Intuitively, when the central bank supports the efficient monetary transmission with banks then it needs to correct distortions in the banking sector, which requires fiscal resources. In a single-tiered system, in contrast, efficiency only calls for the Friedman rule which implies a balanced central bank budget.

We argue that these different fiscal requirements could work towards strengthening the political support for CBDC if taxpayers are politically more influential than the owners of banks. We also discuss implications for central bank independence, which might suffer in a two-tiered system when the optimal policy requires fiscal support from the treasury.

Related Literature The paper relates to an old literature on two-tiered monetary systems and their properties. While money-multiplier analysis dates back at least to the 1940s, Gurley and Shaw (1960) introduce the distinction between inside money issued by banks and outside money supplied by the government. Tobin (1963; 1969; 1985) discusses the fractional reserve banking system and proposes a precursor to CBDC. More recently, Bullard and Smith (2003) study inside and outside money with frictions due to spatial separation. Benes and Kumhof (2012) simulate deposit creation in a New Keynesian DSGE model and find that fractional reserve banking generates instability and higher debt levels. Chari and Phelan (2014) emphasize negative externalities of fractional reserve banking when central bank money is scarce while Taudien (2020) argues that inside money fosters production by lowering producers' financing costs. Andolfatto (2018) contrasts mainstream and heterodox views of the macroeconomic role of fractional reserve banking and Faure and Gersbach (2018) compare allocations with and without private money creation. Farhi and Tirole (2021) emphasize complementarities between bank lending, insured deposits, lender of last resort facilities, and prudential supervision in fractional reserve banking systems. Jackson and Pennacchi (2021) contrast liquidity (safe asset) creation by the private and the public sector.

Following literatures in macroeconomics and banking we assume that reserves affect banks' operating costs in a deposit-based payment system (e.g., Bolton et al., 2020; Vandeweyer, 2019). In Kiyotaki and Moore (2019) a liquid security, like reserves, relaxes resalability constraints and increases productivity (reduces costs). Bianchi and Bigio (2020) carefully model the portfolio choice of banks and monetary policy transmission in a frictional interbank market; the shocks in their framework, which relate to matching frictions and the volatility of deposit withdrawals, map into the structural shocks we posit. Parlour et al. (2020) analyze the "liquidity externality" that arises because banks hold reserves against claims issued by their competitors. In our model reserves holdings exert positive externalities in addition to lowering a bank's operating costs.

A recent literature analyzes potential implications of a retail CBDC.[5] Barrdear and Kumhof (2016) simulate a rich DSGE model in which CBDC lowers real interest rates and transaction costs. In Keister and Sanches (2020) CBDC promotes exchange but may crowd out deposits which finance investment; the resulting tradeoffs vary with the substitutability of cash, deposits, and CBDC. Böser and Gersbach (2020) argue that CBDC in combination with tight central bank collateral requirements can render bank lending nonviable. Schilling et al. (2020) model CBDC as a central bank liability that is demandable against aggregate output, which gives rise to a tradeoff between price stability and the absence of runs. In Piazzesi and Schneider (2021) balance sheet length is costly, credit lines provide liquidity off balance sheet, and precautionary CBDC holdings give rise to real effects. In our model operating costs are proportional to the stock of liquid liabilities, liquidity services are substitutable, and bank funding may come from different sources; equivalence only hinges on the condition that the resource costs to manage payments are commensurate across monetary architectures.

We follow Klein (1971), Monti (1972) and a more recent literature including Drechsler

[5] See Niepelt (2018) for a review of early contributions to this literature.

et al. (2017) in stipulating a noncompetitive deposit market.[6] In Andolfatto (2021) the introduction of CBDC leads noncompetitive banks to raise the deposit rate, with positive effects on financial inclusion, and in Garratt and Zhu (2021) this affects the market structure. Chiu et al. (2019) quantitatively assess the implications of CBDC in a framework that combines elements of Keister and Sanches (2020) and Andolfatto (2021). We show that the disciplining role of CBDC can be more nuanced. When banks set an inefficiently low deposit rate the central bank can correct this by offering CBDC as an off-equilibrium alternative, as in Andolfatto (2021); but when the equilibrium deposit rate is inefficiently high (because the central bank subsidizes reserves to correct externalities) then this is not possible.

Keister and Monnet (2020) analyze how CBDC could shape information asymmetries between banks and the central bank and by implication, financial stability. Garratt and van Oordt (2021) argue that CBDC (electronic cash) could help limit negative externalities due to data collection on digital payment platforms (see also Kahn et al., 2005). In Williamson (2019) CBDC is a more efficient payment instrument than cash but balance sheet lengthening by the central bank creates collateral scarcity. We take a decidedly macroeconomic approach, both in terms of the workhorse model we adopt and the issues we emphasize, namely liquidity provision, the financing of intermediaries, and the role of monetary architecture. Moreover, we abstract from cash but this is without loss of generality as we argue in section 2.[7]

Our normative analysis complements large literatures building on "New Keynesian" and "New Monetarist" frameworks (Woodford, 2003; Galí, 2015; Rocheteau and Nosal, 2017). Unlike the former literature we emphasize the role of money as a means of payment and store of value rather than unit of account and we abstract from nominal rigidities; unlike the latter but in line with the literature following Sidrauski (1967) (e.g., Di Tella, 2020) we emphasize connections with the workhorse macroeconomic model rather than the market micro structure underlying the supply and demand for liquidity.

Our equivalence result builds on Brunnermeier and Niepelt (2019) and Niepelt (2020) and indirectly on seminal contributions such as Wallace (1981), Bryant (1983), Chamley and Polemarchakis (1984), or Sargent (1987, 5.4). Relative to this literature we emphasize liquidity considerations, introduce resource costs of operating single- or two-tiered payment systems, and focus on the CBDC application. Moreover, we exploit our equivalence result to derive an equivalent central bank loan interest rate and to quantify implicit bank subsidies in the U.S. monetary system.

Structure of the Paper The remainder of the paper is structured as follows. Section 2 lays out the monetary economy. Section 3 characterizes general equilibrium and discusses its properties including the transmission mechanisms in different monetary systems. In section 4 we analyze the social planner outcome, the Ramsey government's choice of monetary architecture and monetary policy, as well as second-best policies. Section 5 assesses the equivalence of monetary systems with deposits and reserves versus CBDC,

[6]See also Drechsler et al. (2021) and, related, Di Tella and Kurlat (2021).

[7]Design options for retail CBDC are the focus of Kahn et al. (2018), Kumhof and Noone (2018), Bindseil (2020), and Auer and Böhme (2020), among others.

and it quantifies the funding cost reduction for banks in a deposit-based system. Section 6 considers political economy aspects and section 7 concludes.

2 A Monetary Economy

We consider an infinite horizon production economy with a continuum of mass one of homogeneous infinitely-lived households. The households own a succession of two-period-lived, monopsonistic banks and of one-period-lived, competitive firms. Monetary and fiscal policy is determined by a consolidated government/central bank.

The model extends the environment in Sidrauski (1967) threefold. First, it features two means of payment at the retail level, one issued by banks and the other by the central bank. One can think of the former as deposits and of the latter as retail central bank digital currency (CBDC) or "reserves for all"—a digital medium of exchange issued by the central bank and accessible to households. Second, the model allows for resource costs of managing the payment system.[8] Finally, the model features a reserves layer, consistent with the typical arrangement in modern economies: The payments that banks make on behalf of their customers are settled with reserves through a central-bank-run clearing system.

2.1 Households

The representative household takes prices, returns, profits, and taxes as given and solves

$$\max_{\{c_t,x_t,k_{t+1},m_{t+1},n_{t+1}\}_{t\geq 0}} \sum_{t=0}^{\infty} \beta^t \mathbb{E}_0[u(c_t, z_{t+1}, x_t)]$$

$$\text{s.t.} \quad k_{t+1} + m_{t+1} + n_{t+1} = k_t R_t^k + m_t R_t^m + n_t R_t^n + w_t(1 - x_t) + \pi_t - c_t - \tau_t, \quad (1)$$

$$k_{t+1}, m_{t+1}, n_{t+1} \geq 0.$$

Here, c_t and x_t denote household consumption of the good and leisure at date t, respectively, and z_{t+1} denotes "effective real balances" carried from t into $t+1$.

Effective real balances are a weighted sum of (real) base money held by nonbanks, m_{t+1}, and bank deposits, n_{t+1},

$$z_{t+1} \equiv \lambda_t m_{t+1} + n_{t+1}.$$

We typically refer to m_{t+1} as "money" (CBDC) and to n_{t+1} as "deposits." The parameter $\lambda_t > 0$ indexes the liquidity benefits of money relative to those of deposits. While we allow for $\lambda_t \neq 1$ and for variation of λ_t across time or histories none of this is required for the results. As we will point out along the way several of our results also go through when we allow λ_t to be endogenous.

Our interpretation of m_{t+1} as CBDC does not mean, of course, that a model of cash rather than CBDC would require a modified framework; after all, the two central bank

[8] These costs can alternatively be interpreted as costs of managing the assets backing the payment instruments.

liabilities closely resemble each other from a macroeconomic perspective even if their specific liquidity properties (reflected by λ_t) may differ substantially.[9] Including cash as a third retail means of payment in addition to deposits and CBDC would be largely irrelevant for the results as we discuss at the end of the section.

The felicity function u is increasing, strictly concave, and satisfies Inada conditions; the discount factor β is positive and strictly smaller than unity. As is well known, this "money in the utility function" specification can represent several monetary frictions, including a "shopping time" friction which renders real balances helpful to economize on time spent purchasing consumption goods.[10] All key results of the paper are robust to changes in the assumptions that generate a demand for liquidity. What matters is not why households demand liquidity services but that they do.

Equation (1) represents the household's budget constraint. It states that the household invests in capital, k_{t+1}, as well as real balances. The household finances these investments as well as consumption outlays and taxes, τ_t, out of wage income, which equals the product of the wage, w_t, and labor supply, $1-x_t$; distributed profits, π_t; and the gross return on its portfolio. The latter consists of the returns on capital, $k_t R_t^k$, money, $m_t R_t^m$, and deposits, $n_t R_t^n$. The gross rates of return on money and deposits, R_t^m and R_t^n respectively, reflect both nominal interest rates and inflation; the decomposition between the two components is irrelevant for now.[11]

In equilibrium capital holdings and real balances are strictly positive. The household's optimality conditions are given by the following Euler equations for k_{t+1}, m_{t+1}, n_{t+1}, and x_t:

$$\begin{aligned} u_c(c_t, z_{t+1}, x_t) &= \beta \mathbb{E}_t[R_{t+1}^k u_c(c_{t+1}, z_{t+2}, x_{t+1})], & (2) \\ u_c(c_t, z_{t+1}, x_t) &\geq \beta R_{t+1}^m \mathbb{E}_t[u_c(c_{t+1}, z_{t+2}, x_{t+1})] + \lambda_t u_z(c_t, z_{t+1}, x_t), \quad m_{t+1} \geq 0, \\ u_c(c_t, z_{t+1}, x_t) &\geq \beta R_{t+1}^n \mathbb{E}_t[u_c(c_{t+1}, z_{t+2}, x_{t+1})] + u_z(c_t, z_{t+1}, x_t), \quad n_{t+1} \geq 0, \\ u_x(c_t, z_{t+1}, x_t) &= u_c(c_t, z_{t+1}, x_t) w_t. & (3) \end{aligned}$$

The weak inequality in the Euler equation for m_{t+1} or n_{t+1}, respectively, holds with equality if m_{t+1} or n_{t+1} is strictly positive.

To express the Euler equations for m_{t+1} and n_{t+1} more compactly let $\lambda_t^m \equiv \lambda_t$, $\lambda_t^n \equiv 1$, and define the risk-free interest rate, R_{t+1}^f, as

$$R_{t+1}^f \equiv 1/\mathbb{E}_t[\mathrm{sdf}_{t+1}],$$

where $\mathrm{sdf}_{t+1} \equiv \beta u_c(c_{t+1}, z_{t+2}, x_{t+1})/u_c(c_t, z_{t+1}, x_t)$ denotes the stochastic discount factor. When the household holds payment instruments of type $i \in \{m, n\}$ then the associated

[9] If m_{t+1} represented cash it would seem natural to let λ_t be a function of m_{t+1} and n_{t+1} rather than a parameter to capture limited substitutability between deposit- and cash-based payments. As mentioned before we discuss this possibility along the way.

[10] See Saving (1971), McCallum and Goodfriend (1987), Feenstra (1986), and Croushore (1993).

[11] For simplicity, we do not index variables and parameters by history. Variables c_t, x_t, k_{t+1}, m_{t+1}, n_{t+1}, z_{t+1}, R_t^k, R_{t+1}^m, R_{t+1}^n, w_t, π_t, τ_t and parameter λ_t are measurable with respect to information available at date t. That is, real interest rates on deposits and money are risk-free (inflation risk is negligible, an assumption we make to keep the notation simple) while the rate of return on capital may be risky.

first-order condition reads

$$\lambda_t^i u_z(c_t, z_{t+1}, x_t) = u_c(c_t, z_{t+1}, x_t) \left(1 - \frac{R_{t+1}^i}{R_{t+1}^f}\right). \tag{4}$$

When the household holds both payment instruments then

$$R_{t+1}^f - R_{t+1}^m = \lambda_t (R_{t+1}^f - R_{t+1}^n). \tag{5}$$

According to equation (4) payment instrument i enjoys a liquidity premium when $\lambda_t^i u_z(c_t, z_{t+1}, x_t) > 0$. We denote this liquidity premium on payment instrument i by

$$\chi_{t+1}^i \equiv 1 - \frac{R_{t+1}^i}{R_{t+1}^f}.$$

Equivalently, $-\chi_{t+1}^i$ equals the spread on payment instrument i compared with a risk-free bond that does not provide liquidity services. When the household holds both payment instruments then, according to equation (5), the liquidity premium on money exceeds the premium on deposits if money is more liquid than deposits ($\lambda_t > 1$), and vice versa.

2.2 Banks

There is a finite number of regions of equal size with one bank per region. Households in a region can only access the regional bank so that banks are monopsonists in their regional deposit market.[12] Regional borders do not restrain any other type of transaction and households are residual claimants to aggregate bank profits.

A bank at date t issues deposits and invests in capital and reserves, r_{t+1}.[13] It takes rates of return on capital and reserves as well as the stochastic discount factor as given but chooses deposits subject to the deposit funding schedule of households. The government may subsidize deposits at rate θ_t. In the context of the equivalence analysis in section 5 we also allow the central bank to extend a loan, l_{t+1}, at the gross interest rate R_{t+1}^l to the bank.

To introduce a role for reserves we assume that larger reserves holdings relative to deposits reduce a bank's operating costs. (We need not stipulate a minimum reserves requirement.) A narrative that motivates this feature is that banks settle the payments they make among each other on behalf of their clients by transferring reserves. If a bank lacks reserves to cover net payments to other financial institutions then it needs to transfer ownership of capital but this generates resource costs related to, e.g., fire sales,

[12]Alternatively, we could assume that the banking sector is monopsonistically competitive, with effective real balances a composite of deposit holdings at different banks, or that several banks in a region compete à la Cournot. The optimality conditions in Cournot equilibrium closely parallel the conditions of a monopsony bank derived below (see, e.g., Freixas and Rochet, 2008).

[13]This is equivalent to assuming that banks extend loans which eventually fund physical capital accumulation.

deadweight losses, or legal charges. As a consequence, the unit resource costs of managing deposit-based payments, ν_t, are decreasing in r_{t+1} and increasing in n_{t+1}.[14]

We also allow ν_t to vary with the reserves and deposits of other banks. This allows to capture externalities of reserves holdings and specifically, positive externalities which may be present because fire sales by one bank increase the costs of other institutions. Formally, letting $\zeta_{t+1} \equiv r_{t+1}/n_{t+1}$ denote the reserves-to-deposits ratio (the "liquidity ratio") of a bank and $\bar{\zeta}_{t+1} \equiv \bar{r}_{t+1}/\bar{n}_{t+1}$ the aggregate reserves-to-deposits ratio of all other banks, we make the following assumptions about operating costs: When a bank has no deposits then its operating costs equal zero: $n_{t+1} = 0 \Rightarrow n_{t+1}\nu_t = 0$. When all other institutions have no deposits ($\bar{n}_{t+1} = 0$) then the bank's operating costs per unit of own deposits are large but bounded; this rules out asymmetric equilibria with $\bar{n}_{t+1} = 0$ but $n_{t+1} > 0$. And otherwise ($n_{t+1} > 0$ and $\bar{n}_{t+1} > 0$) ν_t is given by the function $\nu_t(\zeta_{t+1}, \bar{\zeta}_{t+1})$, which is strictly decreasing in both arguments, strictly convex, and satisfies $\nu_{12,t} = 0$ or $\nu_{11,t} \geq \nu_{22,t}$, as well as $\lim_{\zeta_{t+1}\downarrow 0} \nu_{1,t} = \infty$.[15] These assumptions imply that in equilibrium, $\zeta_{t+1} = \bar{\zeta}_{t+1}$ and reserves are strictly positive if and only if deposits are strictly positive.

The program of a bank at date t reads[16]

$$
\begin{aligned}
\max_{n_{t+1}, r_{t+1}} \quad & \pi^b_{1,t} + \mathbb{E}_t[\text{sdf}_{t+1}\ \pi^b_{2,t+1}] \\
\text{s.t.} \quad & \pi^b_{1,t} = -n_{t+1}(\nu_t(\zeta_{t+1}, \bar{\zeta}_{t+1}) - \theta_t), && (6) \\
& \pi^b_{2,t+1} = (n_{t+1} - r_{t+1})R^k_{t+1} + r_{t+1}R^r_{t+1} - n_{t+1}R^n_{t+1}, && (7) \\
& R^n_{t+1} \text{ reflects deposit funding schedule,} \\
& n_{t+1} \geq 0,
\end{aligned}
$$

where $\pi^b_{1,t}$ and $\pi^b_{2,t+1}$ denote the cash flows generated in the first and second period of the bank's operations, respectively. The first constraint relates the cash flow in the first period to the operating costs net of subsidies. The second constraint relates the cash flow in the second period to the gross yield on physical capital and reserves holdings net of gross interest payments to depositors. The third constraint states that the bank takes the deposit funding schedule (rather than the deposit rate) as given because it is a monopsonist. We assume, and later verify, that the funding schedule is differentiable.

