The Declining Role of Deposits in Credit Creation*

Naz Koont
Stanford University

Stefan Walz
Boston College

October 23, 2026

Abstract

Does deposit funding still drive credit creation? Deposit inflows affect credit directly, through banks' balance sheet response, and indirectly, through a feedback loop in which credit generates new deposits. Using exogenous deposit shocks and a structural model, we show that the total bank-credit impact of a $1 deposit inflow has fallen from $2.14 to $0.62 over two decades. Both channels have weakened: banks extend less credit out of each marginal deposit, reflecting greater costs of balance sheet expansion, and fewer credit proceeds recycle into stable deposits, consistent with declining deposit specialness and a greater use of non-bank alternatives. Central bank policy also matters: Quantitative Easing lowers the credit impact of a marginal deposit through higher capital costs, while Quantitative Tightening can raise it through scarcer liquidity. We show that non-bank credit is unlikely to offset the decline. These findings point to an implicit narrowing of banking that dampens credit creation and reshapes monetary and fiscal policy transmission.

*We thank Michele Andreolli, Luigi Bocola, Greg Buchak, Mayur Chaudhary (Discussant), Matteo Crosignani (Discussant), Darrell Duffie, Manasa Gopal (Discussant), Zhiguo He, Nicolas Inostroza (Discussant), Arvind Krishnamurthy, Jane Li (Discussant), Wenhao Li, Silvio Petriconi (Discussant), Rodney Ramcharan, Martin Schneider, Amit Seru, Jeremy Stein, Phil Strahan, Adi Sunderam, Christopher Tonetti, Vikrant Vig and seminar participants at the IU Holden Finance Conference, FRBSF Macro-Finance Conference, UGA Fall Finance Conference, NFA, USC Macro-Finance Workshop, Banking and the Deposit Franchise SITE Session, Boston College, UCSC, and Stanford GSB for helpful comments.

1 Introduction

Does deposit funding still drive credit creation? Stable bank deposits are widely thought to support bank lending and maturity transformation, as they relax banks' liquidity management (Diamond and Dybvig, 1983; Kashyap et al., 2002a; Drechsler et al., 2017). Beyond this direct effect, lending can generate new stable deposits, creating a feedback loop that amplifies the effect of deposit inflows on credit creation (Tobin, 1963; Friedman and Schwartz, 1963). Despite its potential importance, this amplification channel has received relatively little attention in the modern banking literature. As a result, we know little about the total effect of deposit inflows on bank credit provision, or about how that effect changes as the financial system evolves and intermediation shifts outside banks.

In this paper, we decompose the total effect of deposit inflows on credit into a direct credit response and an indirect amplification through the credit-deposit feedback loop. We document a pronounced decline in both channels over the past two decades, and use a structural model to trace these declines to shifts in primitives driven by higher balance sheet costs and non-bank disintermediation. Our results point to an implicit narrowing of banking: deposits have become substantially less effective at generating credit, even as formal deposit-taking remains integrated with lending, with implications for the cyclicality of credit, the transmission of monetary and fiscal policy, and the supply of safe assets.

We begin by providing empirical evidence on the magnitude and evolution of banks' credit-deposit multiplier using weekly data from 2005 to 2025. The multiplier embeds both banks' direct credit response to a deposit inflow and the recycling of credit into new deposits, and thus captures how an initial inflow propagates through successive rounds of credit creation. To estimate it, we use high-frequency variation in the timing of U.S. Treasury cash settlements, purged of predictable seasonal and recurring components, as an instrument for deposit inflows, and trace the cumulative responses of deposits and bank credit over the following weeks.

We find that the credit-deposit multiplier is sizable: a \$1 deposit inflow ultimately raises cumulative deposits by \$1.68 through balance sheet propagation. However, this amplification has weakened markedly over time, with the multiplier declining from \$2.43 in the early period to \$1.04 in the late period, and with the bank-credit response declining in parallel. These declines indicate a weakening of one or both of the margins that govern the loop: banks may lend less against a marginal deposit dollar, or fewer credit proceeds may return as stable deposits that

support another round of lending. The reduced form shows that the loop has weakened, but not why. Post-crisis regulation, central bank balance-sheet policy, the growth of non-bank savings and credit markets, and shifts in household behavior could each affect both margins. To trace the decline to these primitives, ask how credit creation responds as they change, and recover the response of total credit including non-banks, we next develop a structural model of the banking sector.

Our structural model is built around the forces we seek to quantify: balance sheet costs, non-bank competition, and the recycling of credit into deposits. Households are at once savers and borrowers, allocating liquid wealth across bank deposits, market savings, and cash, and borrowing from banks or non-banks; a banking system sets deposit and loan rates, and a non-bank credit intermediary competes with banks for borrowers. The size and composition of the banking system's balance sheet are disciplined by the costs of maintaining capital and liquidity buffers, while non-bank competition shapes the elasticity of bank credit demand. We characterize the equilibrium of this economy in two layers. The first is a static portfolio equilibrium: taking household liquid wealth as given, markets clear as a consistent allocation of that wealth across claims, with banks and non-banks optimizing subject to their constraints. This layer pins down banks' direct credit response to deposit inflows as an equilibrium object shaped by balance sheet costs and competition. The second layer is a monetary-creation channel, summarized by a single recycling rate: the fraction of credit that returns as stable deposits. We microfound this rate in household behavior. Bank credit creates liquid wealth for the households that receive its proceeds, and they allocate it through the same portfolio choice that governs the first layer. The recycling rate is therefore the deposit share of the marginal recipients of credit, and it falls when credit reaches recipients that hold fewer deposits or when deposits become less special.

We recover the direct credit response structurally by differentiating the equilibrium conditions, and discipline the recycling rate with data. We estimate the model's parameters via deterministic minimum distance, following the logic of structural banking models that use externally disciplined demand elasticities and equilibrium restrictions to recover supply-side frictions (Wang et al., 2022). The banking sector's deposit and loan first-order conditions discipline the shadow costs of illiquidity and leverage, while the non-bank first-order condition disciplines the cost of expanding non-bank credit supply. We jointly fit these pricing residuals and the reduced-form credit-deposit multipliers in the early and late periods.

The estimated model delivers the two sufficient statistics that govern the credit-deposit

multiplier: banks' credit response to deposit inflows and the recycling rate of credit proceeds into stable deposits. Both have declined sharply. The credit response falls from 0.86 to 0.58: about a third less of each marginal deposit dollar is channeled into new credit, with the remainder absorbed by reductions in wholesale funding. The recycling rate falls from 0.67 to 0.09, a sharp decline in the share of credit-generated payment flows that return to banks as the stable, low-cost funding that supports further credit expansion. The microfoundation of the recycling rate points to two forces behind its decline: the shift in bank credit from loans toward securities, which sends credit proceeds to institutional sellers rather than to households and firms, and the smaller share of financial assets that households and firms hold as deposits. We attribute the residual to a shift, within sectors, toward recipients of credit proceeds who hold fewer deposits.

Turning to the central question of the paper, we find that the total bank-credit effect of a deposit inflow has declined from 2.14 to 0.62: each dollar of deposits today generates less than a third of the bank credit it once did. The total bank credit effect combines banks' direct credit response with the amplification of deposits through the multiplier. The decline reflects both a weaker direct credit response and diminished deposit recycling, which weakens the amplification from one round to the next. As a result, the financial system has lost an important amplification channel through this implicit narrowing of the banking sector. Consequently, concerns about deposits migrating from banks to non-banks such as stablecoins may have more muted bank-credit effects today than they would have had in earlier periods, when deposit funding translated more powerfully into credit creation.

We next use the model to study the drivers of this decline in the credit impact of deposit flows. We decompose the decline into two channels. The first exercise returns the structural primitives governing bank behavior from their late- to their early-period values, one at a time and together. Two forces emerge as first-order, and they compound: reduced recycling and the higher capital wedge each lower the total bank-credit response on their own, and their interaction, since a weaker lending response leaves less credit to recycle and weaker recycling leaves less funding to lend, accounts for about half of the decline. The liquidity wedge is small and works in the opposite direction, slightly offsetting the decline by raising the liquidity benefit of the marginal deposit dollar. The second exercise varies balance-sheet quantities, asking how Quantitative Easing and Tightening affect the banking sector's capacity to intermediate deposits into credit. The object here is the marginal bank-credit response to a new deposit dollar—not the direct credit-supply effect of QE itself. In the model, the two operate on the same balance-sheet

frictions in opposite directions. QE injects reserves and deposits together: liquidity improves, but the balance sheet expands, tightening the capital position and lowering the marginal value of a deposit. QT instead can tighten the liquidity constraint, as emphasized by Acharya and Rajan (2024), removing reserves without a commensurate decline in deposit claims and raising the marginal value of stable funding. Together, the two exercises show that the decline reflects higher balance sheet costs, reduced deposit specialness, and the transformation of the balance sheet itself.

A natural question is whether the non-bank sector is able to compensate for the dampened credit response in banks. We approach this from two angles. First, we examine the direct competitive response of non-banks to a deposit shock and find that non-bank credit contracts rather than expands: the inflow reaches non-banks only through lower bank loan rates, which pull borrowers away from non-bank credit, so non-bank credit contracts. A more elastic non-bank sector deepens the contraction rather than closing the gap. Second, we ask how strong a spillover from non-recycled credit proceeds to non-bank funding conditions would need to be to restore the aggregate credit response to its early-period level. Even in the most favorable case, in which all non-recycled proceeds flow into money markets, restoring the early-period response requires a sizable pass-through from money-market depth to non-bank intermediation costs. More broadly, because non-bank intermediation chains are longer, neither recycled non-bank proceeds nor leaked bank flows are likely to compensate for the loss in bank recycling. On the direct margin, non-banks contract rather than substitute; on the funding margin, they do not replace the multiplier.