The marginal effect of n_{t+1} on the bank's objective is given by (dropping arguments of the ν_t function)

$$-(\nu_t(\cdot) - \theta_t) + \nu_{1,t}(\cdot)\zeta_{t+1} + \mathbb{E}_t[\text{sdf}_{t+1}(R^k_{t+1} - R^n_{t+1} - n_{t+1}{R^n_{t+1}}'(n_{t+1}))].$$

[14] See the earlier literature review on models of the demand for reserves and of interbank payments. In a more general framework operating costs might also depend on holdings of other securities, for instance government bonds that serve as collateral.

[15] We denote the partial derivatives of ν_t with respect to the first and second argument, respectively, by $\nu_{1,t}$ and $\nu_{2,t}$ and we use similar notation for higher order derivatives.

[16] Variables $r_{t+1}, \zeta_{t+1}, \text{sdf}_t, R^r_{t+1}, \pi^b_{1,t}, \pi^b_{2,t}, \theta_t$, and ν_t (as well as l_{t+1} and R^l_{t+1} in the program with central bank loans) are measurable with respect to information available at date t. We do not normalize the typical bank's portfolio positions by the number of banks. That is, we state the conditions as they apply for the banking sector as a whole.

Using the household's Euler equation the first-order condition with respect to n_{t+1} therefore reads

$$-(\nu_t(\cdot)-\theta_t)+\nu_{1,t}(\cdot)\zeta_{t+1}+\chi^n_{t+1}\leq n_{t+1}R^{n}_{t+1}{}'(n_{t+1})/R^f_{t+1},\quad n_{t+1}\geq 0. \tag{8}$$

The left-hand side of inequality (8) represents the marginal profit from deposit issuance, holding the interest rate on deposits constant: Managing the marginal unit of deposits costs $\nu_t(\cdot)-\theta_t$ and in addition, the marginal unit increases the operating costs for inframarginal units, but it also yields a gain if the deposit liquidity premium is positive, $\chi^n_{t+1}>0$. The right-hand side equals the profit loss on inframarginal deposits, which results because higher deposit issuance is associated with an increase in R^n_{t+1}. When deposits are strictly positive then inequality (8) simplifies to

$$\chi^n_{t+1}-(\nu_t(\cdot)-\theta_t-\nu_{1,t}(\cdot)\zeta_{t+1})=\frac{1}{\eta_{n,t+1}}\frac{R^n_{t+1}}{R^f_{t+1}},$$

where $\eta_{n,t+1}$ denotes the elasticity of deposit funding with respect to R^n_{t+1} (see Klein, 1971; Monti, 1972). This elasticity may depend on central bank choices, in particular on whether—and how elastically—the central bank issues m_{t+1}. We address this in detail in subsequent sections.

Turning to reserves, when $n_{t+1}>0$ such that reserves are interior then the corresponding first-order condition reads

$$-\nu_{1,t}(\zeta_{t+1},\bar{\zeta}_{t+1})=1-\frac{R^r_{t+1}}{R^f_{t+1}}. \tag{9}$$

Intuitively, the optimal choice of reserves equalizes the (private) gain due to lower operating costs and the loss due to the bank's lower return when the spread on reserves, $\chi^r_{t+1}\equiv 1-R^r_{t+1}/R^f_{t+1}$, is positive. Since in equilibrium $\zeta_{t+1}=\bar{\zeta}_{t+1}$ equation (9) implies a unique mapping from the opportunity costs of holding reserves to the equilibrium reserves-to-deposits ratio, which we write as $\zeta_{t+1}=\nu^{-1}_{1,t}(-\chi^r_{t+1})$.[17]

Combining equations (8) and (9) implies that as long as $n_{t+1}>0$,

$$\chi^n_{t+1}-\left(\tilde{\nu}_t(-\chi^r_{t+1})-\theta_t\right)=\frac{1}{\eta_{n,t+1}}\frac{R^n_{t+1}}{R^f_{t+1}}, \tag{8a}$$

where we define

$$\tilde{\nu}_t(-\chi)\equiv\nu_t\left(\nu^{-1}_{1,t}(-\chi),\nu^{-1}_{1,t}(-\chi)\right)+\chi\nu^{-1}_{1,t}(-\chi).$$

Function $\tilde{\nu}_t$ summarizes the direct and indirect effect of deposits on the bank's operating costs: A marginal increase in deposit issuance raises the total operating costs not only

[17] By the mean value theorem, for any $\varepsilon>0$ there exists a $\iota\in(0,\varepsilon)$ such that $\nu_{1,t}(\zeta+\varepsilon,\zeta+\varepsilon)=\nu_{1,t}(\zeta,\zeta)+(\nu_{11,t}(\zeta+\iota,\zeta+\iota)+\nu_{12,t}(\zeta+\iota,\zeta+\iota))\varepsilon$. We claim that $\nu_{11,t}+\nu_{12,t}>0$ such that the function $\nu_{1,t}(\zeta,\zeta)$ is monotonically increasing in ζ and therefore invertible. Strict convexity of ν_t implies that its Hessian is positive definite such that, in particular, $\nu_{11,t}>0$ and $\nu_{11,t}+2\nu_{12,t}+\nu_{22,t}>0$. Recall that we assume $\nu_{12,t}=0$ or $\nu_{11,t}\geq\nu_{22,t}$. In the former case the claim is directly established. In the latter case as well because $\nu_{11,t}+\nu_{12,t}>\nu_{11,t}-(\nu_{11,t}+\nu_{22,t})/2=(\nu_{11,t}-\nu_{22,t})/2\geq 0$.

directly, by $\nu_t(\zeta_{t+1}, \zeta_{t+1})$, but also indirectly because $\nu_t(\zeta_{t+1}, \zeta_{t+1})$ increases. The magnitude of these effects is pinned down by the opportunity costs of reserves and the bank's first-order condition for reserves, equation (9). A parallel condition holds when the bank borrows from the central bank.[18]

It sometimes is instructive to consider a simplified variant of the model without the reserves layer. In that variant ν_t is exogenous and banks therefore do not hold reserves. Accordingly, $\tilde{\nu}_t(-\chi^r_{t+1})$ is replaced by the exogenous ν_t in that case.

2.3 Firms

Firms rent capital, κ_t, and labor, ℓ_t, to produce the output good. They take wages, the rental rate of capital, $R^k_t - 1 + \delta$, and the goods price as given; the rental rate reflects the depreciation rate, δ.[19] Without loss of generality we abstract from liquidity demand by firms.[20] Letting f_t denote a neoclassical production function the representative firm solves

$$\begin{aligned} \max_{\kappa_t, \ell_t} \quad & \pi^f_t \\ \text{s.t.} \quad & \pi^f_t = f_t(\kappa_t, \ell_t) - \kappa_t(R^k_t - 1 + \delta) - w_t \ell_t \end{aligned} \tag{10}$$

and the first-order conditions read

$$R^k_t - 1 + \delta = f_{\kappa,t}(\kappa_t, \ell_t), \tag{11}$$

$$w_t = f_{\ell,t}(\kappa_t, \ell_t). \tag{12}$$

Since f_t exhibits constant returns to scale and firms are competitive, equilibrium profits π^f_t equal zero.

2.4 Government

The consolidated government collects taxes, pays deposit subsidies, invests in capital, k^g_{t+1}, and issues money and reserves. The unit resource costs of managing money-based payments equal μ_t, and the unit resource costs of managing reserves-based payments among banks equal ρ_t. Accordingly, the government budget constraint reads[21]

$$k^g_{t+1} - m_{t+1} - r_{t+1} = k^g_t R^k_t - m_t R^m_t - r_t R^r_t + \tau_t - n_{t+1}\theta_t - m_{t+1}\mu_t - r_{t+1}\rho_t. \tag{13}$$

Central bank liabilities are injected through open market operations. That is, banks exchange some of their capital holdings (which they acquire from households in exchange for deposits) against reserves, and households similarly exchange capital against money.

[18] In this case the bank chooses l_{t+1} subject to the loan funding schedule.

[19] Variables π^f_t, κ_{t+1} and ℓ_t are measurable with respect to information available at date t.

[20] But note the discussion of heterogeneity in section 5.

[21] Variable k^g_{t+1} is measurable with respect to information available at date t. When the central bank extends loans to banks two additional terms are included in the constraint: l_{t+1} on the left-hand side and $l_t R^l_t$ on the right-hand side.

In the central bank's balance sheet reserves and money thus are "backed" by capital. Equivalently (since the government has access to nondistorting taxes) the central bank might inject means of payment by transfer. When a household wants to exchange deposits against money and the central bank accepts the incoming payment from the household's bank then the bank's reserves account at the central bank is debited or the central bank extends a loan to the bank. (See the discussion in section 5.)

2.5 Market Clearing

Each household is endowed with one unit of time per period. Labor and capital market clearing as well as the bank's balance sheet identity and the definition of total profits then imply[22]

$$\ell_t = 1 - x_t, \quad \kappa_t = k_t + k_t^g + n_t - r_t, \quad \pi_t = \pi_{1,t}^b + \pi_{2,t}^b + \pi_t^f. \tag{14}$$

2.6 Resource Constraint

Walras' law implies that market clearing on the markets for labor and capital as well as the budget constraints of households, banks, firms, and the government imply market clearing on the market for the output good. Specifically, combining equations (1), (6), (7), (10), (13), and (14) yields the resource constraint

$$\kappa_{t+1} = f_t(\kappa_t, 1 - x_t) + \kappa_t(1 - \delta) - c_t - m_{t+1}\mu_t - n_{t+1}\nu_t(\zeta_{t+1}, \zeta_{t+1}) - r_{t+1}\rho_t. \tag{15}$$

2.7 Policy and Equilibrium

Let $\xi_{t+1} \equiv n_{t+1}/z_{t+1}$ denote the share of deposits in effective real balances.

A *policy* $\mathcal{P}$ consists of $\{\tau_t, \theta_t\}_{t\geq 0}$; $\{\chi_{t+1}^r\}_{t\geq 0}$ if the central bank issues reserves; and $\{m_{t+1}\}_{t\geq 0}$, $\{\chi_{t+1}^m\}_{t\geq 0}$, or $\{\xi_{t+1}\}_{t\geq 0}$ if the central bank issues money and targets its quantity, spread, or share in effective real balances, respectively.

An *equilibrium* conditional on policy $\mathcal{P}$ consists of

- a positive allocation, $\{c_t, x_t, k_{t+1}, k_{t+1}^g, \kappa_{t+1}, \ell_t\}_{t\geq 0}$;
- positive money, deposit, and reserves holdings, $\{m_{t+1}, n_{t+1}, r_{t+1}\}_{t\geq 0}$;
- and a positive (shadow) price system, $\{w_t, R_{t+1}^k, R_{t+1}^f, \chi_{t+1}^m, \chi_{t+1}^n, \chi_{t+1}^r\}_{t\geq 0}$,

such that (1)–(14) (and by implication (15)) are satisfied and asset markets clear. If the central bank extends loans then the policy also includes a loan funding schedule, the equilibrium objects also include $\{l_{t+1}, R_{t+1}^l\}_{t\geq 0}$, and the loan market must clear as well.

[22] When the central bank extends loans to banks the second identity reads $\kappa_t = k_t + k_t^g + n_t + l_t - r_t$.

2.8 Functional Form Assumptions

Existence and uniqueness of equilibrium in models with money in the utility function may require conditions on primitives (e.g., Walsh, 2017). In parts of the analysis we will impose the following functional form assumptions:

Assumption 1. Preferences and the operating costs for deposit-based payments, respectively, satisfy

$$\begin{aligned} u(c_t, z_{t+1}, x_t) &= \left((1-\vartheta)c_t^{1-\psi} + \vartheta z_{t+1}^{1-\psi}\right)^{\frac{1-\sigma}{1-\psi}} (1-\sigma)^{-1} v(x_t), \\ \nu_t(\zeta_{t+1}, \bar{\zeta}_{t+1}) &= \phi_{1,t}\zeta_{t+1}^{1-\varphi} + \phi_{2,t}\bar{\zeta}_{t+1}^{1-\varphi}, \end{aligned}$$

where $\vartheta, \psi \in (0,1)$; $\sigma > 0, \neq 1$; $\phi_{1,t} > 0, \phi_{2,t} \geq 0$; $\varphi > 1$. Function v is strictly increasing and concave.

This preference specification allows for a balanced growth path along which consumption, wages, and asset holdings grow at the same rate while interest rates, spreads, labor supply and leisure remain constant (King et al., 1988). The elasticity of substitution between c_t and z_{t+1} equals ψ^{-1}, and consumption and real balances are complements as long as $\psi > \sigma$.

Simpler variants of the model result when we assume that $v'(x_t) = 0$ or $\varphi = 1$. When $v'(x_t) = 0$ the model reduces to a specification without leisure. When $\varphi = 1$ the reserves layer disappears.

2.9 Cash

Conceptually, it would be straightforward to include a third central bank liability (in addition to reserves and money) and third retail means of payment (in addition to money and deposits) in the model and to interpret this additional instrument as cash. In the interest of transparency and to maintain the focus of our analysis we opt against this possibility.

There are two ways to think about the interplay between cash and CBDC: Either the introduction of CBDC would leave cash demand largely unaffected, or it would not. In the former case, the inclusion of both means of payment in the model would be uninteresting.[23] In the latter case, it would allow to analyze cash-CBDC substitution, a swap of central bank liabilities with no immediate macroeconomic consequences. Our framework focuses on the "disruptive" deposit-CBDC substitution that affects the banking sector and which is independent of the role of cash.

[23]Cash could affect the bargaining power of households vis-à-vis banks already before CBDC is introduced, as in Lagos and Zhang (2020). Reduced market power of banks would affect the elasticity of deposit funding without changing our main conclusions.

3 General Equilibrium

3.1 Model Solution

We solve the model under assumption 1 focusing on the case of a two-tiered system with deposits and reserves where the spread on reserves and the deposit subsidy are the policy instruments. Appendix A contains detailed derivations. In the appendix we also solve for the equilibrium allocation in a system with money but no deposits, or with both money and deposits under different assumptions about monetary policy.

The key steps in the derivation are first, to show that households hold deposits in proportion to consumption and depending on the deposit spread, χ^n_{t+1}. Second, we solve the bank's problem given the demand for real balances by households. This yields closed form solutions that relate χ^n_{t+1} and ζ_{t+1} to the policy instruments χ^r_{t+1} and θ_t as well as parameters. Finally, we reduce the equilibrium conditions to three fundamental equations,

$$
\begin{aligned}
c_t^{-\sigma} v(x_t) &= \beta \mathbb{E}_t \left[(1 - \delta + f_{\kappa,t+1}(\kappa_{t+1}, 1 - x_{t+1})) c_{t+1}^{-\sigma} v(x_{t+1}) \frac{\Omega^c_{t+1}(\chi^r_{t+2}, \theta_{t+1})}{\Omega^c_t(\chi^r_{t+1}, \theta_t)} \right], \\
\frac{c_t^{1-\sigma}}{1-\sigma} v'(x_t) &= c_t^{-\sigma} v(x_t) f_{\ell,t}(\kappa_t, 1 - x_t) \frac{\Omega^c_t(\chi^r_{t+1}, \theta_t)}{\Omega^x_t(\chi^r_{t+1}, \theta_t)}, \\
\kappa_{t+1} &= f_t(\kappa_t, 1 - x_t) + \kappa_t(1 - \delta) - c_t \Omega^{rc}_t(\chi^r_{t+1}, \theta_t).
\end{aligned}
$$

This system (plus transversality condition) parallels the equilibrium conditions in the basic real business cycle (RBC) model with felicity function $u(c_t, x_t) = (c_t^{1-\sigma} - 1)/(1 - \sigma)v(x_t)$ (King et al., 1988). The only difference are three "pseudo wedges:" A term in the intertemporal first-order condition, $\Omega^c_{t+1}(\chi^r_{t+2}, \theta_{t+1})/\Omega^c_t(\chi^r_{t+1}, \theta_t)$; a term in the intratemporal first-order condition, $\Omega^c_t(\chi^r_{t+1}, \theta_t)/\Omega^x_t(\chi^r_{t+1}, \theta_t)$; and a term multiplying consumption in the resource constraint, $\Omega^{rc}_t(\chi^r_{t+1}, \theta_t)$. The three functions $\Omega^i_t(\chi^r_{t+1}, \theta_t),\ i \in \{c, x, rc\}$, which are defined in appendix A, depend on policy, (χ^r_{t+1}, θ_t), and the payment technology, $\nu_t(\cdot)$ and ρ_t. Conditional on policy the model therefore solves exactly as an RBC model.[24]

The $\Omega^i_t(\chi^r_{t+1}, \theta_t)$ terms can directly be compared with the wedges in an RBC model with distorting labor-income and consumption taxes as well as government consumption. Unlike fiscal policy wedges in the RBC model, however, the pseudo wedges in our monetary framework should not be interpreted as efficiency wedges. They simply represent the additional terms relative to a nonmonetary, frictionless RBC model that deposits, reserves, and banks introduce into the equilibrium conditions. This becomes most evident by noting that $\lim_{\vartheta \to 0} \Omega^i_t(\chi^r_{t+1}, \theta_t) = 1$, i.e., the system reduces to the baseline RBC model when households do not value liquidity services.

Along a balanced growth path with constant policy the intertemporal pseudo wedge equals one. The same holds true when utility is separable between consumption and real balances ($\psi = \sigma$, see the appendix). In contrast, the intratemporal pseudo wedge and the wedge in the resource constraint differ from unity even in these special cases because $\Omega^c_t(\chi^r, \theta) \neq \Omega^x_t(\chi^r, \theta)$ and $\Omega^{rc}_t(\chi^r, \theta) \neq 1$.

[24] This would continue to be the case when distorting taxes were introduced as the derivations in the appendix make clear.

The equilibrium conditions in the environment with money but no deposits have the same structure as the conditions reported above but the expressions for the pseudo wedges differ reflecting the absence of banks. The first difference concerns households' liquidity demand, which depends on the liquidity premium that households face. In the environment with money this premium is given by χ^m_{t+1} and it is under the direct control of the central bank. In the environment with deposits, in contrast, it is given by χ^n_{t+1}, which is not directly controlled by the government but depends on the policy instruments χ^r_{t+1} and θ_t by way of the transmission mechanism in the banking sector. As a consequence the pseudo wedges in the environment with money contain fewer dependencies than with deposits.

The second difference concerns resource costs. In the environment with money the resource costs of the payment system are proportional to μ_t while with deposits, they are proportional to $\nu_t(\zeta_{t+1}, \zeta_{t+1}) + \zeta_{t+1}\rho_t$, which is endogenous because banks choose ζ_{t+1} in response to policy. This is reflected in the pseudo wedges in the resource constraint.

Finally, the equilibrium conditions in the environment with both money and deposits again have the same structure. Appendix A contains a detailed characterization.