Overall, we find that deposits injected into the banking system generate far less cumulative amplification of credit than they did in the past, and non-bank intermediaries are unlikely to fully offset this decline. As a result, both money and credit have become less responsive to deposit inflows, dampening the amplification of the credit cycle. In this sense, our findings document an implicit narrowing of banking, driven not by institutional separation, but by regulation and non-bank competition in the payments market, which has weakened the traditional feedback between deposits and credit. Deposits have become less effective at supporting credit creation, even as formal deposit-taking remains integrated with lending.

Literature Contribution

Whether deposit funding matters for bank credit is a classic question in banking. It con-

nects to seminal theories of why banks jointly provide deposit-taking and credit provision services (Diamond and Dybvig, 1983; Kashyap et al., 2002a; Drechsler et al., 2017), as well as to a more recent debate about the consequences of non-bank intermediaries eroding banks' deposit base (Whited et al., 2022). Our contribution is to quantify how deposit inflows translate into bank credit creation and to link this relationship to underlying economic primitives. By doing so, we provide a framework for evaluating how shifts in the composition of money and savings, including the migration of funds toward non-bank instruments such as money market funds or stablecoins, affect the supply of credit in the modern financial system.

Moreover, the idea that deposits and loans are connected through a feedback loop back at least to Tobin (1963), yet remains underexplored. The literature has tended to emphasize two separate views: one in which deposits fund loans, and another in which loans create deposits. In the traditional "intermediary" view, banks collect deposits from savers, hold a fraction as reserves, and lend out the remainder, so deposit market size constrains lending (Bernanke and Blinder, 1988; Kashyap et al., 2002b; Xiao, 2019). In the "money creation" view, lending itself creates deposits (Tobin, 1963; Donaldson et al., 2018), and in the absence of a binding reserve requirement is shaped by regulations, loan demand, and banks' risk appetite. Empirically, micro evidence shows that banks create deposits when they lend (Werner, 2014), aggregate deposits far exceed cash deposits (Thakor and Edison, 2024), and reduced-form estimates document a positive elasticity from deposit shocks to credit growth (Kundu et al., 2021). Our contribution is to bridge these perspectives by quantifying the strength of the endogenous credit-deposit multiplier and the amount of deposits and credit that are ultimately created from an initial deposit inflow. This multiplier highlights that deposits both constrain and are created by lending: the availability and stability of deposits shape banks' willingness to expand credit, while banks' lending decisions feed back into the deposit base. By decomposing the evolution and drivers of the multiplier, we show how the balance between these forces shifts over time.[1]

The credit-deposit multiplier relates to amplification mechanisms emphasized in the macro-finance literature. A large body of work studies how intermediary balance sheet strength shapes

[1]While we estimate a long-term liquidity multiplier that supports credit provision, Denbee et al. (2021) and Piazzesi and Schneider (2018) study a related short-term payments liquidity multiplier, showing how transactional liquidity shapes banks' reserve management and monetary transmission. Li et al. (2026) show that interbank payment flows redistribute liquidity across banks, so that a bank lends more when payment inflows generated by other banks' lending relax its liquidity constraint, and Li and Li (2022) show that banks facing greater payment risk lend less. While these papers study how payments redistribute liquidity among banks, we trace out how aggregate deposit flows remain within the banking system as stable funding.

credit supply through net-worth-based amplification. Seminal contributions show how collateral values and internal liquidity constrain borrowing and investment (Kiyotaki and Moore, 1997; Holmstrom and Tirole, 1997), while Bernanke et al. (1999) highlights how stronger balance sheets reduce financing frictions and expand borrowing. Building on these foundations, intermediary-based macro models emphasize how intermediary net worth governs leverage and credit creation (Gertler and Kiyotaki, 2010; Gertler and Karadi, 2011), with implications for risk premia and asset prices (Brunnermeier and Sannikov, 2014; He and Krishnamurthy, 2013). More recent heterogeneous-bank frameworks study how these elasticities vary across institutions (Corbae and D'Erasmo, 2021; Jamilov and Monacelli, 2025). Our framework studies a closely related amplification mechanism operating through funding rather than intermediary net worth separately. The marginal lending response captures how strongly banks expand credit in response to deposit inflows and therefore plays a role analogous to the credit supply elasticities emphasized in this literature. While we focus on short-run balance sheet propagation through funding flows, the two mechanisms would interact at longer horizons. Lending generates profits that raise depositor wealth and bank equity, which in turn strengthen the feedback loop between deposit creation and credit supply. In this sense, the credit-deposit multiplier complements the net-worth-based amplification mechanisms emphasized in the macro-finance literature.

We highlight the importance of accounting for the recycling of credit proceeds into stable deposit funding. This connects to recent work on deposit franchise value and depositor heterogeneity in both the cross-section and the time series (Drechsler et al., 2017, 2021; Li et al., 2023; DeMarzo et al., 2024; Blickle et al., 2025; Egan et al., 2025; Lu et al., 2024; Koont et al., 2024), as well as to the growing literature on non-bank intermediaries and the implications of a shrinking traditional banking sector (Begenau and Landvoigt, 2022; Buchak et al., 2024; Schneider et al., 2025). The recycling rate measures how much of banks' own credit creation returns as stable funding: a high rate means that credit provision replenishes the deposit franchise, while its decline implies that credit expansion generates less franchise value per dollar. Similarly, as non-banks disintermediate banks on both sides of their balance sheets, a smaller share of funding flows through this amplification mechanism, raising the question of whether non-bank intermediation can compensate. We show that it is unlikely to: in equilibrium, non-bank credit contracts rather than expands in response to a deposit inflow, and because non-bank intermediation chains are longer, neither recycled non-bank proceeds nor leaked bank flows are likely to make up the loss in bank recycling.

2 Defining the Total Effect of Deposit Inflows on Bank Credit

2.1 Banks' Lending Response to Deposit Inflows

The central question in this paper is how deposit inflows translate into bank credit creation. When deposits enter the banking system, banks may expand their balance sheets by increasing lending or other assets. A natural starting point for thinking about this relationship is banks' marginal credit response to deposits, which we denote by $\lambda = dL/dD$: how much additional credit banks supply when their deposit funding increases.

However, the total effect of deposit inflows on bank credit may differ from this marginal response because deposits and credit are jointly determined within the banking system. When banks extend credit, the associated payment flows circulate, and a fraction, which we denote by θ, returns as the stable deposits that fund further credit. As a result, the ultimate increase in credit following a deposit inflow depends both on banks' marginal credit response λ and on the recycling rate θ.

Our object of interest is the total effect of a deposit inflow D_0 on bank credit L, given by dL/dD_0, which embeds both the marginal response λ and the amplification generated through θ. The remainder of this section develops this object: how deposits and credit are jointly determined in a fractional-reserve system, a simple example of the propagation mechanism, and a formal decomposition into these two margins.

2.2 The Link Between Deposits and Bank Credit

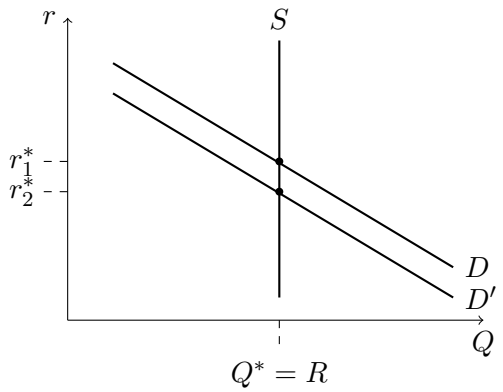
To understand how deposit inflows translate into bank credit creation, we begin by discussing how deposits and credit are jointly determined in the modern banking system, where banks operate with fractional reserves and are not constrained by binding reserve requirements. In such an environment, the effect of deposit inflows on credit depends on how bank balance sheets adjust and whether credit provision generates additional deposits within the financial system.

As a benchmark, consider a fully narrow banking system in which deposits are backed one-for-one by central bank reserves, as illustrated in Figure 1 panel (a). Under narrow banking, the aggregate quantity of deposits is fixed by reserves, so that $Q^* = R$, where R denotes aggregate

Figure 1: Deposit determination under narrow and fractional reserve banking

Narrow Banking

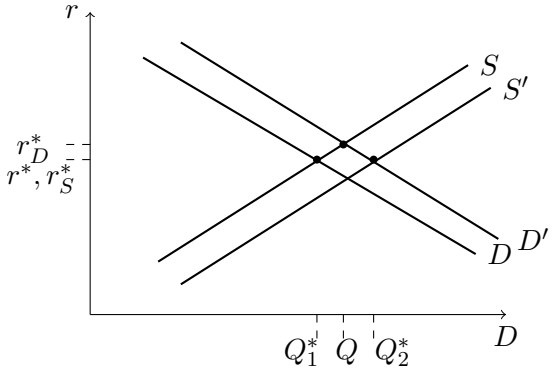
Assets		**Liabilities**	
Reserves	R	Deposits	R
	R		R



Panel A: Narrow banking: deposits are backed one-for-one by reserves, so deposit quantity is pinned down by aggregate reserves. Demand shifts change equilibrium rates but not quantities.

Fractional Reserve Banking

Assets		**Liabilities**	
Reserves	R	Deposits	D
Loans / Securities	A		
	$R + A$		D



Panel B: Non-narrow banking: deposits are jointly determined by an upward-sloping deposit supply curve (banks' willingness to expand balance sheets) and a downward-sloping deposit demand curve (households' willingness to hold deposits). Shifts in either demand or supply move both equilibrium rates and quantities.

reserves. In this case, the aggregate supply of deposits is perfectly inelastic. Changes in the demand for deposits shift equilibrium deposit rates r^* but do not affect the quantity of deposits.

In contrast, in a fractional reserve banking system as in Figure 1 panel (b), deposits are jointly determined by (i) households' willingness to hold deposits and (ii) banks' willingness to expand their balance sheets and issue deposit liabilities (e.g., to fund loans and securities). Unlike in a narrow system, banks are not required to hold reserves one-for-one against deposits. Instead, they choose balance sheet size to maximize their profits subject to liquidity and capital management or regulatory constraints that limit how much asset expansion can be supported by a given stock of liquid resources. In this environment, shifts on either side of the market affect both equilibrium rates and quantities. Bank balance sheet expansion can therefore be initiated either by an increase in deposit demand, arising from changes in household preferences

or portfolio allocations toward deposits, or by a shift in credit supply, reflecting improvements in lending opportunities or changes in banks' regulatory environment that make asset expansion more attractive.[2]

The same balance sheet expansion can also be described from the perspective of the household sector that holds the bank's liabilities. Bank deposits are part of household liquid wealth, and bank loans are household liabilities. When the bank extends credit, it creates a deposit and a loan of equal value, and both appear on the household balance sheet: household liquid wealth and household debt rise by the same amount, and household net worth is unchanged. Bank credit thus adds to the liquid wealth of the household sector without adding to its net worth. For the bank to expand its balance sheet in equilibrium, the household sector must be willing to hold the additional claims. How the households that end up with the proceeds hold them, as deposits or as money market fund shares, determines how much of the credit returns to the banking system as stable funding. This is the recycling margin that we formalize in Section 4, where it is derived from household portfolio choices.

We see evidence of this deposit creation in the U.S. banking sector when we examine the evolution of key monetary aggregates over time. We focus on two standard measures of money that are most directly linked to the banking system's role in money creation. M0, or the monetary base, includes currency in circulation and reserves held by banks at the central bank and represents the supply of base money injected by the central bank. M1, or narrow money, includes currency in circulation as well as demand deposits and other highly liquid checkable deposits. Comparing these two aggregates illustrates how the stock of bank deposits compares to reserves. Figure 2 plots the evolution of M1 relative to M0. The fact that the ratio fluctuates over time, and generally exceeds 1, is consistent with banks creating deposits through credit creation.[3]

The fact that banks create deposits to expand their balance sheets does not, by itself, imply the existence of a feedback loop between deposits and credit. Such amplification arises only if banks respond to additional deposits by expanding credit provision, so that the marginal credit response to deposits, dL/dD, is positive. This condition is likely to hold in practice because deposits, when they constitute stable funding, relax banks' liquidity constraints at the margin

[2]This example abstracts from banks' non-deposit wholesale funding adjustment. Allowing for this margin would imply that balance sheet expansion occurs net of funding adjustments, as formalized in our model in Section 4.

[3]The ratio drops below 1 during the QE period primarily because central bank reserves expanded disproportionately relative to the bank deposit categories in M1.

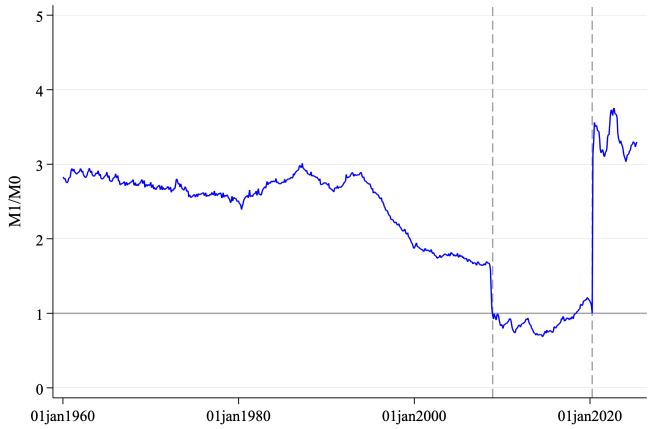


Figure 2: Evolution of M1/M0. Monthly data is from FRED (1960–2025). Vertical lines denote the onset of QE after the 2008 financial crisis and pandemic-related stimulus. In May 2020, savings deposits were reclassified as transaction accounts and added to M1.

Table 1: Simple Example: Credit-Deposit Multiplier

In this numerical example, a \$100 deposit inflow replaces wholesale funding (b). The banking sector expands credit by $\lambda = 0.7$ per dollar of stable deposits, and loan proceeds appear as payment balances (PB) that circulate as borrowers spend (c). A fraction $\theta = 0.9$ of these flows settles as stable deposits and the remainder as wholesale funding (d), initiating the next round of credit creation (e) and recycling (f).

(a) Initial

Assets		**Liabilities**	
R	\$100	W	\$100
	\$100		\$100

(b) Deposit Inflow

Assets		**Liabilities**	
R	\$100	D	\$100
		W	\$0
	\$100		\$100

(c) Credit Creation

Assets		**Liabilities**	
R	\$100	D	\$100
L	\$70	PB	\$70
	\$170		\$170

(d) Deposit Recycling

Assets		**Liabilities**	
R	\$100	D	\$163
L	\$70	W	\$7
	\$170		\$170

(e) Credit Creation

Assets		**Liabilities**	
R	\$100	D	\$163
L	\$114	PB	\$44
		W	\$7
	\$214		\$214

(f) Deposit Recycling

Assets		**Liabilities**	
R	\$100	D	\$203
L	\$114	W	\$11
	\$214		\$214

and thereby support additional balance sheet expansion. As a result, deposits created through credit can return to the banking system as funding and support further balance sheet expansion. The ultimate increase in credit can therefore exceed banks' initial marginal credit response.

2.3 A Simple Example of the Credit-Deposit Feedback Loop

The example in Table 1 illustrates how an initial deposit inflow can propagate through the banking system via successive rounds of balance sheet expansion, using a consolidated balance sheet for the aggregate banking system. A deposit-demand shock first lets banks replace wholesale funding with stable deposits, leaving total assets unchanged but improving the funding mix. Because stable deposits relax banks' liquidity management, suppose that banks expand credit. The credit proceeds initially appear as flighty payment balances at the issuing bank, and the intermediary view of banks stops here, whereby an increase in deposit funding leads to an effect on banks' willingness to extend credit.

However, the quantity of deposits is endogenously determined, and this is what sets the feedback loop in motion. As borrowers spend the credit proceeds, payments circulate through the economy. Over time, some fraction of these proceeds settles as deposit funding in the banking system, and the rest as flightier wholesale funding. The recycled deposits again relax banks' liquidity management and support another round of credit expansion. Some of the resulting credit proceeds again return as deposit funding, continuing the feedback loop.

Each entry in Table 1 also has a counterpart on the balance sheet of the household sector: the borrower owes the loan, and the seller of the asset that the borrower purchases holds the payment balance and subsequently the deposit. Appendix A.2 tracks the household side of the example panel by panel and separates the borrower from the seller. Two features of the household ledger are worth noting. The net worth of the household sector is unchanged throughout, since each claim created through lending is matched by a corresponding liability, while its liquid wealth and its debt both increase by the amount of credit extended. Lending thus expands gross positions rather than net wealth. For the bank to expand its balance sheet in equilibrium, the household sector must be willing to hold the additional claims, which in the aggregate is its willingness to carry the additional debt: payments move deposits between households without extinguishing them, and households carry the debt because borrowing is worthwhile at the loan rate the bank offers, so no change in deposit pricing is required. The fraction of loan proceeds that is retained as deposits reflects the portfolio choices of the households that end up holding them. We formalize this margin in Section 4.

This process continues until it converges, as shown in Table 2. Amplification arises because deposits are superior to wholesale funding: they relax the banking sector's liquidity man-

Table 2: Aggregate Banking System: Incremental Expansion by Round

This table reports the round-by-round expansion of deposits and loans under the credit-deposit multiplier, using the numerical example from Table 1. Each round's new credit creation equals $\lambda = 0.7$ times newly retained deposits, which in turn equal $\theta = 0.9$ times the previous round's credit. Successive rounds therefore shrink by the factor $\lambda\theta = 0.63$, and the final row reports the limiting value of the resulting geometric series.

Round h	**Deposits**	**Credit**
1	\$100.0	\$70.0
2	\$63.0	\$44.1
3	\$39.7	\$27.8
4	\$25.0	\$17.5
Limit $h \to \infty$	**\$270.3**	**\$189.2**

agement, and proceeds from credit creation partially recycle into new deposits. The interaction between banks' credit response and the retention of payment flows governs the strength of the credit-deposit multiplier.

2.4 The Total Effect of Deposit Inflows on Bank Credit

Formally, we define the credit-deposit multiplier m_D as the total change in the deposit base, ΔD_{total}, generated by an initial deposit inflow ΔD_0. Two sufficient statistics govern this process. The first is λ, banks' marginal credit response to deposit inflows.[4] The second is θ, the marginal propensity for credit-financed expenditures to be redeposited within the banking system as stable bank deposits.