3.2 Monetary Neutrality and Monetary-Policy Nonneutrality

In steady state with a stationary consumption-to-real-balances ratio (e.g., under assumption 1) the marginal utility of consumption is constant. Monetary policy does not affect the steady-state capital-labor ratio in this case; irrespective of whether money and/or deposits circulate, the ratio is pinned down by technology and preferences from the Euler equation for capital.[25] Similar conclusions follow for balanced growth paths as long as policy enters the equilibrium expression for the marginal utility of consumption separably (for example under assumption 1).

In contrast, the payment system and monetary policy do affect the consumption-labor ratio along a balanced growth path (and of course during the transition as well). We have already seen this in the case where assumption 1 holds. The nonneutrality arises because real balances affect the marginal utilities of consumption and leisure asymmetrically. Because of the wedge in the resource constraint policy and the payment system also affect output and its composition.

Clearly, the framework exhibits monetary neutrality. Whether it also exhibits superneutrality depends on how inflation is reflected in policy rates. When real risk-free rates are not affected by anticipated inflation while real rates on money or reserves fall because nominal rates respond inelastically, then households (directly or through banks) face higher liquidity premia and the demand for means of payment falls. Inflation differentially affects the marginal utilities of consumption and leisure in this case and therefore changes the allocation.[26]

[25] See Walsh (2017, 2.2.1) for a discussion of steady states in the Sidrauski (1967) model with nonstationary consumption-to-real-balances ratios. If real balances constituted an argument of the production function (with real balances and capital as substitutes) the model would feature a "Tobin effect" (Tobin, 1965; Fischer, 1972).

[26] See Walsh (2017, 2.4.2) for a discussion of nonsuperneutrality in the Sidrauski (1967) model.

3.3 Liquidity Premia and Monetary Policy Transmission

The presence of a "bank layer" in the deposit-based payment system implies a less direct monetary policy transmission than in the money-based system. To affect the liquidity premium on deposits that households face, the central bank must manipulate the cost structure of banks or their competitive environment. In section 4, we consider the latter option in detail, focusing on policies that target the quantity of money, its spread, or the composition of real balances. Here, we focus on the former option and analyze how the spread on reserves and the deposit subsidy affect the equilibrium.

Imposing assumption 1 the liquidity premium as a function of these two policy instruments satisfies

$$\tilde{\chi}^n_{t+1}(\chi^r_{t+1}, \theta_t) = \frac{(\phi_{1,t}\varphi + \phi_{2,t}) \left(\frac{\chi^r_{t+1}}{\phi_{1,t}(\varphi-1)} \right)^{1-\frac{1}{\varphi}} - \theta_t}{1-\psi}, \quad \varphi > 1, \ \psi \in (0,1),$$

(see appendix A). The numerator of the expression on the right-hand side of the equality represents the marginal costs of a bank that optimally chooses the reserves-to-deposits ratio[27]

$$\zeta_{t+1} = \left(\frac{\chi^r_{t+1}}{\phi_{1,t}(\varphi - 1)} \right)^{-\frac{1}{\varphi}}.$$

The equilibrium spread exceeds these marginal costs because the bank charges a markup as long as liquidity demand is not perfectly elastic (ψ^{-1} is not infinite). A deposit subsidy, θ_t, directly reduces marginal costs. In contrast, the spread on reserves, χ^r_{t+1}, is transmitted indirectly: A higher spread raises the interest that a bank foregoes on the share ζ_{t+1} of its assets and this leads the bank to reduce its reserves-to-deposits ratio, pushing operating costs up. In addition, the operating costs also rise because of the externality as other banks adapt correspondingly.

In terms of their fiscal implications the two transmission mechanisms differ. While a marginal increase in θ_t imposes fiscal costs n_{t+1} and lowers the deposit spread by $(1-\psi)^{-1}$, a marginal decrease in χ^r_{t+1} costs $n_{t+1}\zeta_{t+1}$ and lowers the deposit spread by $(\phi_{1,t}\varphi + \phi_{2,t})\zeta_{t+1}/(\varphi(1-\psi)\phi_{1,t})$. A reduction in χ^r_{t+1} thus increases the interest rate on deposits at lower fiscal costs than an increase in θ_t whenever $\phi_{2,t} > 0$, that is, when reserves holdings generate external benefits. This result does not only hold under assumption 1 because it is a consequence of the envelope theorem. It is relevant, for example, in the context of discussions about lower bounds on interest rates or "reversal rates."

Changes in payment technology, $\nu_t(\cdot)$, affect the equilibrium spread depending on whether they act on bank internal or external costs. Stronger externalities (higher $\phi_{2,t}$) raise the deposit spread. They increase the costs for a bank and these costs are partly shifted to households. Higher internal benefits of reserves (higher $\phi_{1,t}$) have two additional effects because they induce a bank to increase its reserves holdings. This lowers the costs for the bank itself and for all other banks (because of the externality) but it increases the interest the bank foregoes on its reserve holdings. A higher $\phi_{1,t}$ therefore raises marginal

[27] See conditions (8a) and (9) as well as conditions (8a') and (9') in the appendix.

costs and the spread through two channels and lowers them through one. We summarize these results as follows:

Proposition 1. With external benefits of reserves holdings,

i. an increase in the interest rate on reserves raises deposit rates at lower fiscal costs than an increase in the deposit subsidy;

ii. higher internal benefits can increase deposit rates. Under assumption 1 this is the case when $\phi_{1,t} + \phi_{2,t}(\varphi^{-1} - 1) < 0$.

Proof. (ii.) The derivative $d\tilde{\chi}^n_{t+1}(\chi^r_{t+1}, \theta_t)/d\phi_{1,t}$ is proportional to $\phi_{1,t} + \phi_{2,t}(\varphi^{-1} - 1)$. $\square$

4 Optimality

We now turn to the normative implications of the model.

4.1 Social Planner Allocation

The social planner solves

$$\max_{\{c_t, x_t, \kappa_{t+1}, m_{t+1}, n_{t+1}, r_{t+1}\}_{t\geq 0}} \sum_{t=0}^{\infty} \beta^t \mathbb{E}_0 \left[u(c_t, z_{t+1}, x_t)\right]$$

$$\text{s.t.} \quad \kappa_{t+1} = f_t(\kappa_t, 1 - x_t) + \kappa_t(1 - \delta) - c_t - m_{t+1}\mu_t - n_{t+1}\nu_t(\zeta_{t+1}, \zeta_{t+1}) - r_{t+1}\rho_t,$$

$$\kappa_{t+1}, m_{t+1}, n_{t+1}, r_{t+1} \geq 0.$$

The optimality conditions for capital, consumption, and leisure are standard and yield the same conditions as in decentralized equilibrium (see equations (2), (3), (11), (12), and (14)),

$$\begin{aligned} u_c(c_t, z_{t+1}, x_t) &= \beta \mathbb{E}_t \left[(1 - \delta + f_{\kappa,t+1}(\kappa_{t+1}, 1 - x_{t+1}))u_c(c_{t+1}, z_{t+2}, x_{t+1})\right], \\ u_x(c_t, z_{t+1}, x_t) &= u_c(c_t, z_{t+1}, x_t) f_{\ell,t}(\kappa_t, 1 - x_t). \end{aligned}$$

Optimal reserves holdings are characterized by

$$\begin{aligned} \nu_{1,t}(\zeta_{t+1}, \zeta_{t+1}) + \nu_{2,t}(\zeta_{t+1}, \zeta_{t+1}) + \rho_t &= 0 \ \text{ if } \ n_{t+1} > 0, \\ r_{t+1} &= 0 \ \text{ otherwise.} \end{aligned}$$

Intuitively, when the planner relies on deposits then it issues reserves up to the point where the cost reduction for deposit payments, $\nu_{1,t}(\zeta_{t+1}, \zeta_{t+1}) + \nu_{2,t}(\zeta_{t+1}, \zeta_{t+1})$, is balanced by the costs of managing the marginal reserve unit, ρ_t. When reserves generate an externality the planner takes this into account. This contrasts with the decentralized equilibrium where a bank only internalizes its own cost reduction from reserves holdings, see equation (9).

Strict convexity of the function ν_t implies a unique mapping from ρ_t to the optimal choice of reserves, conditional on n_{t+1}.[28]

Finally, the first-order conditions for money and deposits yield the optimality condition

$$u_z(c_t, z_{t+1}, x_t) = u_c(c_t, z_{t+1}, x_t) \min[\mu_t/\lambda_t, \nu_t(\zeta^\star_{t+1}, \zeta^\star_{t+1}) + \zeta^\star_{t+1}\rho_t], \tag{SP}$$

where we use the condition for reserves derived above[29] and $\zeta^\star_{t+1}$ denotes the reserves-to-deposits ratio chosen by the social planner when relying on deposits.

To interpret equation (SP), consider first the case where reserves do not affect the operating costs of deposit-based payments (the model variant without the reserves layer). In this case, equation (SP) reduces to

$$u_z(c_t, z_{t+1}, x_t) = u_c(c_t, z_{t+1}, x_t) \min[\mu_t/\lambda_t, \nu_t].$$

Intuitively, the social planner satiates the household with real balances up to the point where the marginal benefit of real balances equals the marginal resource costs, both expressed in utility terms. This is just a variant of the Friedman (1969) rule which, in its more common form, abstracts from resource costs of managing liquidity (i.e., posits $\min[\mu_t/\lambda_t, \nu_t] = 0$).

As for the source of liquidity, the optimality condition implies that the planner generically provides one means of payment (but see the discussion below). When λ_t strictly exceeds μ_t/ν_t then the first best involves money but no deposits, and vice versa. Only when $\mu_t/\nu_t = \lambda_t$—when the relative resource costs of money and deposits and their relative liquidity benefits happen to coincide—is the planner indifferent between the two payment instruments.

Consider next the case of interest where reserves do affect the operating costs of deposit-based payments such that ν_t is a function of ζ_{t+1}. In this case, equation (SP) holds and the optimal allocation still has the two properties emphasized before. The only new element is that the planner compares the resource costs of managing money with those of managing deposits and reserves.

Summarizing:

Proposition 2. The social planner provides the means of payment with the lowest resource costs per effective liquidity. It equalizes the marginal liquidity benefit of real balances and the marginal resource costs.

The result that the social planner generically relies on a single means of payment reflects the assumption that money and deposits enter additively in effective real balances. In a more general setting with a nonlinear aggregation of the two sources of liquidity it would typically be optimal to circulate both money and deposits.[30] For example, if real

[28] By the mean value theorem, for any $\varepsilon > 0$ there exists a $\iota \in (0, \varepsilon)$ such that $\nu_{1,t}(\zeta+\varepsilon, \zeta+\varepsilon) + \nu_{2,t}(\zeta+\varepsilon, \zeta+\varepsilon) = \nu_{1,t}(\zeta, \zeta) + \nu_{2,t}(\zeta, \zeta) + (\nu_{11,t}(\zeta+\iota, \zeta+\iota) + 2\nu_{12,t}(\zeta+\iota, \zeta+\iota) + \nu_{22,t}(\zeta+\iota, \zeta+\iota))\varepsilon$. Strict convexity of ν_t implies that its Hessian is positive definite such that, in particular, $\nu_{11,t} + 2\nu_{12,t} + \nu_{22,t} > 0$. This implies that the function $\nu_{1,t}(\zeta, \zeta) + \nu_{2,t}(\zeta, \zeta)$ is monotonically increasing in ζ and therefore invertible.

[29] The resource costs of a marginal deposit equal $\nu_t(\zeta_{t+1}, \zeta_{t+1}) - \zeta_{t+1}\,(\nu_{1,t}(\zeta_{t+1}, \zeta_{t+1}) + \nu_{2,t}(\zeta_{t+1}, \zeta_{t+1}))$.

[30] A similar qualification applies with respect to operating costs. If μ_t or ν_t depended on money or deposits the planner might rely on both means of payment.

balances were given by $z_{t+1} \equiv \Lambda_t(m_{t+1}) + n_{t+1}$ with $\Lambda_t(m_{t+1}) \equiv \alpha \ln(m_{t+1} + 1)$, the first-order conditions would hold as before except that λ_t would be replaced by $\alpha/(m_{t+1} + 1)$. For $\alpha > \mu_t/(\nu_t(\zeta^\star_{t+1}, \zeta^\star_{t+1}) + \zeta^\star_{t+1}\rho_t)$ but not "too large" the planner would rely on both money and deposits (and reserves).

4.2 Ramsey Policy

In contrast to the social planner, the Ramsey government controls the allocation only indirectly, by way of choosing policy instruments that support an equilibrium. We show next that the Ramsey government nevertheless implements the first best: The optimal policy $\mathcal{P}^\star$ supports an equilibrium with the social planner allocation.

To see this, note first that any equilibrium satisfies the social planner's optimality conditions for capital, consumption, and leisure as well as the resource constraint. The Ramsey policy therefore implements the first best if the first-best quantities of money, deposits, and reserves correspond to the demand and supply of these means of payment under the optimal policy. We can therefore focus on the conditions characterizing equilibrium in the payment system.

Suppose first that money is the more efficient means of payment, $\mu_t/\lambda_t \leq \nu_t(\zeta^\star_{t+1}, \zeta^\star_{t+1}) + \zeta^\star_{t+1}\rho_t$, such that the social planner does not use deposits. The Ramsey government can replicate the first-best outcome in this case by issuing the efficient amount of money and charging the liquidity premium

$$\chi^{m\star}_{t+1} \equiv \mu_t, \tag{RA-1}$$

pricing banks out of the market. Equation (RA-1), which reduces to the traditional Friedman rule when $\mu_t = 0$, follows directly from equations (4) and (SP). To see that banks are priced out of the market, recall from equation (5) that households only hold deposits if $\chi^n_{t+1} \leq \chi^m_{t+1}/\lambda_t$. With $\chi^m_{t+1} = \mu_t$ this would require that banks pay such a high interest rate that they incur losses if the government sets subsidies to zero and increases the liquidity premium on reserves sufficiently.[31]

Suppose next that deposits are more efficient such that the social planner does not issue money. Consider first a relaxed Ramsey program without the bank's optimality condition for deposits, inequality (8). In this relaxed program the Ramsey government can replicate the first best by issuing no money but the efficient amount of deposits, charging the liquidity premium on deposits

$$\chi^{n\star}_{t+1} \equiv \nu_t(\zeta^\star_{t+1}, \zeta^\star_{t+1}) + \zeta^\star_{t+1}\rho_t, \tag{RA-2}$$

and setting the liquidity premium on reserves to

$$\chi^{r\star}_{t+1} \equiv -\nu_{1,t}(\zeta^\star_{t+1}, \zeta^\star_{t+1}). \tag{RA-3}$$

Equation (RA-2) again follows directly from equations (4) and (SP). The liquidity premium on reserves in (RA-3) follows from equation (9) and the planner's optimal choice

[31] In the model variant without a reserves layer where ν_t is exogenous the same argument applies: Without government subsidies, paying a "competitive" interest rate on deposits would generate losses for banks.

of reserves: The premium must induce banks to select the first-best reserves-to-deposits ratio even when there are external effects that the banks do not internalize. We conclude that the Ramsey allocation in the relaxed program coincides with the social planner allocation.[32]

Consider next the full Ramsey program including equation (8) or (8a). The key question is how restrictive this equilibrium condition is: Can banks be induced to charge the liquidity premium on deposits given in (RA-2)—and thus issue the first-best quantity of deposits—when the central bank charges the liquidity premium on reserves given in (RA-3)? They can, when the subsidy θ_t in equation (8a) takes the value $\theta_t^\star$, which renders the equation evaluated at $\chi_{t+1}^{n\star}$ and $\chi_{t+1}^{r\star}$ consistent with the first-best deposit quantity.[33]

From equation (8a) this subsidy satisfies

$$\theta_t^\star \equiv \frac{1}{\eta_{n,t+1}} \frac{R_{t+1}^{n\star}}{R_{t+1}^{f\star}} - \chi_{t+1}^{n\star} + \tilde{\nu}_t(-\chi_{t+1}^{r\star}),$$

where $R_{t+1}^{f\star}$ is pinned down by the first-best allocation and $R_{t+1}^{n\star} \equiv R_{t+1}^{f\star}(1 - \chi_{t+1}^{n\star})$. Using (RA-2) and (RA-3) this can be expressed as

$$\theta_t^\star = \frac{1}{\eta_{n,t+1}} \frac{R_{t+1}^{n\star}}{R_{t+1}^{f\star}} + \zeta_{t+1}^\star \nu_{2,t}(\zeta_{t+1}^\star, \zeta_{t+1}^\star). \tag{RA-4}$$

The optimal subsidy has two components which reflect the two frictions in the banking sector, namely market power and externalities. The effect of market power on the subsidy is positive; the higher the elasticity of deposit funding, the smaller the subsidy. The effect of the externality, in contrast, is negative (recall that $\nu_{2,t} < 0$). Intuitively, with external benefits from reserves the Ramsey government subsidizes reserves by issuing them at a reduced liquidity premium, $\chi_{t+1}^{r\star}$, and it sterilizes the effect of this intervention on the deposit margin by lowering the subsidy on deposits.

Summarizing:

Proposition 3. The Ramsey policy implements the first best independently of whether the social planner relies on money or deposits. In the former case the Ramsey policy satisfies (RA-1). In the latter case it satisfies (RA-3) and (RA-4) and implements the deposit spread (RA-2); deposits may be taxed or subsidized.

Of course, proposition 3 applies accordingly when exogenous restrictions rule out either money or deposits such that a priori only a single retail means of payment is available. More importantly, proposition 3 generalizes to settings with nonlinear aggregation of liquidity (endogenous λ_t) because the policy instruments still suffice to control the amount of liquidity, its composition between deposits and money, and banks' willingness to issue deposits in this case.

[32]In the model variant without a reserves layer the right-hand side of equation (RA-2) is replaced by ν_t. The Ramsey government need not target the reserves-to-deposits ratio in this case but it also has one less instrument, χ_{t+1}^r, at its disposal. As a consequence, the first best can be implemented.

[33]In general the optimal policy may not uniquely implement the first best. When we impose the functional form assumptions this is not an issue, see appendix B.

When optimality is consistent with the joint circulation of money and deposits, i.e., when the condition $\mu_t/\lambda_t = \nu_t(\zeta^\star_{t+1}, \zeta^\star_{t+1}) + \zeta^\star_{t+1}\rho_t$ happens to be satisfied, the Ramsey government may issue both money and reserves. An important question is whether the interest rate on money and reserves should be equalized in this case, i.e., whether CBDC should be remunerated at the interest rate on reserves. Proposition 3 indicates that this would not be optimal: Liquidity premia on money and reserves should equal μ_t and $-\nu_{1,t}(\zeta^\star_{t+1}, \zeta^\star_{t+1})$, respectively, reflecting the social costs of money- and reserves-based payments as well as externalities. These social costs and externalities typically differ between money and reserves—even if the relative liquidity benefits of money and deposits coincide with their relative resource costs.