Given a deposit inflow ΔD_0, banks extend new credit in the amount $\lambda\Delta D_0$, of which a fraction θ returns as deposits, raising the deposit base to $\Delta D_0 + \lambda\theta\Delta D_0$. These new deposits $\lambda\theta\Delta D_0$ then support additional credit provision of $\lambda^2\theta\Delta D_0$, of which a fraction θ recycles as deposits, and the process continues. Iterating forward, we can express the total deposit expansion from an initial inflow ΔD_0 as a geometric series:

$$\Delta D_{t-1,t+h} = \Delta D_0 \left(1 + \lambda\theta + (\lambda\theta)^2 + \cdots + (\lambda\theta)^h\right). \tag{1}$$

[4]In general, the first-round response depends on the source of the inflow—a preference shift, a reserve injection, or recycled payment flows—while every subsequent round involves the same response to recycled deposits. Section 4.7.1 formalizes this distinction; the multiplier logic here requires only the recycled-deposit response, which governs all rounds beyond the first.

The corresponding credit-deposit multiplier at horizon h is:

$$m_D^h = 1 + \lambda\theta + (\lambda\theta)^2 + \cdots + (\lambda\theta)^h. \tag{2}$$

As $h \to \infty$, the process converges (provided $|\lambda\theta| < 1$), yielding the long-run credit-deposit multiplier:

$$m_D = \frac{1}{1 - \lambda\theta}. \tag{3}$$

The credit-deposit multiplier m_D describes how an initial deposit inflow expands the deposit base through successive rounds of credit creation and redepositing. To recover the total effect of deposit inflows on bank credit ΔL_{total}^{bank}, we combine this amplification mechanism with banks' marginal lending response to deposits,

$$\Delta L_{total}^{bank} = \lambda m_D \Delta D_0. \tag{4}$$

In Appendix A.1 we consider whether non-bank institutions can generate a similar funding multiplier. They are unlikely to: because non-bank liabilities do not serve as the economy-wide settlement asset, credit proceeds return to non-banks only through longer intermediation chains, attenuating at each link, and the limited recycling that occurs within these subsystems typically does not support long-term credit provision to the real economy. We therefore focus on banks in the main text and abstract from non-bank multipliers.

3 Empirical Analysis

3.1 Empirical Strategy and Data

To estimate banks' credit-deposit multiplier, we measure how an initial deposit inflow translates into a total change in deposits. A natural way to implement this is to estimate local projections of deposits at horizon h on a shock to deposits at time t:

$$\Delta D_{t+h} = \alpha_h + m_D^h \cdot \Delta D_t + \Gamma_h X_t + \varepsilon_{t+h}. \tag{5}$$

In Equation (5), ΔD_t is deposit growth from period $t-1$ to t and ΔD_{t+h} is cumulative deposit growth from $t-1$ to $t+h$. X_t denotes a vector of controls dated at time t, and ε_{t+h} is the error term. In this setting, the coefficient m_D^h represents the credit-deposit multiplier at horizon h. As h grows large, m_D^h converges to the long-run multiplier m_D.

If all other determinants of deposit growth were fully accounted for in X_t, then local projections would deliver unbiased estimates of m_D^h at each horizon, converging to the true multiplier m_D for large h. However, in practice, many relevant factors that influence deposit dynamics, such as local demand conditions, interest rate expectations, or shifts in portfolio preferences, are difficult or impossible to observe. As a result, ΔD_t may be correlated with the error term ε_{t+h}, giving rise to endogeneity and biasing our estimate of m_D. For instance, we may overestimate m_D if ΔD_t and ΔD_{t+h} co-move positively because of an unobservable factor that drives deposit flows at both horizons, such as a surge in aggregate risk appetite or liquidity demand. To address this endogeneity concern, we seek a shifter of contemporaneous deposit growth, ΔD_t, that is uncorrelated with the unobservable factors that jointly drive both current and future deposit growth. We construct such a shifter using high-frequency innovations to Treasury cash flows from the Daily Treasury Statement (DTS).

Treasury cash flows reflect the government's operating receipts and outlays, including corporate and income taxes, tax refunds, federal salaries, grants to states, vendor payments, and interest on Treasury debt. These flows move funds between the public sector and private bank deposits. We aggregate daily Treasury withdrawals and deposits of operating cash into weekly net cash flows, CF_t, treating deposits of operating cash as negative values so that positive realizations correspond to net Treasury payments to the private sector. Our goal is to isolate variation in routine Treasury operations, so we exclude categories that may reflect financial system transfers or endogenous policy responses.[5] The cash-flow innovation is constructed using the full weekly DTS series. After constructing the innovation, the estimation sample excludes January weeks because the turn of the year combines large scheduled federal payments, tax-withholding timing, and year-end balance-sheet adjustments that add noise to the weekly link between Treasury settlements and core deposit growth.

We next purge this weekly series of predictable variation by residualizing on quarter-year

[5]We include all operating cash categories except for internal transfers to and from the Federal Reserve, the TGA, and depository institutions, and counter-cyclical means-based transfers including unemployment insurance, food assistance, and TANF, which may correlate with broader economic conditions. We exclude October 2008, which coincides with extraordinary fiscal and financial interventions during the Great Financial Crisis.

fixed effects α_q, calendar week-of-year fixed effects α_w, and five consecutive weekly lags, as in Equation (6). The lags absorb short-run settlement dynamics. We normalize the residual by lagged aggregate deposits, $Z_t \equiv \widetilde{u}_t / D_{t-1}$, so that a given dollar shock is scaled relative to the deposit base over which it is absorbed.[6]

$$CF_t = \alpha_q + \alpha_w + \sum_{k \in \mathcal{K}} \rho_k \, CF_{t-k} + u_t, \qquad \mathcal{K} = \{1, 2, 3, 4, 5\}. \tag{6}$$

This innovation captures unexpected timing variation in Treasury settlement flows. The identifying variation is the administrative timing of Treasury settlements after removing predictable seasonal and short-run variation. Because the outcome is cumulative, the instrument must be serially uncorrelated over the projection horizon. If a Treasury innovation in week t predicted further innovations in later weeks, cumulative deposit growth would include those later inflows, and the estimated multiplier would mix the propagation of a single inflow with the arrival of new ones. The residualization in Equation (6) is designed to remove this persistence. Appendix Figure A.2 shows that little remains after normalization and conditioning on our controls. When we project future values of the instrument on its current value, the responses at horizons one through eight are close to zero.

Instrument validity requires that the relevance and exclusion assumptions be satisfied. The relevance condition is supported by the first-stage estimates reported in Appendix Table A.2. Intuitively, Treasury payments and receipts credit and debit private bank accounts, generating changes in aggregate core deposits at weekly horizons. The exclusion restriction requires that innovations in Treasury cash-flow timing affect future deposit growth only through its impact on contemporaneous deposit growth ΔD_t, and not through any other channel. This is plausible because, after removing predictable seasonality and persistence, the remaining high-frequency variation in Treasury cash-flows largely reflects administrative settlement timing across a large number of payment streams.

We implement this strategy in a 2SLS framework. We measure deposits using aggregate core deposits from FRED, defined as seasonally adjusted total deposits held by all U.S. commercial banks excluding seasonally adjusted large time deposits. Deposit growth ΔD_t is the log difference of core deposits, winsorized at the 1% level. Our sample runs weekly from

[6] We focus on the surprise component of government spending, while Greenwood et al. (2015) instrument for Treasury bill issuance using the predictable seasonal component of bill supply, the fitted values from regressing four-week changes in bills outstanding on week-of-year dummies.

2005 to 2025; we estimate the multiplier over the full sample and separately for an early period (2005–2010) and a late period (2011–2025). In the vector of controls X_t we include four lags of deposit growth and the contemporaneous change in the effective federal funds rate, measured by the change in its Thursday-Wednesday weekly average. Because some Treasury payments settle with short delays, we instrument contemporaneous deposit growth with both the contemporaneous and one-period-lagged innovation, so that the first stage accommodates settlement effects that reach deposit balances in the following week. Our first-stage specification is:

$$\Delta D_t = \alpha + \beta_1 Z_t + \beta_2 Z_{t-1} + \Gamma_h X_t + \mu_t. \tag{7}$$

The first-stage estimates are reported in Appendix Table A.2. In the second stage, we regress cumulative deposit growth at horizon h on instrumented deposit growth at time t within the local projection framework. For horizon h, the second stage regression equation is,

$$\Delta D_{t-1,t+h} = \alpha_h + m_D^h \widehat{\Delta D_t} + \Gamma_h X_t + \varepsilon_{t+h}. \tag{8}$$

We consider three potential threats to identification. First, in the late period, Treasury settlements move reserves as well as deposits, because changes in the Treasury's account at the Federal Reserve are no longer offset through balances held at commercial banks. Reserve movements could then drive subsequent deposit growth independently of the deposit inflow itself. Appendix Figure A.4 shows that the late-period estimates are nearly unchanged when we control for either the lagged level or the contemporaneous change in aggregate reserves, indicating that the estimates primarily reflect deposit inflows rather than reserve co-movement at weekly frequency. The model in Section 4 accommodates both cases, allowing deposit inflows to arrive either on their own or together with reserves. Second, persistence in Treasury settlement timing could cause current innovations to predict further inflows within the projection horizon. As discussed above, Appendix Figure A.2 shows that the instrument has little persistence. Third, the deposit response could reflect real-side feedback rather than balance-sheet propagation: deposit-funded lending may stimulate economic activity, raising incomes and wealth and shifting out deposit demand. Such adjustments operate over business-cycle horizons, however, and the eight-week window we study is too short for them to play a meaningful role.