4.3 Policy Rules

Under the Ramsey policy liquidity premia respond to developments in the payment sector: $\chi^{m\star}_{t+1}$ varies with μ_t while $\chi^{n\star}_{t+1}$ and $\chi^{r\star}_{t+1}$ depend on ρ_t as well as the cost function $\nu_t(\cdot)$. The subsidy $\theta^\star_t$ also reflects shocks in the payment sector and in addition, preferences and firm technology, through their effects on the allocation $(R^{f\star}_{t+1})$.

Focusing on the case with deposits and reserves and imposing assumption 1 the optimal spread on reserves satisfies (see appendix B for all derivations)

$$\chi^{r\star}_{t+1} = \rho_t \frac{\phi_{1,t}}{\phi_{1,t} + \phi_{2,t}} \leq \rho_t.$$

The optimal spread is proportional to the operating costs of the reserves layer, ρ_t. When reserves holdings generate externalities ($\phi_{2,t} > 0$) the optimal policy subsidizes reserves holdings, $\chi^{r\star}_{t+1} < \rho_t$, in order to induce banks to internalize the external benefits. Note that the expression for $\chi^{r\star}_{t+1}$ can be expressed as a simple interest rate rule for reserves,

$$R^{r\star}_{t+1} = R^f_{t+1} \cdot \left(1 - \rho_t \frac{\phi_{1,t}}{\phi_{1,t} + \phi_{2,t}}\right).$$

According to this rule $R^{r\star}_{t+1}$ moves one-to-one with R^f_{t+1}, and with small interest rates ($R^f_{t+1} \approx 1$) the interest rate on reserves should equal the risk-free rate minus $\rho_t\phi_{1,t}/(\phi_{1,t} + \phi_{2,t})$.

The optimal reserves-to-deposits ratio is given by the expression reported in subsection 3.3 evaluated at $\chi^{r\star}_{t+1}$,

$$\zeta^\star_{t+1} = \left(\frac{(\phi_{1,t} + \phi_{2,t})(\varphi - 1)}{\rho_t}\right)^{\frac{1}{\varphi}}.$$

The optimal ratio increases in $\phi_{1,t}$ and $\phi_{2,t}$; that is, when reserves affect the costs of deposit-based payments more strongly, either bank internally or externally, then the reserves-to-deposits ratio should be higher. The optimal ratio falls with the central bank's resource costs of reserves operations, ρ_t.

The optimal subsidy,

$$\theta_t^\star = \left(\frac{\rho_t}{(\phi_{1,t}+\phi_{2,t})(\varphi-1)}\right)^{1-\frac{1}{\varphi}} \left(\psi(\phi_{1,t}+\phi_{2,t})\varphi - \phi_{2,t}(\varphi-1)\right),$$

may have a positive or negative sign (cf. proposition 3). We have the following result:

Proposition 4. Under assumption 1, the optimal subsidy in an economy with deposits is strictly negative when

$$\frac{\phi_{2,t}}{\phi_{1,t}+\phi_{2,t}}\frac{\varphi-1}{\varphi} > \psi,$$

and strictly positive when the reverse inequality holds.

The left-hand side of the inequality represents the importance of externalities as well as the strength with which reserves reduce operating costs; the right-hand side represents the extent of bank market power. When reserves do not generate externalities ($\phi_{2,t} = 0$) the left-hand side equals zero and the optimal subsidy is nonnegative. Reserves are priced at their marginal resource costs in this case and there is therefore no need for the deposit subsidy to sterilize a reserves subsidy; the deposit subsidy only must correct the distortion due to monopsony power, which it achieves by subsidizing deposits. When monopsony power is absent, in contrast, because consumption and effective real balances are perfect substitutes ($\psi \to 0$) then the right-hand side equals zero and the subsidy is nonpositive. Absent externalities and monopsony power the optimal subsidy equals zero.

Under the optimal policy the liquidity premium on deposits is given by the expression reported in subsection 3.3 evaluated at $\chi_{t+1}^{r\star}$ and $\theta_t^\star$,

$$\chi_{t+1}^{n\star} = \left(\frac{\rho_t}{(\phi_{1,t}+\phi_{2,t})(\varphi-1)}\right)^{1-\frac{1}{\varphi}} (\phi_{1,t}+\phi_{2,t})\varphi.$$

The premium reflects all social cost components of deposit-based payments, namely those in the interbank and the retail payment system. When reserves-based payments do not generate resource costs ($\rho_t = 0$) then the optimal liquidity premium on deposits equals zero because satiation with reserves eliminates bank operating costs (unless $\varphi = 1$, see below). When reserves operations require resources ($\rho_t > 0$), in contrast, then the liquidity premium accounts for this cost as well as for the cost structure in the retail payment system (represented by $\phi_{1,t}$, $\phi_{2,t}$, and φ). Greater importance of reserves for the retail payment system, either through bank internal or external channels, increases the deposit liquidity premium because it raises reserves holdings and the costs associated with them.

The parameter φ determines the elasticity of ν_t with respect to the reserves-to-deposits ratio. For $\varphi \to 1$ this elasticity equals zero and the model variant without a reserves layer results. In this variant the optimal reserves-to-deposits ratio equals zero; the optimal subsidy satisfies $\theta_t^\star = \psi(\phi_{1,t}+\phi_{2,t})$; and the liquidity premium on deposits equals $\chi_{t+1}^{n\star} = \phi_{1,t}+\phi_{2,t}$.

4.4 Alternative Policy Instruments and Second-Best Policies

In addition to the spread on reserves—a common instrument in actual monetary policy operations—the Ramsey policy relies on a less common deposit subsidy, θ_t, or similar instruments such as regulatory constraints to affect bank behavior. One may ask whether standard monetary policy instruments such as money supply or spread targets could replace θ_t without restricting policy makers. As we show next this is only the case in specific circumstances. Accordingly, we also analyze second-best policies.

To fix ideas suppose that the social planner relies on deposits because money is an inefficient means of payment and that reserves generate externalities. The Ramsey policy then sets the spread on reserves to $-\nu_{1,t}(\zeta^{\star}_{t+1}, \zeta^{\star}_{t+1})$ and the deposit subsidy to $\theta^{\star}_t$ (see equations (RA-3) and (RA-4)). If the subsidy instrument is not admissible an alternative second instrument is needed to induce banks to charge the socially optimal deposit spread, $\chi^{n\star}_{t+1}$. This instrument should not waste resources.

Any policy that involves the issuance of money would fail to achieve the latter goal because it would require households to hold money rather than deposits and thus waste resources. In contrast, an appropriately set target for the spread on money, χ^m_{t+1}, possibly can achieve that goal because it does not entail that households actually hold money in equilibrium; in some circumstances the spread target can also induce the socially desirable bank behavior.

To see this, note from equation (8a) that a bank whose deposits are not subsidized or taxed sets the spread

$$\chi^n_{t+1} = \tilde{\nu}_t(-\chi^{r\star}_{t+1}) + \frac{1}{\eta_{n,t+1}} \frac{R^n_{t+1}}{R^f_{t+1}} \equiv \frac{\tilde{\nu}_t(-\chi^{r\star}_{t+1})}{1-\psi},$$

where the identity defines the markup $(1-\psi)^{-1} - 1$. (Under assumption 1 ψ denotes a preference parameter and the equilibrium markup equals $(1-\psi)^{-1} - 1$. Here we use the same notation for the markup without imposing assumption 1 or implying that ψ is a (preference) parameter.) When the elasticity $\eta_{n,t+1}$ approaches infinity the bank becomes a price taker and the markup disappears, $\psi \to 0$. We refer to the profit maximizing spread χ^n_{t+1} in the foregoing equation as the "monopsony spread."

A monetary policy that sets the spread on money higher than λ_t times the monopsony spread is irrelevant because it renders money unattractive for households even when banks exploit their market power to the fullest. When the central bank sets the spread on money lower than λ_t times the monopsony spread, however, then monetary policy is relevant. Banks have two options in this case. Either they raise the interest rate on deposits to the competitive level given in equation (5) (plus epsilon), thereby pricing the central bank out of the market, or they do not raise the rate and are priced out of the market themselves. Banks will choose the former option as long as this generates nonnegative profits given $\chi^{r\star}_{t+1}$, which requires (from conditions (6) and (7))

$$\chi^n_{t+1} \geq \nu_t(\zeta^{\star}_{t+1}, \zeta^{\star}_{t+1}) + \zeta^{\star}_{t+1}\chi^{r\star}_{t+1} = \tilde{\nu}_t(-\chi^{r\star}_{t+1}).$$

In conclusion the money-spread target $\chi^m_{t+1} = \lambda_t \chi^{n\star}_{t+1}$ implements the first-best deposit spread, $\chi^{n\star}_{t+1}$, if and only if

$$\tilde{\nu}_t(-\chi^{r\star}_{t+1}) \leq \chi^{n\star}_{t+1} < \frac{\tilde{\nu}_t(-\chi^{r\star}_{t+1})}{1-\psi}.$$

In this case banks raise the deposit rate to stay competitive and no money is held in equilibrium. When $\tilde{\nu}_t(-\chi^{r\star}_{t+1})/(1-\psi) \leq \chi^{n\star}_{t+1}$, in contrast, then monetary policy is powerless because it cannot push the spread on deposits up relative to the monopsony spread.[34] Summarizing:

Proposition 5. When the social planner relies on deposits, targeting χ^m_{t+1} (in addition to χ^r_{t+1}) can only substitute for the first-best deposit subsidy, $\theta^\star_t$, when that subsidy is positive.

Figure 1 illustrates this result under assumption 1 and for parameter values that render the two-tiered system with unconstrained $\chi^{r\star}_{t+1}$ and $\theta^\star_t$ preferable to a money-based system.[35] The solid line in the figure represents the deposit spread under the unconstrained Ramsey policy, $\chi^{n\star}_{t+1} = 0.04$, and the dotted line represents $\theta^\star_t$ as a function of ψ; a higher ψ, which is associated with more bank market power, raises the optimal subsidy. The two dashed lines represent the bounds $\tilde{\nu}_t(-\chi^{r\star}_{t+1})$ and $\tilde{\nu}_t(-\chi^{r\star}_{t+1})/(1-\psi)$, which are relevant when the θ_t instrument is not admissible; specifically, the upper dashed line represents the monopsony spread.

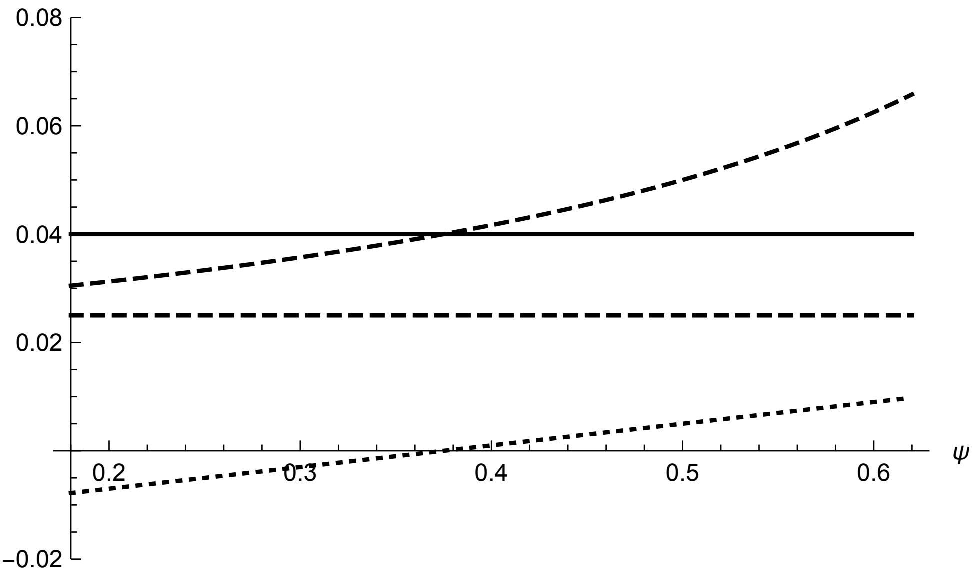

Figure 1: Targeting χ^m_{t+1} cannot always replicate $\theta^\star_t$.

[34] In Andolfatto (2021) this problem is not present since there are no reserves with positive externalities. The case of $\chi^{n\star}_{t+1} < \tilde{\nu}_t(-\chi^{r\star}_{t+1})$ cannot arise.

[35] We posit $\sigma = 0.5$, $\vartheta = 0.1$, and in the baseline $\psi = 0.2$. Function ν_t is characterized by $\phi_{1,t} = 0.01$, $\phi_{2,t} = 0.03$, and $\varphi = 2$. Finally, $\mu_t = 0.06$, $\rho_t = 0.01$, and $\lambda_t = 1$. These values imply $\nu_t(\zeta^\star_{t+1}, \zeta^\star_{t+1}) + \zeta^\star_{t+1}\rho_t = 0.04 < \mu_t$, $\chi^{r\star}_{t+1} = 0.0025$, $\chi^{n\star}_{t+1} = 0.04$, and $\theta^\star_t = -0.007$.

Note that this upper bound crosses the $\chi^{n\star}_{t+1}$ line when $\theta^\star_t = 0$, consistent with proposition 5. For high values of ψ the monopsony spread of banks is high, the optimal unconstrained subsidy is positive, and if the subsidy instrument is not admissible the central bank can nevertheless implement the first best by targeting $\chi^m_{t+1} = \lambda_t \chi^{n\star}_{t+1}$. When ψ is smaller than 0.38, in contrast, the monopsony spread is lower than $\chi^{n\star}_{t+1}$, the optimal unconstrained subsidy is negative, and if the subsidy instrument is not admissible targeting χ^m_{t+1} does not constitute a substitute; absent the θ_t instrument the first best is not implementable in this latter case.

What is the second-best policy when $\theta^\star_t < 0$ but the θ_t instrument is not admissible? We consider three candidates in addition to a policy that solely relies on money and which is unattractive given our assumption of high resource costs of money-based payments, μ_t. First, an "n policy" that solely relies on χ^r_{t+1} subject to $\theta_t = m_{t+1} = 0$. Second, a "ξ policy" that targets χ^r_{t+1} as well as the composition of real balances, $\xi_{t+1} \equiv n_{t+1}/z_{t+1}$, letting banks issue their preferred quantity of deposits with deposit and money spreads adjusting subject to condition (5). And third, an "m policy" that targets χ^r_{t+1} as well as the quantity of money, again letting banks issue their preferred quantity of deposits and letting the deposit and money spreads adjust. Appendix A contains a detailed characterization of the different equilibria.

The constrained optimal n policy attains the first-best welfare level when $\psi = 0.38$, setting the spread on reserves to $\chi^{r\star}_{t+1}$; no subsidy is needed in this case. For any other value of ψ the constrained optimal n policy does worse than the unconstrained policy. When ψ is smaller (larger) than 0.38 then the constrained optimal spread on reserves is higher (lower) than $\chi^{r\star}_{t+1}$ because χ^r_{t+1} partly compensates for the missing θ_t instrument.

The constrained optimal ξ policy strictly dominates the constrained optimal n policy for any value of ψ other than 0.38. This follows from the fact that a ξ_{t+1} target does not affect the markup that banks charge (see appendix A). As a consequence a ξ_{t+1} target slightly below unity allows the central bank to expand liquidity provision conditional on the monopsony spread, a possibility that n policies do not afford. This lowers the equilibrium spread χ^n_{t+1} and raises welfare in spite of the fact that it requires issuing a small quantity of (resource intensive) money.

Importantly, the constrained optimal ξ policy also dominates the constrained optimal m policy because targeting ξ_{t+1} is not equivalent to, and in fact cannot be replicated by, targeting a specific m_{t+1}. The reason for this is that an m_{t+1} target affects the markup that banks charge while a ξ_{t+1} target does not (see appendix A). We summarize these findings in the following proposition:

Proposition 6. Under assumption 1 the unconstrained Ramsey policy generically strictly dominates the constrained optimal ξ policy which in turn generically strictly dominates the constrained optimal n policy. Generically, an m policy cannot replicate a ξ policy.

We close the section with a remark on a recurrent theme in our first- and second-best analyses: Reserves and money, the two central bank liabilities, should pay different interest rates. This holds true independently of whether money is used as a means of payment or as a device to discipline banks when the θ_t instrument is not admissible. When both means of payment circulate in a first-best equilibrium their spreads should reflect operating costs and externalities, which generally differ (see the discussion after proposition 3). When money only serves to discipline banks then its spread should reflect the central bank's target for the deposit rate which is only indirectly linked to the social costs and benefits of reserves. And when both money and reserves circulate in a second-best equilibrium then the spreads must additionally account for the tradeoffs that the missing θ_t instrument creates. Summarizing:

Proposition 7. When the optimal policy relies on both central bank liabilities, reserves and money, then the spreads on both liabilities generically optimally differ.

5 Equivalence

In this section we establish an equivalence class of monetary systems with different compositions of real balances. We use the equivalence result to characterize scenarios in which the introduction of CBDC is "irrelevant" from a macroeconomic perspective. We also quantify implicit subsidies to banks in a two-tiered monetary system.

5.1 Equivalence of Monetary Systems

The equivalence result, which extends findings in Brunnermeier and Niepelt (2019), establishes that the composition of real balances is "irrelevant" when the private and the public sector are equally efficient in operating payment systems.[36] In this case the central bank can ensure that a portfolio shift from deposits into money leaves equilibrium consumption, capital accumulation, and the price system unchanged. Private and public means of payment thus are substitutes in general equilibrium as long as the central bank intervenes appropriately. If it does, portfolio shifts out of deposits into money (CBDC) do not undermine bank intermediation.

Importantly, and as will become clear, equivalence follows under much more general conditions than those laid out in section 2. The type of monetary friction (money in the utility function or otherwise), structure of preferences, or competitive environment in the banking sector are not important for the result. What is critical is that deposits as a source of bank funding can be substituted by a central bank loan.[37]

The following condition stipulates that the resource costs per unit of effective real balances are the same for money and deposits:

Condition 1. $\mu_t/\lambda_t = \nu_t(\zeta_{t+1}, \zeta_{t+1}) + \zeta_{t+1}\rho_t$ where ζ_{t+1} denotes the equilibrium reserves-to-deposits ratio.

[36] Brunnermeier and Niepelt (2019) consider the case without reserves nor resource costs of providing liquidity. Substitutability between money and deposits guarantees "liquidity neutrality" in their terminology. See also Niepelt (2018; 2020).

[37] Of course, our specific assumptions about how reserves holdings enter ν_t and how ν_t and θ_t affect bank operating costs matter for the exact formula of the equivalent central bank loan rate derived below. In general this rate reflects all factors that determine banks' costs of sourcing funding through deposits as opposed to from the central bank.

Consider a reduction in household deposit holdings by λ_t units at date t, complemented by a unit increase in money holdings also at date t. We claim that this swap need not change the allocation. Specifically, we have the following result, which is formally stated and proved in appendix C:

Proposition 8. Suppose condition 1 holds. Consider a policy and equilibrium with deposits, reserves, and no central bank loan. There exists another policy and equilibrium with fewer deposits and reserves, more money, a central bank loan, a different ownership structure of capital, and otherwise the same allocation and price system.[38] The central bank loan carries the interest rate

$$R^l_{t+1} = \frac{R^n_{t+1} + (\nu_t(\zeta_{t+1}, \zeta_{t+1}) - \theta_t)R^f_{t+1} - \zeta_{t+1}R^r_{t+1}}{1 - \zeta_{t+1}}.$$

The logic underlying the proposition is simple: When households hold fewer deposits but more money then the central bank can pass this new funding back to banks as loans. With an appropriate loan supply schedule the central bank perfectly replicates the choice sets that banks had before the intervention. The central bank thus can guarantee that banks' lending choices remain unchanged and more generally, that the initial equilibrium prices and allocation together with the modified portfolios constitute an equilibrium. Establishing these results requires an analysis of the budget and choice sets of banks, households, firms, and the government. Appendix C contains this analysis which, as emphasized before, applies under quite general conditions.

According to proposition 8 the loan interest rate, R^l_{t+1}, which supports the equivalent portfolio positions, reflects the (old and new) equilibrium reserves-to-deposits ratio, ζ_{t+1}. There are three reasons for this dependence; they relate to the quantity of bank funding and its price. First, one dollar of deposit funding only results in $1 - \zeta_{t+1}$ dollars to be invested in capital because the remaining ζ_{t+1} dollars are invested in reserves. With a central bank loan, in contrast, the bank borrows $1 - \zeta_{t+1}$ dollars and invests that amount fully in capital. This difference explains the $1 - \zeta_{t+1}$ term in the denominator of the expression for the loan rate. Second, the cost of deposit sourced funding depends on the share of deposits that is invested in reserves. This explains the term $-\zeta_{t+1}R^r_{t+1}$ in the numerator. And third, the reserves-to-deposits ratio affects the operating costs, $\nu_t(\zeta_{t+1}, \zeta_{t+1})$.

A couple of further remarks are in order. First, in the model variant without a reserves layer where ν_t is exogenous the proposition applies subject to the obvious modifications; in particular, condition 1 reduces to $\mu_t/\lambda_t = \nu_t$.

Second, a special case in which the equivalence condition is trivially satisfied is when the marginal resource costs of money-, deposit-, and reserves-based payments all equal zero, $\mu_t = \nu_t = \rho_t = 0$, which might be plausible at least locally. The three payment networks might have substantial fixed costs, however. The equivalence result still holds in this case if the initial equilibrium features all means of payment such that the fixed costs are borne both in the initial and the new equilibrium.

[38] The new policy may also include state contingent taxes whose market value equals zero.

Third, since the first-best reserves-to-deposits ratio, $\zeta^{\star}_{t+1}$, may differ from the ratio in equilibrium the condition for equivalence need not coincide with the condition under which the social planner is indifferent between the two means of payment, $\mu_t/\lambda_t = \nu_t(\zeta^{\star}_{t+1}, \zeta^{\star}_{t+1}) + \zeta^{\star}_{t+1}\rho_t$ (see condition (SP)).[39] That is, it is conceivable that the introduction of money is "irrelevant" as far as equilibrium outcomes are concerned although the social planner or the Ramsey government strictly prefers money over deposits or vice versa.

Fourth, the equivalence proposition relies on the substitutability of deposits and other sources of bank funding. When there are complementarities between bank assets and deposits, but not between bank assets and a central bank loan, then the refinancing operation by the central bank may change the choice sets of banks and equivalence may not be guaranteed.[40] Deposits may also be "special" in the sense that they induce more or less monitoring. But if banks are mainly monitored by the central bank and bank supervisors then the substitution of a loan for deposits is unlikely to make a significant difference.

Fifth, the central bank loan in the equivalent arrangement is extended against no collateral. This stands in contrast with most actual central bank lending operations, reflecting concerns about central bank net worth and independence from the treasury—an issue we take up in section 6.[41] At the same time it raises interesting questions about implicit lender-of-last-resort guarantees: When deposits circulate as means of payment they are not secured even if households perceive their liquidity to derive from central bank guarantees (see the discussion in the next subsection). But if the implicit guarantees were made explicit, by substituting money for deposits, collateral would have to be posted under most real-world operating procedures. The equivalence result highlights this inconsistency.

Sixth, heterogeneity across households or banks including household specific λ_t parameters or bank specific ν_t functions may or may not undermine the equivalence result. When condition 1 is satisfied for a specific group but not for others the substitution of money for deposits accompanied by the appropriate central bank policy is neutral as long as the swap can be targeted to the specific group. Even if this is not possible the substitution may still be neutral. Consider for example a monetary union composed of two countries of which one satisfies condition 1 while the other does not, due to a high regional λ_t value. The substitution of money for deposits would not have differential effects on banks in the two countries but it would lead to different adjustments in household portfolios. Equivalence therefore would require country specific taxes as the proof of proposition 8 makes clear.

Seventh, the equivalence result does not hinge on specific assumptions about the market structure in the banking sector. While the model in section 2 assumed segmented markets with monopsonist banks this assumption plays no role in condition 1, the logic of the equivalence result, or the formula for R^l_{t+1}. The market structure only affects the

[39] However, in the model variant without a reserves layer these two conditions coincide.

[40] For examples of such complementarities see Kashyap et al. (2002).

[41] In the model the central bank need not guard against losses on its assets because the government sector is consolidated, has access to nondistorting taxes, and households are homogeneous. Williamson (2019) emphasizes that the Federal Reserve System before the 1920s and the European Central Bank have been set up as institutions that lend to banks—against collateral.

properties of the loan funding schedule that preserves the choice sets of banks.

Eighth, if deposits were fully replaced by money, the volume of central bank loans under the equivalent policy would be a decreasing function of the reserves-to-deposits ratio in the initial equilibrium: With high initial ζ_{t+1} values the central bank would have to extend a small loan and vice versa. Currently, central banks in many monetary areas have substantially longer balance sheets than before the financial crisis, reflecting high ζ_{t+1} values in the national banking systems. An equivalent substitution of CBDC for deposits in these areas thus would require smaller central bank loans than fifteen years ago. Of course, this is not to suggest that such a complete substitution may be realistic or even feasible. After all, defining "money" in legally cogent terms is most challenging and curtailing its private creation even harder.

Ninth, we have assumed that deposits are risk-free, implicitly ruling out bank runs with incomplete deposit insurance. Relaxing this assumption would not change the results. If deposits were subject to bank runs but money were risk-free equivalence would require state-contingent transfers (with a market value of zero) to counterbalance the modified return characteristics of household portfolios.[42] More generally, the introduction of money into a two-tiered system would not undermine financial stability in the sense of putting bank funding at risk because any run from deposits into money would automatically—as a matter of accounting—trigger a central bank loan. If the central bank rejected incoming payments from depositors who seek to swap their deposits for money then a run could not occur in the first place.

Finally, the equivalence result allows to draw some normative conclusions. We know that if condition 1 is met the central bank can secure the status quo by acting appropriately. This implies that welfare under the status quo is a lower bound to what the substitution of money for deposits could achieve if the central bank chose a "better" policy than the equivalent one.[43]

5.2 Bank Profits and Lender-of-Last-Resort Support

Banks earn profits by providing liquidity. But do banks themselves create this liquidity? Or does the central bank create it and banks benefit from beliefs held by the public (and upheld by institutions such as deposit insurance and lender-of-last-resort facilities) that deposits are implicitly or explicitly guaranteed by the central bank? If one accepts the latter view then it is natural to ask how valuable the guarantees are, that is, how highly banks are subsidized by letting them issue means of payment. If one accepts the former view then it is interesting to ask how much value banks create. The equivalent loan rate encountered in proposition 8, R^l_{t+1}, provides answers to these questions.

Recall that R^l_{t+1}, which can be expressed as

$$R^r_{t+1} + \frac{R^n_{t+1} - R^r_{t+1}}{1 - \zeta_{t+1}} + \frac{(\nu_t(\zeta_{t+1}, \zeta_{t+1}) - \theta_t) R^f_{t+1}}{1 - \zeta_{t+1}}, \tag{16}$$

[42] See Brunnermeier and Niepelt (2019) for a discussion.

[43] Brunnermeier and Niepelt (2019) argue that rather than replicating households' deposit holding strategies the central bank as a "large" lender could for example do better by internalizing bank run externalities.

represents the costs for a bank in the two-tiered system to finance one unit of capital investment with deposits. If the bank financed investment with an "illiquid" security that does not serve as a means of payment then the costs would be represented by R^f_{t+1} rather than R^l_{t+1}. Expressed as a share of GDP banks in the two-tiered monetary system thus benefit from a funding cost reduction of

$$\text{fcr}_t \equiv \frac{R^f_{t+1} - R^l_{t+1}}{R^f_{t+1}} \frac{n_{t+1}(1-\zeta_{t+1})}{\text{GDP}_t}.$$

In line with the different perspectives on bank profits described above this funding cost reduction can be interpreted as value added by banks or as an implicit subsidy. Against the background of the equivalence result, which relates the equilibrium with deposits to an equilibrium with money and central bank funding of commercial banks, the latter interpretation seems natural. But whatever interpretation one favors, fcr_t clearly represents a measure of the political risks associated with the introduction of CBDC. After all, this introduction would only be neutral if the central bank refinanced banks at the equivalent loan rate, but such refinancing at nonmarket terms could be very difficult or impossible on political grounds. We return to political economy considerations in the subsequent section. Before that, we quantify R^l_{t+1} and fcr_t.

Figure 2 displays quarterly U.S. time series data for the (inflation adjusted) gross reserves rate, R^r_{t+1}, the gross risk-free rate, R^f_{t+1}, and the gross deposit rate, R^n_{t+1}; the data spans the interval 1999q1–2021q1. To construct these series we use FRED data and Kurlat's (2019) estimates of the risk-free, "illiquid" interest rate as well as the deposit rate. Appendix D contains detailed information.

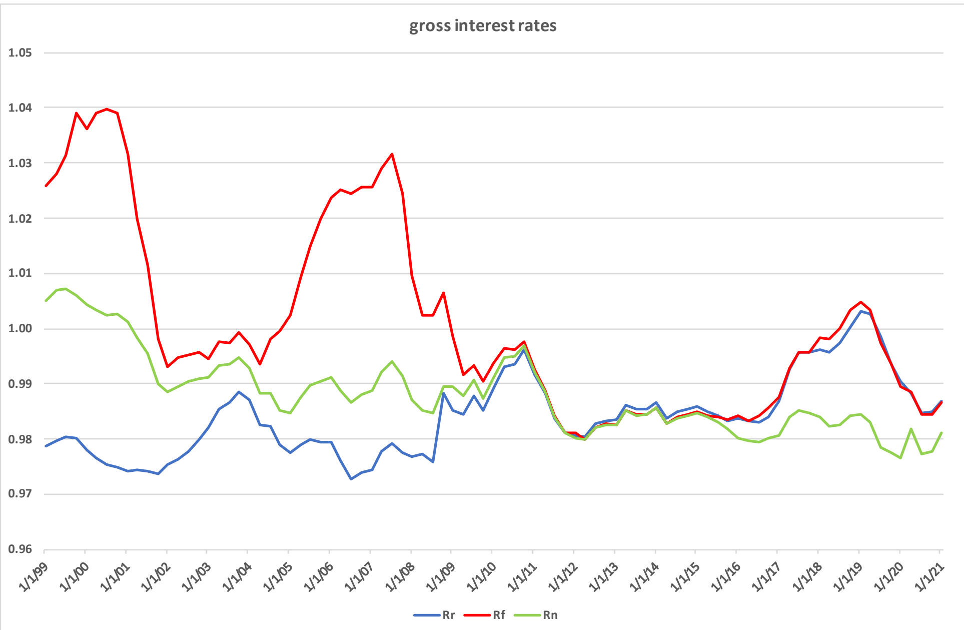

Figure 2: Gross interest rates.

The figure shows that the (inflation adjusted) interest rate on reserves fluctuates in a band between −3 and 0.5 percent while the risk-free rate varies between −2 and 4 percent. After 2010, the two rates nearly coincide. The deposit rate lies between the reserves rate and the risk-free rate before the financial crisis and below the two other rates at the end of the sample. In the first half of the 2010s, there are only tiny spreads between the three rates.

Figure 3 displays quarterly time series data for the reserves-to-deposits ratio, ζ_{t+1}, over the same period. We use FRED data as well as data constructed by Lucas and Nicolini (2015) for reserves and deposits. For the deposit series, we use two alternative measures. The first, indicated by [a], is the sum of checkable and savings deposits. The second, indicated by [b], is the sum of checkable deposits and money market deposit accounts as specified by Lucas and Nicolini (2015). Appendix D contains more information.

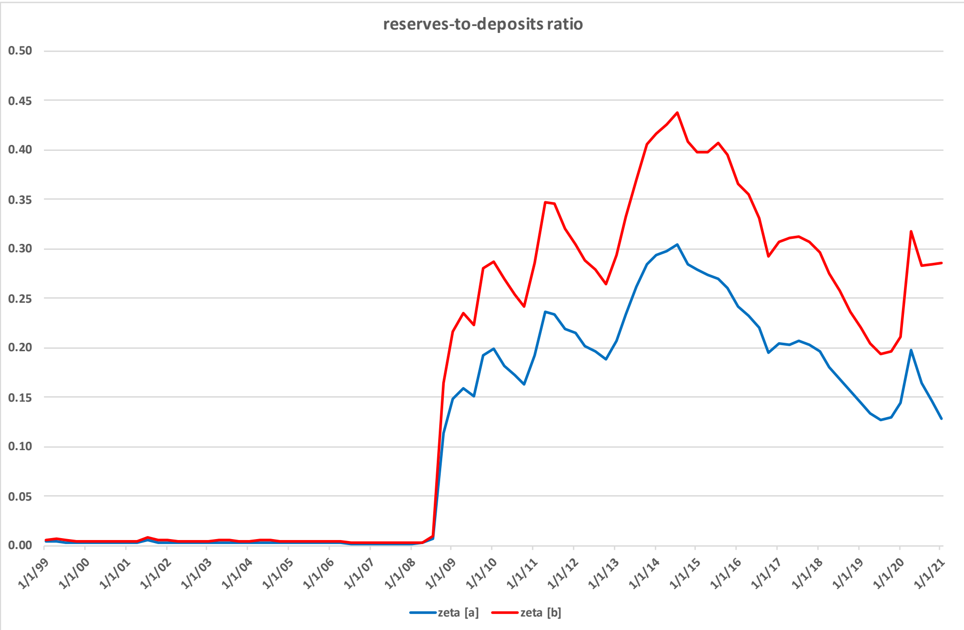

Figure 3: Two measures of the reserves-to-deposits ratio.

Irrespective of the exact measure, the reserves-to-deposits ratio strongly increases in mid 2008, from a very low level (at which it had been since the early 1980s). It reaches a maximum of 30 or 45 percent, depending on the measure, in mid 2014 and falls afterwards before increasing again at the beginning of the COVID-19 crisis.

Figures 4 and 5 display the implied loan rate, R^l_{t+1}, and funding cost reduction for banks, fcr_t. In each case we report the results for either measure of the reserves-to-deposits ratio; the differences are minor. In figure 4 we compute R^l_{t+1} under the assumption that $\nu_t(\zeta_{t+1}, \zeta_{t+1}) - \theta_t = 0.01$, in line with estimates of banks' operating costs for payments.[44] In figure 5 we infer $\nu_t(\zeta_{t+1}, \zeta_{t+1}) - \theta_t$ from the data under the assumption that banks behave according to the model.

[44]Lucas and Nicolini (2015, p. 57) assume that banks' costs of check processing equal 1 percent of GDP. Philippon (2015) estimates that the costs of financial intermediation equal 1.5 to 2 percent of intermediated assets. The ratio of deposits to GDP equals one third to one half.

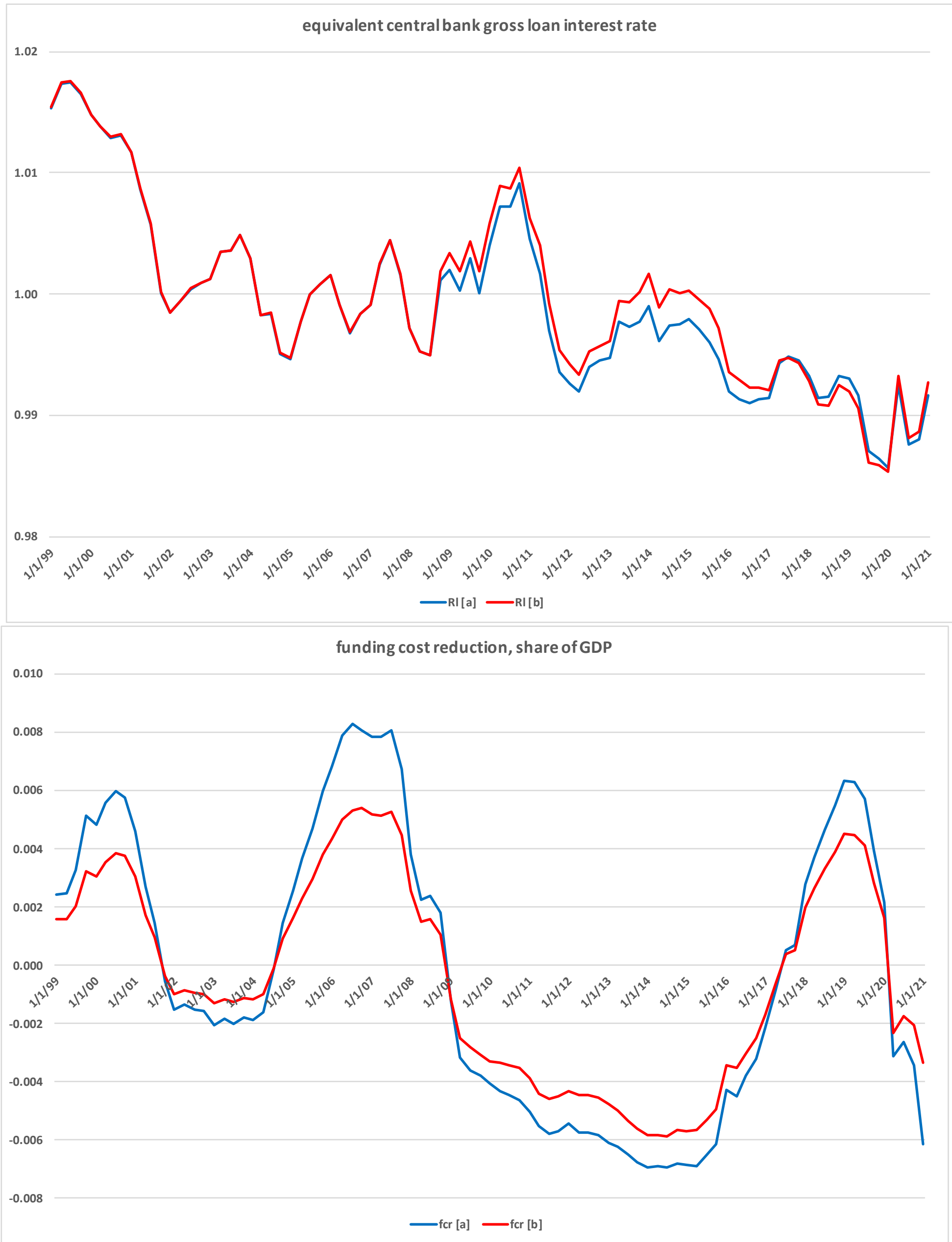

Figure 4: Two measures of the equivalent central bank gross loan interest rate (top) and banks' funding cost reduction (bottom) under the assumption that $\nu_t - \theta_t = 0.01$.