Thus, our empirical strategy allows us to estimate how cumulative deposit growth responds to an initial exogenous inflow through balance sheet propagation, thereby recovering a causal

estimate of the deposit multiplier.

3.2 Estimates of the Credit-Deposit Multiplier

Figure 3 visualizes the coefficient estimates for our IV specification in Equation (8) at horizons h from 0 to 8, with shaded bands indicating 95% confidence intervals based on Newey-West standard errors with a lag length equal to the horizon h. Panel (a) captures the estimation for the complete sample. This specification captures how an initial funding inflow propagates through the banking system over time, including through flows of funds between banks and across institutions. We define the long-run multiplier m_D by averaging the coefficients at horizons five and six. The full-sample multiplier is 1.68, implying that a $1 exogenous deposit inflow raises cumulative deposits by approximately $1.68. The impulse response indicates that initial deposit inflows are gradually re-deposited within the banking system, leading to additional rounds of deposit growth across local institutions. These effects level off after several weeks, suggesting that the cumulative response has converged to the overall deposit multiplier.

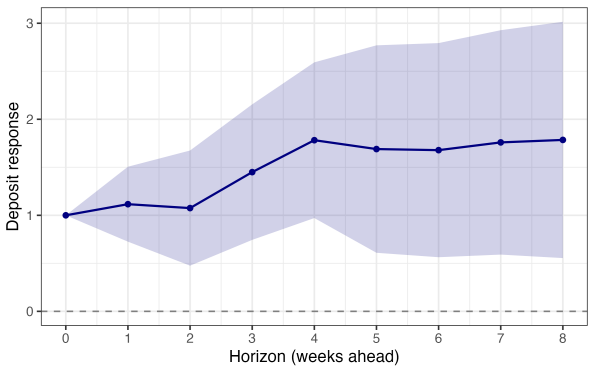
We next examine its time variation. We split our time sample into early and late periods, estimating the multiplier over the years 2005–2010 and 2011–2025. Panel (b) compares the dynamics of the deposit multiplier across the two samples. Comparing the magnitudes shows a marked decrease in the latter sample: the early-period response is 2.43, while the late-period response is 1.04. We summarize each response by its average at horizons five and six. The multiplier operates through successive rounds of lending and redepositing that play out over several weeks, and by horizon five the full-sample response has leveled off. Longer horizons allow slower forces outside the multiplier mechanism, such as shifts in deposit demand or real-side feedback, to enter the response.

We also use the same instrumental variables design to estimate the response of bank credit to deposit inflows, as reported in Appendix Table A.5 and Figure A.3. We find that bank credit rises following a deposit inflow and that, like the deposit response, it is markedly weaker in the late period. This parallel decline is consistent with the multiplier mechanism. Deposit inflows support new credit, and that credit in turn generates new deposits, so a weakening of either link dampens both the deposit and the credit response.

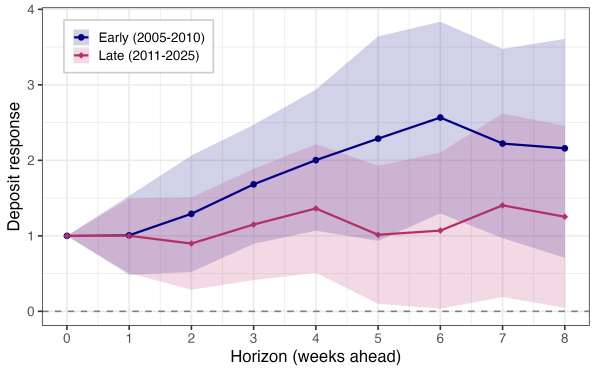
The reduced-form responses embed two distinct economic forces: a direct effect, captured by banks' marginal propensity to expand credit out of deposit inflows, and an indirect effect,

Figure 3: Deposit Multiplier Estimates

This figure reports local projection estimates of the dynamic response of deposits to a deposit shock at time t, estimated using an instrumental variables approach. Panel A shows the full sample estimate from 2005 to 2025. Panel B compares estimates for the early (2005–2010) and late (2011–2025) samples. Shaded bands show 95% confidence intervals based on Newey-West standard errors with lag h.



Panel A: Full Sample



Panel B: Early and Late Periods

captured by the share of credit proceeds that recycle back into deposit funding. Each is itself an equilibrium outcome, shaped by banks' capital and liquidity costs, competition from non-bank lenders, and households' portfolio choices. The reduced form tells us that the feedback loop has weakened, but not why: post-crisis regulation, central bank balance-sheet policy, the growth of non-bank intermediaries, and shifts in household behavior could all have contributed. Tracing the two forces to these primitives, and asking how credit creation responds as they change, requires additional structure, which we turn to next.

4 Model

4.1 Overview

The model features three types of agents: households, a representative banking system, and a representative non-bank credit intermediary.[7] Households are at once savers and borrowers: they hold liquid wealth, which they allocate across bank deposits, market savings, and cash, and they borrow from banks or non-banks to finance illiquid assets. The key exogenous variables are the

[7]We use "households" as shorthand for the non-bank private sector that holds financial claims: households, nonfinancial firms, and institutional investors.

current policy rate r^S, the equity capital of the banking sector E, and the level of central bank reserves R, which together characterize the state of the world. Given these variables, banks set deposit and loan rates to maximize profits, subject to capital and liquidity constraints, and the non-bank credit intermediary sets its lending rate to maximize profits. Finally, competitive funding markets clear the associated balance sheets. Money market funds allocate household market savings between short-term non-bank funding claims and safe assets, while a broader institutional sector supplies bank wholesale funding and the remaining market funding of non-bank credit.

We characterize the equilibrium in two layers. The first is a static portfolio equilibrium: taking household financial wealth as given, markets clear as a consistent allocation of financial claims, with the banking sector and the non-bank intermediary optimizing subject to their constraints. Within this layer there are no payment flows to trace: the equilibrium is a settled allocation of stocks where all markets clear. In particular, the lending response λ captures the banking sector's optimal reallocation of a marginal dollar of stable funding between new lending and reduced wholesale funding, given how the capital and liquidity slacks re-bind.

The second layer is the monetary-creation channel. Bank credit creates liquid wealth for the households who receive its proceeds, and they allocate it by the same portfolio choice that governs the first layer. The fraction of lending that returns as stable deposits, the recycling rate θ, is therefore not a separate primitive: it is the deposit share of the marginal recipients of credit proceeds, which depends on who those recipients are. Because we seek the marginal response to a shock rather than the full monetary equilibrium, money creation enters the model only through θ.

Our objects of interest, bank credit and total credit, are therefore governed by two objects: the lending response λ, evaluated at the equilibrium balance sheet, and the recycling rate θ. We recover λ structurally by differentiating the equilibrium, and use the data to discipline the monetary layer summarized by θ.

4.2 Households

4.2.1 Balance Sheet and Liquid Wealth

The household balance sheet is

$$W^d + \mathcal{A} \;=\; \mathcal{N} + \mathcal{B}, \tag{9}$$

where W^d is liquid wealth, $\mathcal{A}$ denotes illiquid assets (real assets and long-duration securities), $\mathcal{B}$ debt, and $\mathcal{N}$ net worth. Liquid wealth is the sum of bank deposits, market savings, and cash, $W^d = D + M + C$. It is a gross position: it can change while net worth does not.

Households are gross savers and gross borrowers at the same time: a household keeps liquid wealth while it carries debt, even though the debt costs more than the liquid wealth earns. Keeping a dollar liquid therefore has a cost. The dollar could have repaid debt or been invested in illiquid assets, so it forgoes the household's opportunity cost of funds, r^o, and earns instead the return on liquid wealth, r^W, the yields and liquidity services of the deposits, market savings, and cash it is allocated to. The difference, $r^o - r^W$, is the cost of carrying liquid wealth. Households bear it because liquid wealth is a buffer against payment needs, whose value is increasing and concave in total liquid wealth, as in money-in-the-utility-function and convenience-yield models (Sidrauski, 1967; Krishnamurthy and Vissing-Jorgensen, 2012; Nagel, 2016). A household holds liquid wealth up to the point where the value of the last dollar of buffer equals the cost of carrying it; under a Cobb-Douglas value of liquid and illiquid holdings, for example, it spends a fixed amount of interest on carrying each (Appendix A.5). Gross positions therefore matter, not only net worth, and desired liquid wealth rises when the cost of carry falls. The allocation of liquid wealth across deposits, market savings, and cash, to which we now turn, does not depend on its level, so in the static portfolio equilibrium we take liquid wealth as given at its observed level; the level matters when bank credit changes it, in Section 4.7.2.

4.2.2 The Allocation of Liquid Wealth

Households allocate liquid wealth across a menu $k \in \{D, M, C\}$ that includes a bank deposit D, market savings (e.g., money market fund shares) M, and cash C. These options differ in the yields they offer and the liquidity services they provide. Household i's utility from option k is

$$u_i^k \;=\; \alpha_d\, r^k + \xi_k + \varepsilon_i^k, \tag{10}$$

where α_d is the rate sensitivity of households and ε_i^k is the idiosyncratic taste shock for household i over option k. Bank deposits pay rate r^d and provide liquidity services ξ_D, market savings pay the policy rate r^S with liquidity services normalized to zero ($\xi_M = 0$), and cash provides liquidity services ξ_C but pays no interest ($r^C = 0$).

We assume that the idiosyncratic taste shock ε_i^k is i.i.d. Type-1 extreme value, the standard multinomial logit assumption, allowing for a closed-form solution of the choice probabilities. The share of liquid wealth held in option k is

$$s_k(r^d) \;=\; \frac{e^{\,\xi_k + \alpha_d r^k}}{e^{\,\xi_D + \alpha_d r^d} + e^{\,\alpha_d r^S} + e^{\,\xi_C}}, \qquad k \in \{D, M, C\}, \tag{11}$$

and we write $s_d \equiv s_D$ for the deposit share. The demand for each instrument is

$$X_k(r^d) \;=\; s_k(r^d) \cdot W^d, \qquad k \in \{D, M, C\}, \tag{12}$$

and because the shares sum to one, allocations satisfy the wealth identity

$$D + M + C \;=\; W^d. \tag{13}$$

Equation (13) closes the allocation of liquid wealth: every dollar is held in one of the three options, so any reallocation across categories must net to zero. Two properties of (11) are used below:

$$\frac{\partial D}{\partial W^d} \;=\; s_d, \qquad\qquad \frac{\partial D}{\partial r^d} \;=\; \alpha_d \, D \, (1 - s_d). \tag{14}$$

The first states that the deposit share does not depend on the level of liquid wealth, so the marginal dollar of liquid wealth is allocated like the average dollar. The second is the slope of deposit demand in the deposit rate.

4.2.3 Demand for Credit

The liability side of the household balance sheet is a demand for credit: a household borrows the amount $\mathcal{B} = W^d + \mathcal{A} - \mathcal{N}$ by which its desired holdings exceed its net worth, and this amount falls as the loan rate rises, since both desired liquid wealth and desired illiquid assets fall (Appendix A.5). To parameterize this demand and the choice between bank and non-bank credit, we follow the depositor side and use a discrete-choice formulation. There is a mass W^b

of potential borrowing. Each unit i is allocated by a discrete choice over a menu $k \in \{\ell, nb, \emptyset\}$ that includes bank credit ℓ, non-bank credit nb, and not borrowing $\emptyset$, with utility

$$u_i^k \;=\; -\alpha_\ell\, r^k + \xi_k + \eta_i^k, \tag{15}$$

where α_ℓ is the rate sensitivity of borrowers and η_i^k is the idiosyncratic taste shock over option k. Bank credit charges rate r^ℓ and provides non-rate services ξ_ℓ, non-bank credit charges rate r^{nb} and provides non-rate services ξ_{nb}, and the utility of not borrowing is normalized to zero ($r^\emptyset = \xi_\emptyset = 0$).

As with deposit demand, we assume that the idiosyncratic taste shock η_i^k is i.i.d. Type-1 extreme value. The market share of option $k \in \{\ell, nb, \emptyset\}$ is

$$s_k(r^\ell, r^{nb}) \;=\; \frac{e^{\xi_k - \alpha_\ell r^k}}{1 + e^{\xi_\ell - \alpha_\ell r^\ell} + e^{\xi_{nb} - \alpha_\ell r^{nb}}}, \tag{16}$$

and credit quantities are given by[8]

$$L_k(r^\ell, r^{nb}) \;=\; s_k(r^\ell, r^{nb}) \cdot W^b, \qquad k \in \{\ell, nb\}. \tag{17}$$

Because the logit shares sum to one, borrower allocations satisfy the identity

$$L + L_{nb} + L_\emptyset \;=\; W^b, \tag{18}$$

where $L_\emptyset = s_\emptyset(r^\ell, r^{nb}) \cdot W^b$ denotes potential borrowing that is not taken up at prevailing rates. Equation (18) is parallel to the wealth identity (13) on the savings side, with the understanding that the $\emptyset$ component reflects unrealized borrowing capacity rather than a claim held on any balance sheet. Household debt in (9) is the credit actually extended, $\mathcal{B} = L + L_{nb} = (1 - s_\emptyset)\, W^b$; the share $s_\emptyset$ that goes unused rises with the loan rate, which is how the downward-sloping demand for credit is parameterized.[9]

[8] For ease of notation we denote the banking sector's loan demand $L_\ell(r^\ell, r^{nb})$ by $L(r^\ell, r^{nb})$ in the remainder of the paper.

[9] Appendix A.5 derives the demand for credit implied by the household problem. Its semi-elasticity with respect to the loan rate is the counterpart of $\alpha_\ell(1 - s_\ell)$ here, so matching the two is a normalization on the curvature of the liquidity value and of the payoff from illiquid assets. The household problem also implies that credit demand rises with the deposit rate, because a higher return on liquid wealth lowers its cost of carry; the demand system here omits this term, which is the response of liquid wealth to the deposit rate from which we abstract.

4.3 Banking System

4.3.1 Profit Maximization

We model a representative banking sector that chooses the deposit rate r^d and loan rate r^ℓ to maximize profits. It takes as given the policy rate r^S, equity capital E, and the demand functions of depositors $D(r^d)$ and borrowers $L(r^\ell)$. The bank is a price-taker in the wholesale funding market, where it borrows at the policy rate r^S.[10] The banking sector also holds central bank reserves R, which we treat as an exogenous, system-level quantity. The balance sheet constraint is

$$L(r^\ell) + R = E + D(r^d) + W, \tag{19}$$

so that, for any (r^d, r^ℓ), wholesale funding is determined residually as

$$W = L(r^\ell) + R - E - D(r^d). \tag{20}$$

The banking sector's profit maximization problem is given by[11],

$$\max_{r^d, r^\ell} (r^\ell - r^S - \mu_\ell)L \; + \; (r^S - r^d - \mu_d)D \; + \; \xi \ln\big(K\big) \; + \; \tau \ln\big(\Lambda\big). \tag{21}$$

In (21), the first two terms capture profits from loan and deposit markets respectively, net of per-unit operating costs μ_ℓ and μ_d, which capture expenses such as borrower screening and servicing on the loan side and branch operations and payment services on the deposit side. Unlike these per-unit costs of intermediation, the remaining terms capture the costs associated with capital and liquidity management, which depend on the size and composition of the bank's balance sheet and which we describe in detail below.

We model the costs of capital and liquidity management with log penalties, which reduce profits at an increasing rate as the banks' capital or liquidity slack approaches zero. These penalties are a reduced-form representation of the banking sector's revealed preference for maintain-

[10] We use "wholesale funding" as shorthand for non-core funding, encompassing rate-sensitive or runnable liabilities. The economically relevant distinction in the model is the funding's stability: D denotes stable core deposits while W absorbs the remainder.

[11] We present the bank's profit maximization after substituting out the balance-sheet identity. Reserves earn the policy rate r^S and are financed at the margin by wholesale funding at the same rate, so reserve income nets out. The substitution places the opportunity cost r^S inside the loan and deposit margins and leaves a term $r^s E$, which is constant with equity fixed and does not affect the banking sector's choice of rates.

ing capital and liquidity buffers. This preference may stem from regulatory requirements; from internal risk management, such as avoiding costly recapitalization, default, or liquidity shortfalls; or from market discipline, as wholesale funding spreads widen when capital and liquidity positions weaken.

First, the bank faces a risk-invariant capital constraint captured by $\xi \ln\big(K\big)$. Let $\kappa \in [0,1]$ denote the banking sectors' minimum desired capital ratio. Define the capital slack K as

$$K \;=\; E - \kappa\,(L+R) \;=\; (1-\kappa)E \;-\; \kappa(D+W) \geq 0. \tag{22}$$

Thus, the capital slack K is the excess of equity capital E over a fixed fraction κ of total assets, with loans and reserves weighted uniformly. This risk-invariant leverage ratio resembles the Supplementary Leverage Ratio (SLR) implemented in the Basel III regulatory framework, though κ need not correspond to a regulatory minimum; it may equally capture an internally chosen target above any mandated level.[12] By rearranging with the balance sheet identity, we see that higher deposits D and wholesale funding W tighten the constraint, while greater equity E relaxes it. As the balance sheet grows relative to equity, the capital slack K shrinks, and the penalty $\xi \ln(K)$ reduces profits, making additional asset expansion increasingly costly.

Second, the bank faces a liquidity constraint captured by $\tau \ln\big(\Lambda\big)$. Let $\chi \in [0,1]$ denote the banking sectors' minimum desired ratio of liquid asset holdings relative to wholesale funding. Define the liquidity slack Λ as,

$$\Lambda = R - \chi W \geq 0 \tag{23}$$

Thus, the liquidity slack Λ is the excess of reserves R over a fraction χ of wholesale funding W. This formulation resembles the Liquidity Coverage Ratio (LCR) in Basel III, which requires banks to hold high-quality liquid assets against runnable short-term liabilities. More broadly, however, it captures banks' precautionary liquidity management against the withdrawal and rollover risk inherent in demandable and short-term funding, in the spirit of Diamond and Dybvig (1983). From (23), higher wholesale funding tightens the constraint, while greater reserves relax it. As wholesale funding rises without a matching increase in liquid assets, the liquidity slack Λ shrinks, and the penalty $\tau \ln(\Lambda)$ reduces profits, making reliance on unstable funding increasingly costly.