Consider first figure 4. The equivalent central bank loan interest rate displayed in the top panel falls from nearly 2 percent early in the sample to -1 percent at the end, with a temporary increase to 1 percent in late 2010. We can distinguish several phases:

- Prior to 2008 the equivalent loan rate roughly equals $R^n_{t+1} + 0.01$ as the reserves-to-deposits ratio is tiny and $R^f_{t+1} \approx 1$ (see equation (16)).
- In 2008 the reserves-to-deposits ratio strongly increases; one dollar of funding for

capital investment now requires substantially more than one dollar of deposits. Since the deposit rate exceeds the interest rate on reserves, the equivalent loan rate rises.

- Between 2009 and 2015, the equivalent loan rate follows the interest rate on reserves—due to the strong compression of interest rates—plus a term that reflects operating costs and the reserves-to-deposits ratio.

- Finally, after 2015 the reduction of the deposit rate relative to the reserves rate contributes negatively to the equivalent loan rate.

The bottom panel of figure 4 illustrates the implied funding cost reduction for banks. The time series fluctuates between -0.7 and 0.8 percent of GDP, reflecting several drivers: The long-term decline in $R^f_{t+1} - R^l_{t+1}$ and fluctuations around this trend; the U-shaped path of $1-\zeta_{t+1}$ after 2008; and an increasing deposits-to-GDP ratio, in particular towards the end of the sample. Again, we can distinguish several phases:

- In the beginning of the sample, money creation by banks reduces their funding costs by roughly 0.5 percent of GDP because the risk-free rate exceeds the costs of deposit funding. Equivalently, banks benefit from the equivalent central bank loan in the money-based system because the risk-free rate exceeds the equivalent loan rate.

- From 2002 to 2004 the risk-free rate is low and so is the spread between the risk-free rate and the equivalent loan rate. As a consequence the funding cost reduction is negative.

- In 2005 and 2006 the risk-free rate rises again, pushing the funding cost reduction to $0.5 - 0.8$ percent. In the following two years this effects weakens.

- From 2008 the three market interest rates converge and between 2011 and 2015 they practically coincide. The funding cost reduction therefore increasingly mimics $-(\nu_t - \theta_t)n_{t+1}/\mathrm{GDP}_t$ (see equation (16) and the formula for fcr_t).

- From 2016 to 2019 the fall in the equivalent loan rate and the rise in the risk-free rate contribute to an increase in the funding cost reduction.

- At the end of the sample the fall in the risk-free rate relative to the equivalent loan rate reduces the funding cost reduction and pushes it back into negative territory.

Turn next to figure 5. Rather than positing $\nu_t(\zeta_{t+1}, \zeta_{t+1}) - \theta_t = 0.01$ we now infer the operating costs from the data assuming that banks set the deposit rate according to the model. Specifically, imposing assumption 1 and using the bank's first-order condition $\nu_t(\zeta_{t+1}, \zeta_{t+1}) - \theta_t = (1-\psi)\chi^n_{t+1} - \zeta_{t+1}\chi^r_{t+1}$ (see appendix A), we find the simple expressions

$$\begin{aligned} R^l_{t+1} &= R^f_{t+1} - \frac{\psi}{1-\zeta_{t+1}}(R^f_{t+1} - R^n_{t+1}), \\ \mathrm{fcr}_t &= \psi \chi^n_{t+1} \frac{n_{t+1}}{\mathrm{GDP}_t}, \end{aligned}$$

according to which stronger market power (higher ψ) reduces the equivalent loan rate and increases the funding cost reduction. For our calculations we let $\psi = 0.5$.

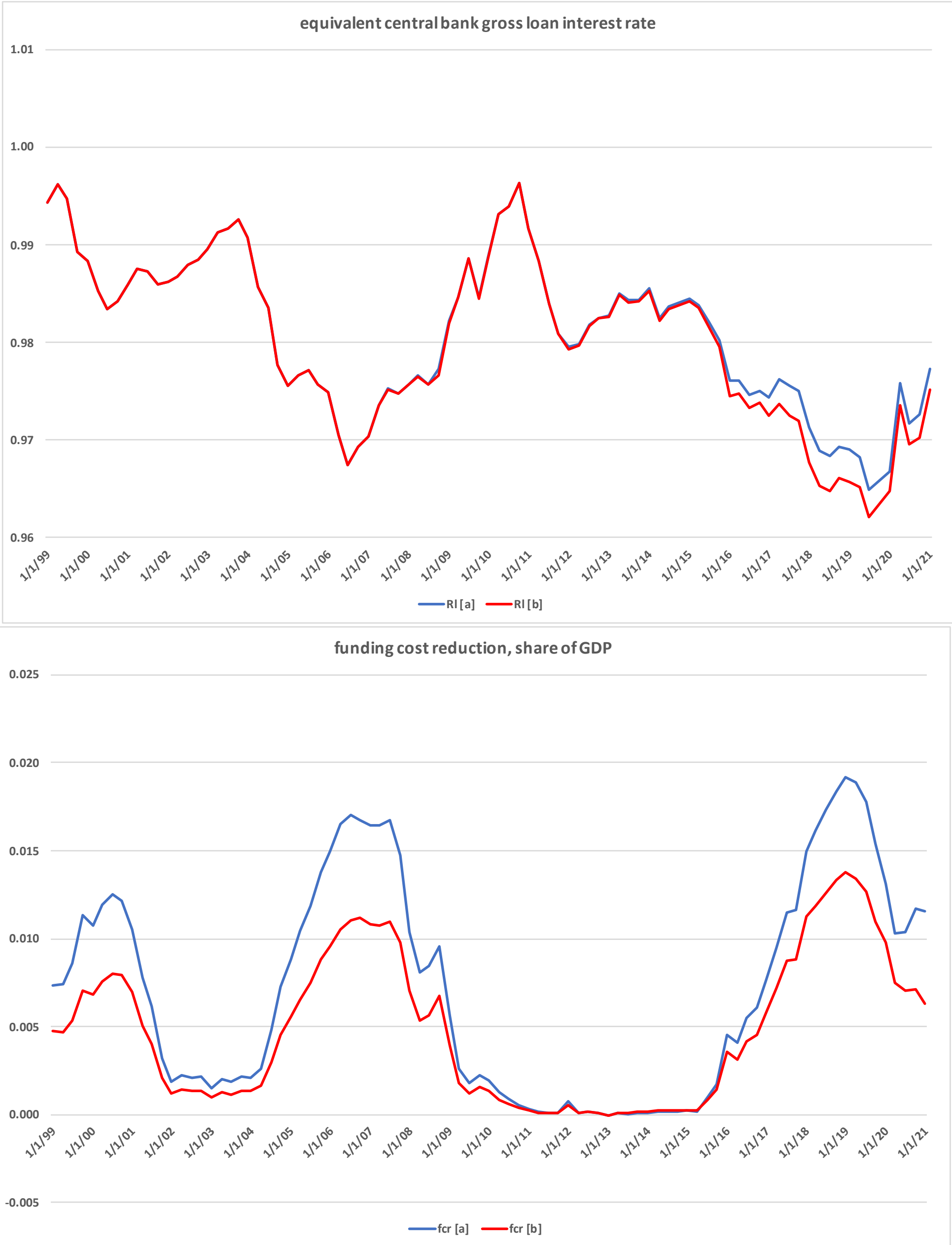

Figure 5: Two measures of the equivalent central bank gross loan interest rate (top) and banks' funding cost reduction (bottom) under the assumption that the optimality condition of banks holds and $\psi = 0.5$.

The equivalent loan rate again displays a downward trend, falling by roughly 2 percent over the sample period with a temporary reversal around the time of the financial crisis. The funding cost reduction varies between 0 and 2 percent of GDP. It exhibits the same cyclical behavior as in figure 4 but now is always positive since R^f_{t+1} always exceeds R^n_{t+1}.

In conclusion, irrespective of assumptions about bank operating costs the U.S. banking sector benefited from a substantial funding cost reduction prior to the financial crisis but did not benefit from such a cost reduction, or even bore additional funding costs, during and after the crisis. A similar picture starts to emerge for the period of the COVID-19 crisis. The funding cost reductions of between -0.7 and 2 percent of GDP compare with NIPA data for financial sector profits on the order of 3 percent of GDP prior to the financial crisis, negative profits during the crisis, and 2 to 3 percent of GDP after the financial crisis.

6 Politics

The analysis of section 5 suggests that the introduction of CBDC would expose banks and the economy to political rather than macroeconomic risks; while the central bank could in principle neutralize a swap of deposits into money this would require cheap refinancing of banks and such refinancing may be politically unpalatable.[45] In this section we focus on another channel through which political factors may undermine economic equivalence.

Recall that the central bank's profit between dates t and $t+1$ in a single-tiered monetary system equals $m_{t+1}\left(\chi^m_{t+1} - \mu_t\right)$.[46] Since the optimal policy implements the Friedman rule and the liquidity premium on money equals the operating costs (see equation (RA-1)) the central bank balances its budget, at least in present value.

In a two-tiered system, in contrast, the central bank's profit between dates t and $t+1$ equals $n_{t+1}\left\{\zeta_{t+1}\left(\chi^r_{t+1} - \rho_t\right) - \theta_t\right\}$.[47] Two effects structurally undermine budget balance in this case. First, the central bank makes losses on its reserves management operations if reserves generate externalities.[48] This follows from equation (RA-3), which implies

$$\chi^{r\star}_{t+1} - \rho_t = \chi^{r\star}_{t+1} + \nu_{1,t}(\zeta^\star_{t+1}, \zeta^\star_{t+1}) + \nu_{2,t}(\zeta^\star_{t+1}, \zeta^\star_{t+1}) = \nu_{2,t}(\zeta^\star_{t+1}, \zeta^\star_{t+1}) < 0.$$

Second, the central bank (or the treasury) optimally pay a positive or negative subsidy to induce banks to set the first-best liquidity premium on deposits, see equation (RA-4). Summing the two effects yields, from equations (RA-3) and (RA-4),

$$\zeta^\star_{t+1}\left(\chi^{r\star}_{t+1} - \rho_t\right) - \theta^\star_t = -\frac{1}{\eta_{n,t+1}} \frac{R^{n\star}_{t+1}}{R^{f\star}_{t+1}} < 0;$$

that is, the total budgetary impact under the optimal policy is unambiguously negative unless banks are competitive.

[45] More generally, the larger central bank balance sheet in a system with CBDC could invite more lobbying and political pressure on the central bank. See Niepelt (2020) for a discussion.

[46] This equals $m_{t+1}(\mathbb{E}_t[\mathrm{sdf}_{t+1}(R^k_{t+1} - R^m_{t+1})] - \mu_t)$.

[47] This equals $n_{t+1}\left\{\zeta_{t+1}(\mathbb{E}_t[\mathrm{sdf}_{t+1}(R^k_{t+1} - R^r_{t+1})] - \rho_t) - \theta_t\right\}$.

[48] To reduce the burden imposed on its budget the central bank could impose a minimum reserves requirement rather than subsidizing reserves holdings. This would be equivalent to subsidizing reserves holdings and simultaneously levying an inframarginal tax.

To understand this result, recall that the subsidy has two components. The first addresses the distortion due to market power, and the second neutralizes the effect of the central bank's low spread on reserves when reserves generate positive externalities. Since the second component of the subsidy exactly balances the fiscal effects of the reserves subsidy the net budgetary impact fully reflects the subsidy component that addresses market power.

We summarize this finding in the following proposition:

Proposition 9. Under the optimal policy, the government fiscally supports a deposit-based system unless banks are competitive. This is not the case with a money-based system.

Proposition 9 implies that—even outside crisis periods—policies to optimally support liquidity provision require less fiscal support when they rely on money rather than deposits. Fundamentally, this reflects distortions in the private sector and fiscal costs of correcting them; no such distortions are present when the government directly controls the public provision of liquidity.

As a consequence private liquidity provision suffers from a structural disadvantage whenever the shadow value of public relative to private funds exceeds unity. In the model of section 2 this is not the case because the government has access to nondistorting taxes. But in an environment with distorting taxes social welfare under the Ramsey policy would be strictly higher in a system with money rather than deposits as long as the resource costs of operating payments in the two systems were (approximately) equal. Even if this condition were not met the optimal money-based policy would be preferable to the optimal deposit-based policy from a narrow fiscal point of view.

This has implications for the political support for CBDC: Even if CBDC would be approximately "irrelevant" from a macroeconomic perspective (given our equivalence result) one should still expect some political support for the introduction of CBDC. Of course, this effect would interact with others and it could be outweighed by opposing forces, for instance when large taxpayers also collect a large share of bank profits. In a richer framework with heterogeneous groups of households the political support for CBDC would depend on the distribution of tax burdens, the ownership structure of banks, and the aggregation of preferences in the political process.

Another implication concerns central bank independence which is widely seen as a key pillar of stability oriented monetary policies. One factor contributing to this independence is a structural central bank surplus. A central bank with structural deficits that regularly requires fiscal support from the treasury is in a weaker position to defend its independence than a central bank that generates surpluses. From this perspective the model suggests that a CBDC-based payment system could better safeguard central bank independence.

7 Conclusion

We have analyzed monetary architectures and policies and in particular, the introduction of CBDC into a two-tiered monetary system. Our extension of the standard macroeconomic workhorse model suggests that there is no a priori case in favor or against CBDC;

whether the optimal monetary system is single-tiered, two-tiered or mixed depends on the resource costs per effective liquidity services in the alternative systems. In any optimal monetary system, spreads should satisfy modified Friedman (1969) rules; deposits should be taxed when the deposit market is sufficiently competitive; interest rates on reserves and CBDC should differ; and a second-best policy should target the composition of real balances between deposits and CBDC.

Plausibly, the marginal resource costs per effective liquidity services of CBDC- and deposit-based payments (the latter backed up by reserves operations) are approximately equal. Our equivalence analysis suggests that in this case the introduction of CBDC need not be disruptive for banks and the wider economy. In fact, the central bank can guarantee (approximate) equivalence when it refinances banks at their effective funding costs in the two-tiered system. We have characterized these costs, computed them for the U.S., and quantified the funding cost reduction for banks due to their ability to issue liquid liabilities.

We caution against hasty conclusions, however. On the one hand, the macroeconomic equivalence logic could break down when deposits are "special" or certain dimensions of heterogeneity important. On the other hand, the introduction of CBDC would likely change the political economy of banking.[49] Our analysis suggests that taxpayers might support a shift to a single-tiered system (if such a shift could be enforced at all) because it requires less taxpayer support. At the same time central banks could find it politically infeasible to refinance banks at the equivalent loan interest rate, rendering it impossible to neutralize macroeconomic effects of the introduction of CBDC. In other words, the main risks of the introduction of CBDC for banks and the economy may well be political, not macroeconomic.

Our framework lends itself to many extensions. One such extension could be to integrate price rigidities and policy motives related to demand stabilization. Another natural extension would be to introduce different dimensions of heterogeneity and to study political economy considerations in more detail, along the lines suggested above.

A General Equilibrium under Assumption 1

A.1 Households and Firms

Let $\mathcal{A}_t \equiv (1-\vartheta)c_t^{1-\psi} + \vartheta z_{t+1}^{1-\psi}$ such that the marginal utilities of consumption and real balances, respectively, are given by

$$\begin{aligned} u_c(c_t, z_{t+1}, x_t) &= (1-\vartheta)\mathcal{A}_t^{\frac{1-\sigma}{1-\psi}-1} c_t^{-\psi} v(x_t), \\ u_z(c_t, z_{t+1}, x_t) &= \vartheta \mathcal{A}_t^{\frac{1-\sigma}{1-\psi}-1} z_{t+1}^{-\psi} v(x_t). \end{aligned}$$

When deposits circulate, $n_{t+1} > 0$ and possibly $m_{t+1} > 0$ as well, the Euler equation

[49]See Gonzalez-Eiras and Niepelt (2015) for an abstract analysis of "politico-economic equivalence."

for real balances, equation (4), reduces to

$$\frac{\vartheta z_{t+1}^{-\psi}}{(1-\vartheta)c_t^{-\psi}} = \chi_{t+1}^n \quad \text{or} \quad z_{t+1} = c_t \left(\frac{\vartheta}{1-\vartheta} \frac{1}{\chi_{t+1}^n} \right)^{\frac{1}{\psi}}. \tag{4'}$$

When deposits do not circulate, $n_{t+1} = 0$ and $m_{t+1} > 0$, equation (4) reduces to

$$\lambda_t \frac{\vartheta(\lambda_t m_{t+1})^{-\psi}}{(1-\vartheta)c_t^{-\psi}} = \chi_{t+1}^m \quad \text{or} \quad z_{t+1} = \lambda_t m_{t+1} = c_t \left(\frac{\vartheta}{1-\vartheta} \frac{\lambda_t}{\chi_{t+1}^m} \right)^{\frac{1}{\psi}}. \tag{4''}$$

Let χ_{t+1} denote "the" spread on household means of payment. By convention, $\chi_{t+1} = \chi_{t+1}^m/\lambda_t$ when money circulates; $\chi_{t+1} = \chi_{t+1}^n$ when deposits circulate; and $\chi_{t+1} = \chi_{t+1}^m/\lambda_t = \chi_{t+1}^n$ when both money and deposits circulate (see equation (5)). Moreover, define

$$\begin{aligned}
\Omega_t^x(\chi_{t+1}) &\equiv (1-\vartheta)^{\frac{1-\sigma}{1-\psi}} \left(1 + \left(\frac{\vartheta}{1-\vartheta} \right)^{\frac{1}{\psi}} \chi_{t+1}^{1-\frac{1}{\psi}} \right)^{\frac{1-\sigma}{1-\psi}}, \\
\Omega_t^c(\chi_{t+1}) &\equiv (1-\vartheta)^{\frac{1-\sigma}{1-\psi}} \left(1 + \left(\frac{\vartheta}{1-\vartheta} \right)^{\frac{1}{\psi}} \chi_{t+1}^{1-\frac{1}{\psi}} \right)^{\frac{\psi-\sigma}{1-\psi}}.
\end{aligned}$$

(In the main text where we report the case with deposits and reserves, we write Ω_t^x and Ω_t^c as functions of the policy instruments χ_{t+1}^r and θ_t which determine χ_{t+1}^n.) We can then express the marginal utility of leisure and consumption, respectively, as

$$\begin{aligned}
\frac{\mathcal{A}_t^{\frac{1-\sigma}{1-\psi}}}{1-\sigma} v'(x_t) &= \frac{c_t^{1-\sigma}}{1-\sigma} \Omega_t^x(\chi_{t+1}) v'(x_t), \\
(1-\vartheta)\mathcal{A}_t^{\frac{1-\sigma}{1-\psi}-1} c_t^{-\psi} v(x_t) &= c_t^{-\sigma} \Omega_t^c(\chi_{t+1}) v(x_t),
\end{aligned}$$

independently of whether money, deposits, or both means of payment circulate. The household's optimality conditions conditional on optimal real balances holdings thus read

$$c_t^{-\sigma} v(x_t) \Omega_t^c(\chi_{t+1}) = \beta \mathbb{E}_t \left[R_{t+1}^k c_{t+1}^{-\sigma} v(x_{t+1}) \Omega_{t+1}^c(\chi_{t+2}) \right] \tag{2'}$$

for capital and

$$\frac{c_t^{1-\sigma}}{1-\sigma} v'(x_t) \Omega_t^x(\chi_{t+1}) = c_t^{-\sigma} v(x_t) \Omega_t^c(\chi_{t+1}) w_t \tag{3'}$$

for leisure.

A.2 Resource Constraint

Recall that $\xi_{t+1} \equiv n_{t+1}/z_{t+1}$ denotes the share of deposits in effective real balances. Defining

$$\Omega_t^{rc}(\chi_{t+1}, \zeta_{t+1}, \xi_{t+1}) \equiv 1 + \left(\frac{\vartheta}{1-\vartheta} \frac{1}{\chi_{t+1}} \right)^{\frac{1}{\psi}} \left(\xi_{t+1} \left\{ \nu_t(\zeta_{t+1}, \zeta_{t+1}) + \zeta_{t+1}\rho_t \right\} + (1-\xi_{t+1}) \frac{\mu_t}{\lambda_t} \right),$$

we can express the resource constraint subject to optimal real balances holdings as

$$\kappa_{t+1} = f_t(\kappa_t, 1-x_t) + \kappa_t(1-\delta) - c_t \Omega_t^{rc}(\chi_{t+1}, \zeta_{t+1}, \xi_{t+1}), \tag{15'}$$

where we use equations (4') and (4").

A.3 Banks

When deposits circulate, $n_{t+1} > 0$, reserves circulate as well, $r_{t+1} > 0$ and $\zeta_{t+1} > 0$. The bank's optimality condition for reserves, equation (9), implies

$$-\phi_{1,t}(1-\varphi)\zeta_{t+1}^{-\varphi} = \chi_{t+1}^r \quad \Rightarrow \quad \zeta_{t+1} = \left(\frac{\phi_{1,t}(\varphi-1)}{\chi_{t+1}^r}\right)^{\frac{1}{\varphi}}. \tag{9'}$$

Accordingly, in equilibrium,

$$\nu_t(\zeta_{t+1}, \zeta_{t+1}) = (\phi_{1,t} + \phi_{2,t})\left(\frac{\phi_{1,t}(\varphi-1)}{\chi_{t+1}^r}\right)^{\frac{1-\varphi}{\varphi}} \tag{17}$$

and

$$\begin{aligned}\tilde{\nu}_t(-\chi_{t+1}^r) &= (\phi_{1,t} + \phi_{2,t})\left(\frac{\phi_{1,t}(\varphi-1)}{\chi_{t+1}^r}\right)^{\frac{1-\varphi}{\varphi}} + \chi_{t+1}^r\left(\frac{\phi_{1,t}(\varphi-1)}{\chi_{t+1}^r}\right)^{\frac{1}{\varphi}} \\ &= (\phi_{1,t}\varphi + \phi_{2,t})\left(\frac{\chi_{t+1}^r}{\phi_{1,t}(\varphi-1)}\right)^{1-\frac{1}{\varphi}}.\end{aligned} \tag{18}$$

The spread on deposits depends on monetary policy. We consider several cases:

Central Bank Targets Composition of Real Balances When the central bank targets $\xi_{t+1} \equiv n_{t+1}/z_{t+1}$ equation (4') yields

$$n_{t+1} = \xi_{t+1} c_t \left(\frac{\vartheta}{1-\vartheta}\frac{1}{\chi_{t+1}^n}\right)^{\frac{1}{\psi}} \quad \Rightarrow \quad \eta_{n,t+1} \equiv \frac{n_{t+1}{}'(R_{t+1}^n)R_{t+1}^n}{n_{t+1}(R_{t+1}^n)} = \frac{1}{\psi}\frac{R_{t+1}^n}{R_{t+1}^f - R_{t+1}^n}$$

and the bank's optimality condition for deposits, equation (8a), reduces to

$$\chi_{t+1}^n = \frac{\tilde{\nu}_t(-\chi_{t+1}^r) - \theta_t}{1-\psi}. \tag{8a'}$$

Central Bank Targets Quantity of Money In this case, $n_{t+1} = z_{t+1} - \lambda_t m_{t+1}$ and ξ_{t+1} is endogenous. Equation (4') yields

$$n_{t+1} = c_t \left(\frac{\vartheta}{1-\vartheta}\frac{1}{\chi_{t+1}^n}\right)^{\frac{1}{\psi}} - \lambda_t m_{t+1} \quad \Rightarrow \quad \eta_{n,t+1} = \frac{1}{\psi}\frac{R_{t+1}^n}{R_{t+1}^f - R_{t+1}^n}\frac{1}{\xi_{t+1}}.$$

Accordingly, the bank's optimality condition reduces to

$$\chi_{t+1}^n = \frac{\tilde{\nu}_t(-\chi_{t+1}^r) - \theta_t}{1-\psi\xi_{t+1}}. \tag{8a''}$$

When the central bank does not issue money, $\xi_{t+1} = 1$ and (8a') and (8a") coincide.

Central Bank Targets Spread on Money We refer to the profit maximizing deposit spread χ^n_{t+1} in equation (8a') subject to $\xi_{t+1} = 1$ as the "monopsony spread." When the central bank targets χ^m_{t+1} and sets it in excess of λ_t times the monopsony spread (rendering money unattractive for households) then monetary policy is irrelevant and $m_{t+1} = 0$.

When the central bank targets χ^m_{t+1} and sets it below λ_t times the monopsony spread (rendering money attractive) then banks have a choice between raising the interest rate on deposits to the competitive rate given in equation (5) (plus epsilon), thereby pricing the central bank out of the market, or not raising the rate and being priced out of the market themselves. As long as the deposit spread covers the average operating costs,

$$\chi^n_{t+1} \geq \nu_t(\zeta_{t+1}, \zeta_{t+1}) - \theta_t + \zeta_{t+1}\chi^r_{t+1} = \tilde{\nu}_t(-\chi^r_{t+1}) - \theta_t,$$

where $\tilde{\nu}_t(-\chi^r_{t+1})$ is given by condition (18), banks optimally choose the former option, that is, they satisfy equation (5).

In conclusion, when the central bank targets χ^m_{t+1} there are three cases to distinguish:

(a) When $\frac{\chi^m_{t+1}}{\lambda_t} < \tilde{\nu}_t(-\chi^r_{t+1}) - \theta_t$ subject to (18) banks would lose money when trying to compete with the central bank. In equilibrium, $n_{t+1} = 0$.

(b) When $\tilde{\nu}_t(-\chi^r_{t+1}) - \theta_t \leq \frac{\chi^m_{t+1}}{\lambda_t} < \frac{\tilde{\nu}_t(-\chi^r_{t+1}) - \theta_t}{1-\psi}$ subject to (18) banks issue deposits at a spread lower than the monopsony spread and equal to (slightly less than) χ^m_{t+1}/λ_t in order to compete with the central bank. In equilibrium, $m_{t+1} = 0$.

(c) When $\frac{\tilde{\nu}_t(-\chi^r_{t+1}) - \theta_t}{1-\psi} \leq \frac{\chi^m_{t+1}}{\lambda_t}$ subject to (18) banks issue deposits at the monopsony spread; monetary policy is irrelevant. In equilibrium, $m_{t+1} = 0$.

A.4 General Equilibrium

Combining these results we conclude that an equilibrium allocation, price, and payment system satisfies conditions (11), (12), (2'), (3'), (15') as well as the following restrictions:

- In a monetary system with deposits, $\chi_{t+1} = \chi^n_{t+1}$ as well as conditions (4'), (9'), (17), (18) and
  - when the central bank targets ξ_{t+1}, condition (8a') and $m_{t+1} = (1-\xi_{t+1})z_{t+1}/\lambda_t$;
  - when it targets m_{t+1}, condition (8a") and $\xi_{t+1} = 1 - \lambda_t m_{t+1}/z_{t+1}$;
  - when it targets χ^m_{t+1}, the conditions described under (b) or (c) above.
- In a monetary system without deposits, $\chi_{t+1} = \chi^m_{t+1}/\lambda_t$ as well as condition (4").

B Optimality under Assumption 1

Social Planner Allocation When the social planner opts for deposits then $\zeta^\star_{t+1}$ satisfies

$$\nu_{1,t}(\zeta^\star_{t+1}, \zeta^\star_{t+1}) + \nu_{2,t}(\zeta^\star_{t+1}, \zeta^\star_{t+1}) = -\rho_t \quad \Rightarrow \quad \zeta^\star_{t+1} = \left(\frac{(\phi_{1,t} + \phi_{2,t})(\varphi - 1)}{\rho_t}\right)^{\frac{1}{\varphi}}.$$

The shadow liquidity premium on reserves thus equals

$$-\nu_{1,t}(\zeta^\star_{t+1}, \zeta^\star_{t+1}) = \rho_t \frac{\phi_{1,t}}{\phi_{1,t} + \phi_{2,t}}.$$

Note that $\nu_{1,t}(\zeta^\star_{t+1}, \zeta^\star_{t+1})$ equals its equilibrium counterpart, $\nu_{1,t}(\zeta_{t+1}, \zeta_{t+1})$, when $\chi^r_{t+1} = \rho_t\phi_{1,t}/(\phi_{1,t} + \phi_{2,t})$. Also,

$$\nu_t(\zeta^\star_{t+1}, \zeta^\star_{t+1}) = (\phi_{1,t} + \phi_{2,t}) \left(\frac{(\phi_{1,t} + \phi_{2,t})(\varphi - 1)}{\rho_t}\right)^{\frac{1-\varphi}{\varphi}}.$$

The shadow liquidity premium on deposits therefore is given by

$$\nu_t(\zeta^\star_{t+1}, \zeta^\star_{t+1}) + \zeta^\star_{t+1}\rho_t = \left(\frac{\rho_t}{(\phi_{1,t} + \phi_{2,t})(\varphi - 1)}\right)^{1-\frac{1}{\varphi}} (\phi_{1,t} + \phi_{2,t})\varphi.$$

Ramsey Policy The optimal liquidity premia under the Ramsey policy follow directly from the shadow premia in the social planner allocation. As for the optimal subsidy, note from equation (8a') that optimal bank choices are consistent with the first-best allocation if

$$\chi^{n\star}_{t+1} = \frac{\tilde{\nu}_t(-\chi^{r\star}_{t+1}) - \theta^\star_t}{1 - \psi}.$$

Solving this equation for $\theta^\star_t$ (or substituting the expression for equilibrium $\eta_{n,t+1}$ into the expression for $\theta^\star_t$ given in proposition 3) yields

$$\theta^\star_t = \left(\frac{\rho_t}{(\phi_{1,t} + \phi_{2,t})(\varphi - 1)}\right)^{1-\frac{1}{\varphi}} (\psi(\phi_{1,t} + \phi_{2,t})\varphi - \phi_{2,t}(\varphi - 1)).$$

C Proof of Proposition 8

We state proposition 8 formally, derive its implications for portfolios and budget sets, and prove it. We consider an intervention which reduces deposit holdings at date t by $\Delta > 0$ and increases money holdings at date t by $\lambda_t^{-1}\Delta$. To guarantee nonnegativity of deposits, capital holdings, and reserves the intervention Δ must not be too large.[50]

Proposition Suppose condition 1 holds. Consider a policy and equilibrium with deposits, reserves, and no central bank loan. There exists another policy and equilibrium, indicated by circumflexes, with fewer deposits and reserves, more money, a central bank loan, a different ownership structure of capital, possibly household taxes at dates t and $t + 1$ whose market value equals zero, and the same allocation and price system.

[50] Specifically, we impose $\Delta \leq n_{t+1}$, $(1 - \lambda_t^{-1})\Delta \leq k^g_{t+1}$, $(1 - \lambda_t^{-1})\Delta \geq -k_{t+1}$, and $\zeta_{t+1}\Delta \leq r_{t+1}$. The assumption that the central bank loan in the initial equilibrium equals zero is a convenient normalization.

The two policies and equilibria coincide except that

$$\hat{m}_{t+1} = m_{t+1} + \lambda_t^{-1}\Delta, \quad \hat{n}_{t+1} = n_{t+1} - \Delta, \quad \hat{l}_{t+1} = \Delta(1-\zeta_{t+1}), \quad \hat{r}_{t+1} = r_{t+1} - \zeta_{t+1}\Delta,$$
$$\hat{k}_{t+1} = k_{t+1} + (1-\lambda_t^{-1})\Delta, \quad \hat{k}^g_{t+1} = k^g_{t+1} - (1-\lambda_t^{-1})\Delta.$$

The household tax at date t, $\hat{T}_{1,t}$, and the state contingent tax at date $t+1$, $\hat{T}_{2,t+1}$, satisfy

$$\begin{aligned}\hat{T}_{1,t} &= \Delta\{\nu_t(\zeta_{t+1},\zeta_{t+1}) - \theta_t\},\\ \hat{T}_{2,t+1} &= \Delta\{(1-\lambda_t^{-1})R^k_{t+1} + \lambda_t^{-1}R^m_{t+1} - R^n_{t+1} - (\nu_t(\zeta_{t+1},\zeta_{t+1}) - \theta_t)R^f_{t+1}\}.\end{aligned}$$

The central bank loan carries the interest rate

$$R^l_{t+1} = \frac{R^n_{t+1} + (\nu_t(\zeta_{t+1},\zeta_{t+1}) - \theta_t)R^f_{t+1} - \zeta_{t+1}R^r_{t+1}}{1-\zeta_{t+1}}.$$

Implications for Portfolios and Budget Sets Note that in an equilibrium with deposits reserves are strictly positive such that $0 < \zeta_{t+1} < 1$.

The portfolio and policy changes described in the proposition have several implications. First, $\hat{z}_{t+1} = z_{t+1}$, $\hat{m}_{t+1}\mu_t + \hat{n}_{t+1}\nu_t(\zeta_{t+1},\zeta_{t+1}) + \hat{r}_{t+1}\rho_t = m_{t+1}\mu_t + n_{t+1}\nu_t(\zeta_{t+1},\zeta_{t+1}) + r_{t+1}\rho_t$, and $\hat{\kappa}_{t+1} = \kappa_{t+1}$; that is, the portfolio changes do not alter effective real balances, the aggregate capital stock, or the total resource costs of operating the payment system. Note that the portfolio changes also leave the reserves-to-deposits ratio, ζ_{t+1}, and thus $\nu_t(\zeta_{t+1},\zeta_{t+1})$ unchanged.

Second, the length of household balance sheets does not change since households swap assets (deposits, money, and capital). In contrast, the balance sheets of banks shorten because banks hold fewer reserves and reduce total borrowing (in the form of deposits and central bank loans) but invest the same amount in capital as before the intervention. The balance sheet of the consolidated government expands by $(\lambda_t^{-1} - \zeta_{t+1})\Delta \lesseqgtr 0$: The central bank raises additional funds Δ/λ_t from households but fewer funds from banks, $-\zeta_{t+1}\Delta$; it passes $\Delta(1-\zeta_{t+1})$ through to the banking sector and increases capital holdings by $(\lambda_t^{-1} - 1)\Delta$.

Third, the new tax at date t compensates for the reduced bank losses borne by households, $\hat{\pi}^b_{1,t} - \pi^b_{1,t} = \hat{T}_{1,t}$. Similarly, the new state contingent taxes at date $t+1$ compensate for the change in the return on the household portfolio as well as the change in bank profits that households collect at date $t+1$.[51] The market value of the two taxes as of date t equals zero, $\hat{T}_{1,t} + \mathbb{E}_t[\mathrm{sdf}_{t+1}\hat{T}_{2,t+1}] = 0$, and so does the market value of the changes in bank profits.[52] As a consequence, the household's dynamic and intertemporal budget constraints continue to be satisfied with the modified portfolios and policy.

[51] The former component equals $(\hat{k}_{t+1} - k_{t+1})R^k_{t+1} + (\hat{m}_{t+1} - m_{t+1})R^m_{t+1} + (\hat{n}_{t+1} - n_{t+1})R^n_{t+1} = (1-\lambda_t^{-1})\Delta R^k_{t+1} + \lambda_t^{-1}\Delta R^m_{t+1} - \Delta R^n_{t+1}$ and the latter $\hat{\pi}^b_{2,t+1} - \pi^b_{2,t+1} = -\Delta(R^k_{t+1} - R^n_{t+1}) + \Delta(1-\zeta_{t+1})(R^k_{t+1} - R^l_{t+1}) - \zeta_{t+1}\Delta(R^r_{t+1} - R^k_{t+1}) = -\Delta(\nu_t(\cdot) - \theta_t)R^f_{t+1}$, where we use the expression for the loan rate.