Importantly, these frictions link deposit-taking to loan supply. Since wholesale funding

[12] The model is qualitatively similar if we consider a risk-based capital requirement.

is the residual funding margin, a deposit inflow reduces wholesale borrowing dollar-for-dollar for a given balance sheet. The liquidity penalty is what makes this relevant for lending: because wholesale funding requires liquid backing while deposits do not, substituting deposits for wholesale funding relaxes the liquidity constraint and lowers the marginal cost of lending. Without the liquidity penalty, marginal deposit inflows would be irrelevant for the marginal cost of lending and thus would have no effect on lending behavior. The capital penalty and market power in deposit and loan markets govern the strength of this link.[13]

4.3.2 First Order Conditions

Given the profit maximization problem above, we next derive the corresponding first order conditions. Throughout, starred quantities denote values at the optimum, in particular $K^\star$ and $\Lambda^\star$ are the capital and liquidity slack evaluated at the equilibrium balance sheet implied by $(r^{d\star}, r^{\ell\star})$.

We begin with the condition for the loan rate $r^{\ell\star}$,

$$r^{\ell\star} = r^S + \mu_\ell + \frac{1}{\alpha_\ell(1 - s_\ell(r^{\ell\star}))} + \chi \frac{\tau}{\Lambda^\star} + \kappa \frac{\xi}{K^\star}. \tag{24}$$

This expression shows that the banking sector's optimal loan rate $r^{\ell\star}$ can be written as the sum of five terms. The first is the policy rate, which provides the baseline cost of funds in the economy. The second is the per-unit operating cost of bank lending. The third is a demand markup, reflecting the bank's market power arising from borrowers' preference for bank loans: when loan demand is less elastic, the bank sets a higher rate above the policy rate. The fourth is a liquidity wedge, which arises because extending loans must be funded with additional wholesale liabilities and therefore tightens the liquidity slack. The fifth is a capital wedge, which reflects the marginal cost of balance sheet expansion: when capital slack is scarce and $K^\star$ is small, the marginal cost of expanding assets rises, increasing lending rates. The term $\tau/\Lambda^\star$ can be interpreted as the shadow cost of liquidity, while $\xi/K^\star$ can be interpreted as the shadow cost of capital.

[13] Making these funding margins imperfect substitutes is a deliberate departure from Modigliani-Miller neutrality, and is well grounded for banks: equity is costly to raise, and deposits, which households hold for their safety and liquidity services, are more stable than wholesale funding. Because the penalties are reduced form representations of banks' revealed preferences for buffers, our analysis requires only that banks behave as if these margins are imperfect substitutes, not a stand on the deeper source of the friction.

Next, we solve for the optimal deposit rate $r^{d\star}$,

$$r^{d\star} \;=\; r^{S} \;-\; \mu_d \;-\; \frac{1}{\alpha_d\left(1 - s_d(r^{d\star} \mid r^{S})\right)} \;+\; \chi\frac{\tau}{\Lambda^{\star}}. \tag{25}$$

The deposit first order condition shows that the banking sector's optimal deposit rate $r^{d\star}$ is determined by four terms. The first is the policy rate, which provides the baseline return on funds. The second is the per-unit operating cost of deposit provision. The third is a deposit markdown, reflecting the bank's market power arising from depositors' preference for bank deposits: when deposit demand is less elastic, the bank can offer a lower rate below the policy rate. The fourth is a liquidity wedge, which raises the deposit rate banks are willing to pay: deposits displace wholesale funding and thereby relax banks' liquidity slack.

Taken together, equations (24) and (25) reveal a direct link between deposit taking and credit provision. Deposits reduce reliance on wholesale funding and relax the liquidity requirement. The banking sector's willingness to extend credit is therefore tied to its ability to attract deposits, providing the mechanism through which the credit-deposit multiplier operates.

4.4 Non-bank Credit Intermediary

We model a non-bank credit intermediary that sets the non-bank credit rate r^{nb} to maximize profits. It takes as given the policy rate r^S and borrower demand $L_{nb}(r^{nb})$. The credit intermediary earns a spread over the policy rate and faces a per-unit cost μ_{nb} which can capture, for example, costs related to funding or monitoring. In the baseline specification this cost is constant; in Section 6 we consider an extension that allows non-bank funding costs to vary with conditions in the short-term funding market. The non-bank's profit maximization problem is,

$$\max_{r^{nb}} \;\; \Pi_{nb} = (r^{nb} - r^{S} - \mu_{nb})\, L_{nb}(r^{nb}). \tag{26}$$

The optimal non-bank loan rate satisfies

$$r^{nb} = r^{S} + \mu_{nb} + \frac{1}{\alpha_{\ell}(1 - s_{nb})}. \tag{27}$$

This pricing condition is analogous to the bank loan pricing equation in Equation (24). The first term is the policy rate, which provides the baseline cost of funds in the economy. The second is

the per-unit cost of non-bank lending. The third is a demand markup, reflecting the non-bank's market power arising from borrowers' preference for non-bank credit: when loan demand is less elastic, the non-bank sets a higher rate above the policy rate.

Because the non-bank bears no state-dependent balance-sheet costs, the bank to non-bank spread reveals the shadow price of intermediating through the banking sector's balance sheet: it widens exactly when the bank's capital and liquidity slack are scarce. To the extent that bank's wedges reflect regulation in excess of banks' internalized economic costs, this spread is the price of regulatory arbitrage.

The non-bank intermediary plays two roles in the model. First, it disciplines the elasticity of bank credit demand in equilibrium: because borrowers substitute between bank and non-bank credit, the non-bank's cost structure shapes the residual demand curve facing banks and thereby influences banks' marginal lending response λ to deposit inflows. Second, it provides the margin along which we evaluate whether non-bank credit can offset the decline in the bank channel, as we explore in Section 6.

4.5 Money Markets and Institutional Market Funding

Market funding is supplied through two sectors. Households place market savings M with money market funds, which allocate these balances between short-term claims that finance non-bank credit, N, and a residual safe asset B_M. The money market balance sheet is therefore

$$M = N + B_M. \tag{28}$$

Alongside money markets, a broader institutional sector supplies market funding P. It allocates these funds across bank wholesale claims W, non-bank funding P_{nb}, and a residual safe asset B_P. Non-bank credit is financed jointly by money market and institutional claims. The corresponding balance sheets are

$$P = W + P_{nb} + B_P, \tag{29}$$

$$L_{nb} = N + P_{nb}. \tag{30}$$

Both funding sectors are competitive, and all claims earn the policy rate r^S. The residual safe assets are supplied elastically at that rate, so the allocation of market funding is pinned down by the funding needs of banks and non-banks. Bank wholesale funding W is determined residually by the bank balance sheet in (19), while non-bank lending L_{nb} is determined by the intermediary's optimization in (26). The two funding sectors supply these claims and place any remaining balances in the residual safe assets.

Combining the sectoral balance sheets gives aggregate market clearing,

$$M + P = W + L_{nb} + B, \qquad B \equiv B_M + B_P. \tag{31}$$

This structure separates household market savings from the broader pool of institutional funding while allowing both sectors to finance non-bank credit. In the baseline, the funding allocation clears balance sheets without introducing an additional pricing wedge. Section 6 extends the model by allowing money-market depth to affect the marginal cost of non-bank intermediation.

4.6 Equilibrium

An equilibrium consists of rates $\{r^{d\star}, r^{\ell\star}, r^{nb\star}\}$, quantities $\{D^\star, M^\star, C^\star, L^\star, L^\star_{nb}, W^\star\}$, and funding allocations $\{P^\star, N^\star, P^\star_{nb}, B^\star_M, B^\star_P\}$, given the exogenous state (r^S, E, R) and household liquid wealth W^d, such that:

1. Households allocate liquid wealth across bank deposits, market savings, and cash according to (11), yielding demands $D(r^{d\star})$, $M(r^{d\star})$, and $C(r^{d\star})$; their liquid wealth is consistent with (54) at the prevailing loan rate and asset returns.

2. Households borrow by choosing among bank loans, non-bank loans, and not borrowing according to (16), yielding loan demands $L(r^{\ell\star})$ and $L_{nb}(r^{nb\star})$.

3. The banking sector chooses $(r^{d\star}, r^{\ell\star})$ to maximize its profits given by (21), subject to its balance sheet identity in (19) and the definitions of capital slack K in (22) and liquidity slack Λ in (23), which residually determines its wholesale funding demand $W^\star$.

4. The non-bank credit intermediary sets $r^{nb\star}$ to maximize its profits given by (26).

5. All markets clear. Banks and non-banks accommodate deposit and credit demand at their chosen rates, and the money-market and institutional funding balances satisfy (28)–(30).

4.7 Recovering the Credit-Deposit Multiplier

With the equilibrium in hand, we can now address the central question of the paper: how much credit does a deposit inflow ultimately create? We develop the answer in three steps. The first two introduce the banking sector's marginal lending response λ and the recycling rate θ; the third combines them into the bank-credit response to a deposit inflow, and extends it to the total credit response by incorporating the non-bank sector.

4.7.1 The Marginal Lending Response λ

Deposit inflows can arise from several distinct shocks: a depositor preference shock ξ_D that shifts household allocation towards deposits, a reserve injection dR that expands household financial wealth, or a loan demand shock ξ_ℓ that raises lending and generates deposit growth through the recycling of payment flows. Each shock produces a marginal lending response, λ^0, which depends on what else moves alongside the deposit inflow. The subsequent amplification, however, is shock-independent: at every round beyond the first, the bank faces the same inflow of recycled stable deposits regardless of which shock initiated the chain.