[52] The market value of the two taxes equals Δ times $(\nu_t(\cdot) - \theta_t) + (1-\lambda_t^{-1}) + \lambda_t^{-1}R^m_{t+1}/R^f_{t+1} - R^n_{t+1}/R^f_{t+1} - (\nu_t(\cdot) - \theta_t) = 0$, where we use equation (5). The market value of the change in bank profits equals $\hat{\pi}^b_{1,t} - \pi^b_{1,t} + \mathbb{E}_t[\mathrm{sdf}_{t+1}(\hat{\pi}^b_{2,t+1} - \pi^b_{2,t+1})] = \hat{T}_{1,t} - \Delta(\nu_t(\cdot) - \theta_t) = 0$.

Fourth, the same holds true for the government. From equation (13), the government budget constraint at date t reads $\hat{k}^g_{t+1} + \hat{l}_{t+1} - \hat{m}_{t+1} - \hat{r}_{t+1} = k^g_t R^k_t - m_t R^m_t - r_t R^r_t + \tau_t - \hat{n}_{t+1}\theta_t - \hat{m}_{t+1}\mu_t - \hat{r}_{t+1}\rho_t + \hat{T}_{1,t}$. Condition 1 implies that this equality is satisfied if and only if the budget constraint was satisfied before the intervention. Similarly, the government budget constraint at date $t+1$, $k^g_{t+2} - m_{t+2} - r_{t+2} = \hat{k}^g_{t+1} R^k_{t+1} + \hat{l}_{t+1} R^l_{t+1} - \hat{m}_{t+1} R^m_{t+1} - \hat{r}_{t+1} R^r_{t+1} + \tau_{t+1} - n_{t+2}\theta_{t+1} - m_{t+2}\mu_{t+1} - r_{t+2}\rho_{t+1} + \hat{T}_{2,t+1}$, is equivalent to the constraint before the change of portfolios and policy.

Proof Conjecture that the price system (including the interest rates on money and deposits, which satisfy $R^f_{t+1} - R^m_{t+1} = \lambda_t(R^f_{t+1} - R^n_{t+1}))$ does not change, as claimed in the proposition. The optimal production decisions of firms are unchanged in this case, as are firm profits. Moreover, the market values of households' time endowments, taxes, and bank profits (as shown above) also do not change. As a consequence, household wealth is unaffected, and so are the optimal consumption, labor supply, and real balances sequences. As shown above, these sequences are supported by the modified portfolios. It remains to be shown that the modified bank portfolios are optimal; in that case, all budget constraints are satisfied at the optimal choices, equilibrium capital accumulation is unchanged, and the conjecture is verified.

To render deposits $\hat{n}_{t+1}$, loans $\hat{l}_{t+1}$, and reserves $\hat{r}_{t+1}$ optimal for a bank it suffices for the central bank to structure the loans such that the bank's choice sets before and after the intervention coincide. Before the intervention, this choice set is determined by the cost function, ν_t; the subsidy, θ_t; the deposit funding schedule; the stochastic discount factor; and the returns on capital and reserves. After the intervention, it is defined by the same cost function, subsidy rate, stochastic discount factor, and returns; a modified deposit funding schedule (because households hold more money); and a central bank loan funding schedule.

To assure identical choice sets it therefore suffices for the central bank to post an appropriate loan funding schedule. This schedule makes up for the shift in the deposit funding schedule, corrected for the fact that $1-\zeta_{t+1}$ dollars of central bank loans provide the same net funding as one dollar of deposits of which ζ_{t+1} dollars are invested in reserves.[53] Subject to the appropriate loan funding schedule a bank chooses loans that make up for the reduction in funding (net of reserves) from households, at the same effective price; moreover, it chooses the same reserves-to-deposits ratio such that $\nu_t(\zeta_{t+1}, \zeta_{t+1})$ is unaffected.

D Data

We use the quarterly average of the FRED series `IOER` (2008q4–2021q1) for the nominal interest rate on reserves. To compute the gross real interest rate we divide the gross nominal rate by the gross inflation rate, Π_{t+1} (see below). Since no interest on reserves

[53] If the banking sector were competitive no pass through funding would be required; equivalence could also be achieved if banks shed assets, see Brunnermeier and Niepelt (2019).

was paid prior to 2008[54] we set $R^r_{t+1} = 1/\Pi_{t+1}$ prior to 2008.

We use the quarterly average of the FRED series `CPILFESL_PC1` (Consumer Price Index for All Urban Consumers: All Items Less Food and Energy in U.S. City Average, Percent Change from Year Ago, Index 1982–1984=100, Seasonally Adjusted) for gross inflation.

We use quarterly averages of Kurlat's (2019) monthly estimates (1999m01–2017m12) of the risk-free rate and the deposit rate.[55] Kurlat (2019) provides two estimates of the latter (based on RateWatch data), one based on data for checking accounts and the other for money market accounts. We compute a weighted average of the two estimates where the weights correspond to the relative size of checkable and savings deposits (see below). We adjust the constructed interest rate series using the inflation series defined before. Since Kurlat's (2019) series end in 2017 we extrapolate them at the end of the sample using projections on the quarterly averages of the FRED series `TB3MS` (3-Month Treasury Bill: Secondary Market Rate, Percent, Not Seasonally Adjusted) and `AAA` (Moody's Seasoned Aaa Corporate Bond Yield, Percent, Not Seasonally Adjusted).

We use quarterly averages of the FRED series `TCDSL` (Total Checkable Deposits, Billions of Dollars, Seasonally Adjusted) and `SAVINGSL` (Savings Deposits - Total, Billions of Dollars, Seasonally Adjusted) for checkable and savings deposits, respectively. Since the series `TCDSL` and `SAVINGSL` were discontinued after 2020q4 and 2020q1, respectively, we extrapolate them at the end of the sample using projections on the quarterly average of the FRED series `DEMDEPSL` (Demand Deposits: Total, Billions of Dollars, Seasonally Adjusted).

We use the quarterly average of the FRED series `RESBALNS` (Total Reserve Balances Maintained with Federal Reserve Banks, Billions of Dollars, Not Seasonally Adjusted) for reserves. Since the series `RESBALNS` was discontinued after 2020q2 we use the quarterly average of the FRED series `BOGMBBM` (Reserve Balances, Millions of Dollars, Not Seasonally Adjusted; divided by thousand) for the most recent periods.

We use two alternative series for deposits. The first series ([a]) is the sum of the quarterly averages of the FRED series `TCDSL` and `SAVINGSL` defined before. The second series ([b]) is the sum of the quarterly average of the FRED series `TCDSL` and the quarterly money market deposit account (MMDA) series constructed by Lucas and Nicolini (2015).[56] Since the updated MMDA series ends in 2020q2 we extrapolate it at the end of the sample using projections on the cumulative sums of the quarterly FRED series `HNOMMFQ027S` (Households and Nonprofit Organizations; Money Market Fund Shares; Asset, Flow, Millions of Dollars, Seasonally Adjusted; divided by thousand), `BOGZ1FA103034000Q` (Nonfinancial Corporate Business; Money Market Fund Shares; Asset, Flow, Millions of Dollars, Seasonally Adjusted; divided by thousand), and `NNBMMFQ027S` (Nonfinancial Noncorporate Business; Money Market Fund Shares; Asset, Flow, Millions of Dollars, Seasonally Adjusted; divided by thousand) as well as lags of the series.

We compute ζ_{t+1} as the ratio of the reserve series and either of the two deposit series. We compute deposits as a share of GDP as the ratio of either of the two deposit series and the FRED series `GDP` (Gross Domestic Product, Billions of Dollars, Quarterly, Seasonally

[54]See `https://www.federalreserve.gov/monetarypolicy/reqresbalances.htm`.

[55]We thank Pablo Kurlat for sharing his data.

[56]We thank Luca Benati for sharing the MMDA series.

Adjusted Annual Rate).

We use the quarterly FRED series `A587RC1Q027SBEA` (Corporate profits with inventory valuation and capital consumption adjustments: Domestic industries: Financial) for financial sector profits.

References

Andolfatto, D. (2018). Reconciling orthodox and heterodox views on money and banking, *Review of Economic Analysis* **10**(4): 351–370.

Andolfatto, D. (2021). Assessing the impact of central bank digital currency on private banks, *Economic Journal* **131**(634): 525–540.

Auer, R. and Böhme, R. (2020). The technology of retail central bank digital currency, *BIS Quarterly Review* pp. 85–100.

Auer, R., Cornelli, G. and Frost, J. (2020). Rise of the central bank digital currencies: Drivers, approaches and technologies, *Working Paper 880*, Bank for International Settlements, Basel.

Barrdear, J. and Kumhof, M. (2016). The macroeconomics of central bank issued digital currencies, *Staff Working Paper 605*, Bank of England, London.

Benes, J. and Kumhof, M. (2012). The Chicago plan revisited, *Working Paper 12/202*, International Monetary Fund, Washington.

Bianchi, J. and Bigio, S. (2020). Banks, liquidity management, and monetary policy. Unpublished, Federal Reserve Bank of Minneapolis and UCLA.

Bindseil, U. (2020). Tiered CBDC and the financial system, *Working Paper 2351*, European Central Bank, Frankfurt.

BIS (2018). Central bank digital currencies, *Report*, BIS: Committee on Payments and Market Infrastructures, and Markets Committee, Basel.

Bolton, P., Li, Y., Wang, N. and Yang, J. (2020). Dynamic banking and the value of deposits, *Working Paper 2020-13*, Ohio State University, Fisher College of Business, Columbus.

Böser, F. and Gersbach, H. (2020). Monetary policy with a central bank digital currency: The short and the long term, *Discussion Paper 15322*, CEPR, London.

Brunnermeier, M. K. and Niepelt, D. (2019). On the equivalence of private and public money, *Journal of Monetary Economics* **106**: 27–41.

Bryant, J. (1983). Government irrelevance results: A simple exposition, *American Economic Review* **73**(4): 758–761.

Bullard, J. and Smith, B. D. (2003). The value of inside and outside money, *Journal of Monetary Economics* **50**(2): 389–417.

Chamley, C. and Polemarchakis, H. (1984). Assets, general equilibrium and the neutrality of money, *Review of Economic Studies* **51**(1): 129–138.

Chari, V. V. and Phelan, C. (2014). On the social usefulness of fractional reserve banking, *Journal of Monetary Economics* **65**(C): 1–13.

Chiu, J., Davoodalhosseini, M., Jiang, J. and Zhu, Y. (2019). Bank market power and central bank digital currency: Theory and quantitative assessment, *Staff Working Paper 2019-20*, Bank of Canada, Ottawa.

Croushore, D. (1993). Money in the utility function: Functional equivalence to a shoppingtime model, *Journal of Macroeconomics* **15**(1): 175–182.

Di Tella, S. (2020). Risk premia and the real effects of money, *American Economic Review* **110**(7): 1995–2040.

Di Tella, S. and Kurlat, P. (2021). Why are banks exposed to monetary policy?, *American Economic Journal: Macroeconomics* **forthcoming**.

Drechsler, I., Savov, A. and Schnabl, P. (2017). The deposits channel of monetary policy, *Quarterly Journal of Economics* **132**(4): 1819–1876.

Drechsler, I., Savov, A. and Schnabl, P. (2021). Banking on deposits: Maturity transformation without interest rate risk, *Journal of Finance* **76**(3).

Farhi, E. and Tirole, J. (2021). Shadow banking and the four pillars of traditional financial intermediation, *Review of Economic Studies* **forthcoming**.

Faure, S. and Gersbach, H. (2018). Money creation in different architectures, *Discussion Paper 13156*, CEPR.

Feenstra, R. C. (1986). Functional equivalence between liquidity costs and the utility of money, *Journal of Monetary Economics* **17**(2): 271–291.

Fischer, S. (1972). Keynes-Wicksell and neoclassical models of money and growth, *American Economic Review* **62**(5): 880–890.

Fisher, I. (1935). *100% Money*, Adelphi, New York.

Freixas, X. and Rochet, J.-C. (2008). *Microeconomics of Banking*, 2. edn, MIT Press, Cambridge, Massachusetts.

Friedman, M. (1969). The optimum quantity of money, *in* M. Friedman (ed.), *The Optimum Quantity of Money and Other Essays*, Aldine, Chicago, chapter 1, pp. 1–50.

Galí, J. (2015). *Monetary Policy, Inflation, and the Business Cycle*, 2. edn, Princeton University Press, Princeton.

Garratt, R. J. and van Oordt, M. R. C. (2021). Privacy as a public good: A case for electronic cash, *Journal of Political Economy* **129**(7): 2157–2180.

Garratt, R. J. and Zhu, H. (2021). On interest-bearing central bank digital currency with heterogeneous banks. Unpublished, University of California Santa Barbara.

Gonzalez-Eiras, M. and Niepelt, D. (2015). Politico-economic equivalence, *Review of Economic Dynamics* **18**(4): 843–862.

Gurley, J. G. and Shaw, E. S. (1960). *Money in a Theory of Finance*, Brookings Institution, Washington.

Jackson, T. and Pennacchi, G. (2021). How should governments create liquidity?, *Journal of Monetary Economics* **118**: 281–295.

Kahn, C. M., McAndrews, J. and Roberds, W. (2005). Money is privacy, *International Economic Review* **46**(2): 377–399.

Kahn, C. M., Rivadeneyra, F. and Wong, T.-N. (2018). Should the central bank issue e-money?, *Staff Working Paper 2018-58*, Bank of Canada, Ottawa.

Kashyap, A. K., Rajan, R. and Stein, J. C. (2002). Banks as liquidity providers: An explanation for the coexistence of lending and deposit-taking, *Journal of Finance* **57**(1): 33–73.

Keister, T. and Monnet, C. (2020). Central bank digital currency: Stability and information. Unpublished, Rutgers University and University of Bern.

Keister, T. and Sanches, D. (2020). Should central banks issue digital currency? Unpublished, Rutgers University and Federal Reserve Bank of Philadelphia.

King, R. G., Plosser, C. I. and Rebelo, S. T. (1988). Production, growth, and business cycles I: The basic neoclassical model, *Journal of Monetary Economics* **21**: 195–232.

Kiyotaki, N. and Moore, J. (2019). Liquidity, business cycles, and monetary policy, *Journal of Political Economy* **127**(6): 2926–2966.

Klein, M. A. (1971). A theory of the banking firm, *Journal of Money, Credit, and Banking* **3**(2): 205–218.

Knight, F. H., Cox, G. V., Director, A., Douglas, P. H., Fisher, I., Hart, A. G., Mints, L. W., Schultz, H. and Simons, H. C. (1933). Memorandum on banking reform. Henry C. Wallace papers.

Kumhof, M. and Noone, C. (2018). Central bank digital currencies—design principles and balance sheet implications, *Staff Working Paper 725*, Bank of England, London.

Kurlat, P. (2019). Deposit spreads and the welfare cost of inflation, *Journal of Monetary Economics* **106**: 78–93.

Lagos, R. and Zhang, S. (2020). The limits of onetary economics: On money as a latent medium of exchange, *Working Paper 26756*, NBER, Cambridge, Massachusetts.

Lucas, R. E. and Nicolini, J. P. (2015). On the stability of money demand, *Journal of Monetary Economics* **73**: 48–65.

McCallum, B. T. and Goodfriend, M. S. (1987). Demand for money: Theoretical studies, *in* J. Eatwell, P. Newman and M. Milgate (eds), *The New Palgrave: A Dictionary of Economics*, Macmillan Press, London, pp. 775–781.

Monti, M. (1972). Deposit, credit and interest rate determination under alternative bank objective functions, *in* K. Shell and G. P. Szegö (eds), *Mathematical Methods in Investment and Finance*, North-Holland, Amsterdam, pp. 431–454.

Niepelt, D. (2018). Reserves for all? Central Bank Digital Currency, deposits, and their (non)-equivalence, *Discussion Paper 13065*, CEPR.

Niepelt, D. (2020). Reserves for all? Central Bank Digital Currency, deposits, and their (non)-equivalence, *International Journal of Central Banking* **16**(3): 211–238.

Parlour, C. A., Rajan, U. and Walden, J. (2020). Payment system externalities. Unpublished, University of California Berkeley and University of Michigan.

Philippon, T. (2015). Has the US finance industry become less efficient? On the theory and measurement of financial intermediation, *American Economic Review* **105**(4): 1408–1438.

Piazzesi, M. and Schneider, M. (2021). Credit lines, bank deposits or CBDC? competition and efficiency in modern payment systems. Unpublished, Stanford University.

Rocheteau, G. and Nosal, E. (2017). *Money, Payments, and Liquidity*, 2. edn, MIT Press, Cambridge, Massachusetts.

Sargent, T. J. (1987). *Dynamic Macroeconomic Theory*, Harvard University Press, Cambridge, Massachusetts.

Saving, T. R. (1971). Transactions costs and the demand for money, *American Economic Review* **61**(3): 407–420.

Schilling, L., Fernández-Villaverde, J. and Uhlig, H. (2020). Central bank digital currency: When price and bank stability collide, *Working Paper 28237*, NBER, Cambridge, Massachusetts.

Sidrauski, M. (1967). Rational choice and patterns of growth in a monetary economy, *American Economic Review* **57**(2): 534–544.

Taudien, R. (2020). Borrowing costs, liquidity and full reserve banking. Thesis paper, University of Bern.

Tobin, J. (1963). Commercial banks as creators of "money", *Discussion Paper 159*, Cowles Foundation, New Haven.

Tobin, J. (1965). Money and economic growth, *Econometrica* **33**(4): 671–684.

Tobin, J. (1969). A general equilibrium approach to monetary theory, *Journal of Money, Credit, and Banking* **1**(1): 15–29.

Tobin, J. (1985). Financial innovation and deregulation in perspective, *Bank of Japan Monetary and Economic Studies* **3**(2): 19–29.

Vandeweyer, Q. (2019). Treasury debt and the pricing of short-term assets. Unpublished, European Central Bank and Sciences Po.

Wallace, N. (1981). A Modigliani-Miller theorem for open-market operations, *American Economic Review* **71**(3): 267–274.

Walsh, C. E. (2017). *Monetary Theory and Policy*, 4. edn, MIT Press, Cambridge, Massachusetts.

Williamson, S. (2019). Central bank digital currency: Welfare and policy implications. Unpublished, University of Western Ontario.

Woodford, M. (2003). *Interest and Prices*, Princeton University Press, Princeton.