We therefore focus on the marginal lending response to recycled deposits, $\lambda \equiv dL/dD$: how much additional credit the banking sector extends per dollar of stable deposit inflow arriving on its own, with reserves and household wealth held fixed, taking into account the induced reduction in wholesale funding and the re-binding of the capital and liquidity slacks. Because it governs every recycling round, λ is the object that determines amplification. Appendix A.2 derives λ, along with the shock-specific first-round responses λ^0, by differentiating the bank's balance sheet identity, the capital and liquidity slacks, and the bank and non-bank pricing conditions around an interior equilibrium. For the recycled-deposit case, this yields,

$$\lambda = \frac{dL}{dD} = \frac{\dfrac{\tau\chi^2}{\Lambda^2}}{\dfrac{1}{\tilde{\kappa}_\ell^{nb}} + \dfrac{\tau\chi^2}{\Lambda^2} + \dfrac{\xi\kappa^2}{K^2}}. \tag{32}$$

In Equation (32), L is total bank credit, loans and securities together, so λ measures the balance sheet expansion that a deposit inflow induces without regard to which asset the bank adds. This follows from the capital constraint, which weighs all assets uniformly in the spirit of the supplementary leverage ratio, so that the marginal cost of expansion is the same for a loan and for a security. The capital constraint thus pins down the size of the expansion but not its composition. Composition matters for recycling, because loan proceeds and securities purchases reach different households, and it enters through θ, to which we turn next. Appendix A.6 separates the loan and securities responses that add up to λ.

Several economic forces shape λ. First, liquidity regulation plays a dual role. Stable deposits relax the liquidity constraint by reducing reliance on wholesale funding, a benefit captured by $\tau\chi^2/\Lambda^2$ in the numerator. At the same time, expanding loans tightens the liquidity constraint, a cost captured by the same term in the denominator, so the net liquidity effect depends on how binding the constraint is at the margin. Second, competition from non-banks shapes λ through the coefficient $\tilde{\kappa}_\ell^{nb}$, which captures the residual loan-demand curvature facing banks after borrower substitution and non-bank repricing.[14] A larger $\tilde{\kappa}_\ell^{nb}$ means that banks face a more elastic residual loan-demand curve, so a relaxation of balance-sheet constraints translates into a larger lending response. Third, capital constraints discipline balance sheet expansion through $\xi\kappa^2/K^2$ in the denominator. When capital slack is limited, additional lending raises the shadow cost of leverage, reducing λ. Overall, λ reflects the joint interaction of liquidity regulation, capital regulation, and credit-market competition in determining how strongly recycled deposit growth translates into bank lending.

4.7.2 The Deposit Recycling Rate θ

The deposit recycling rate θ is the fraction of lending-generated payment flows that settles back into the banking system as stable deposits. Unlike λ, it is not a margin the banking system optimizes over: the proceeds of a loan may settle as stable deposits at other banks, or not return to the banking system at all, and the originating bank does not internalize where they settle, so θ does not enter its first-order conditions.[15] It is determined instead by what banks fund, which decides who receives the proceeds, and by the portfolio choices of the recipients. We derive it from the household block, discipline it with the reduced-form multiplier estimates, and use the

[14] The Appendix derives this term from the baseline non-bank pricing condition

[15] Appendix A.5 considers a bank that expects to recapture a share of the deposits its own lending creates.

derivation to interpret it.

Bank credit creates liquid wealth. Households hold the additional liquid wealth because the credit expansion lowers the cost of carrying it: the loan rate falls for the households that borrow, and the return on illiquid assets falls for the households that sell them. Appendix A.5 shows that these two prices clear the loan and asset markets, so that the increase in desired liquid wealth equals dL.[16] The deposit rate need not move for this to happen; it governs how the proceeds are allocated, and we hold it fixed in the baseline.

Recipients of credit proceeds need not have the same deposit share as the average household. Let households fall into two groups, $j \in \{H, I\}$: households and nonfinancial firms (H), and institutional investors (I). They share the demand system of Section 4.2.2 but may differ in the value they attach to deposit liquidity services, ξ_D^j, so that their deposit shares s_H and s_I need not be the same; the estimation uses only the aggregate share s_d, the liquid-wealth-weighted average of the two. Let $\nu \in [0, 1]$ denote the share of a marginal dollar of bank credit that is held, in equilibrium, as liquid wealth of group H:

$$dW_H^d = \nu\, dL, \qquad\qquad dW_I^d = (1-\nu)\, dL. \tag{33}$$

Aggregate deposit demand is $D = s_H\, W_H^d + s_I\, W_I^d$, so differentiating at an unchanged deposit rate, using (14),

$$\theta \;\equiv\; \left.\frac{\partial D}{\partial L}\right|_{r^d} \;=\; \nu\, s_H + (1-\nu)\, s_I : \tag{34}$$

the recycling rate is the deposit share of the marginal recipients of credit proceeds. Like λ, it is a derivative at the equilibrium balance sheet, the same in every round of the feedback loop. The remaining proceeds are held as market savings and cash:

$$dD = dD_0 + \theta\, dL, \tag{35}$$

$$dM + dC = (1-\theta)\, dL, \tag{36}$$

where dD_0 is a deposit inflow that originates outside the lending process, such as a preference shift toward deposits. In the baseline, M and C enter no pricing condition, so the bank's response to each round's recycled deposits is $dL = \lambda\, dD$; Section 6 lets non-bank costs depend on M, with ω_M the share of non-recycled proceeds held as market savings.

[16] QE enters the model in the same way, as an expansion of W^d (Section 6).

The recycling rate need not equal the aggregate deposit share: $\theta - s_d = (\nu - W_H^d/W^d)(s_H - s_I)$, so the two differ whenever credit proceeds are distributed across households differently from the stock of liquid wealth, as marginal and average propensities differ (Appendix A.5).

The share ν depends on what banks fund. Loan proceeds are spent and settle largely with other households and firms; securities purchases pay the seller, who is typically an institutional investor. With π the loan share of a marginal dollar of bank credit, and g_ℓ and g_s the shares of loan and securities proceeds that come to rest with group H,

$$\nu \;=\; \pi\, g_\ell + (1-\pi)\, g_s, \tag{37}$$

with $\nu = \pi$ in the simplest case, $g_\ell = 1$ and $g_s = 0$. In the baseline, λ is the total balance sheet expansion and does not distinguish securities from loans. Appendix A.6 derives loan- and securities-specific responses λ_ℓ and λ_S, with securities counting toward the liquidity buffer, and there $\pi = \lambda_\ell/\lambda$. Written by asset class, with $\theta_\ell \equiv g_\ell\, s_H + (1-g_\ell)\, s_I$ and $\theta_S \equiv g_s\, s_H + (1-g_s)\, s_I$ the deposit shares of the recipients of loan and securities proceeds,

$$\theta \;=\; \pi\, \theta_\ell + (1-\pi)\, \theta_S, \qquad\qquad \lambda\theta \;=\; \lambda_\ell\, \theta_\ell + \lambda_S\, \theta_S : \tag{38}$$

each asset contributes its own lending response times its own recycling rate to the feedback between deposits and credit, and the multiplier estimates discipline the product $\lambda\theta$.

Equation (34) isolates two reasons the recycling rate can change over time. With bars denoting averages across two periods,

$$\Delta\theta \;=\; \underbrace{\Delta\nu\left(\bar{s}_H - \bar{s}_I\right)}_{\text{composition}} \;+\; \underbrace{\bar{\nu}\,\Delta s_H + (1-\bar{\nu})\,\Delta s_I}_{\text{deposit specialness}}. \tag{39}$$

The first term is a change in who receives credit proceeds, such as a shift of bank credit from loans toward securities; the second is a change in the deposit shares of the recipients. Because the shares in (11) do not depend on the level of liquid wealth, neither does θ: an expansion of liquid wealth, such as QE, affects θ only through the identity of the marginal recipients. Section 5.4 uses this decomposition to interpret the estimated decline in θ.

Appendix A.5 relaxes two features of the baseline. If the deposit rate adjusts according to (25) as lending changes the bank's funding mix, λ is unchanged, because the deposit rate does

not enter the loan pricing condition (24), and the recycling rate becomes a weighted average of θ and one, as banks bid for deposits and pull a larger share of the proceeds back. The estimates are the same whether or not deposit rates moved; what changes is how much of measured recycling reflects who receives the proceeds and how much reflects pricing, which the estimates bound because θ cannot be negative. Since the propagation runs its course within weeks while deposit rates adjust slowly, we take the fixed-rate case as the baseline. Second, a bank that expects to recapture part of the deposits its own lending creates lends more in levels but responds less to deposit inflows, so the bank-credit response is smaller; the recapture is small for any one bank.

4.7.3 The Bank-Credit and Total-Credit Responses

Equation (35) and the lending responses of Section 4.7.1 form a two-equation system in total deposits and total lending: deposits equal the initial inflow plus the recycled share of lending, while lending responds with the shock-specific λ^0 to the initial inflow and with λ to recycled deposits at every subsequent round. Substituting one into the other gives the bank-credit response,

$$dL_{total}^{bank} = \lambda^0 \, m_D \, dD_0, \qquad m_D = \frac{1}{1 - \lambda\theta}, \tag{40}$$

where m_D is the deposit multiplier. The first-round lending response λ^0 governs how much credit the initiating inflow creates, while λ and θ jointly determine how strongly each round's lending sustains the next. For a pure recycled-deposit shock, $\lambda^0 = \lambda$.[17]

Total credit adds the induced response of non-bank credit, which in the baseline operates through borrower substitution alone: as bank lending expands and bank loan rates fall, borrowers shift away from non-bank credit. Let $\lambda_{nb,D}$ denote this non-bank response per dollar of recycled deposits. Equivalently, since each recycled dollar induces λ of bank lending, the non-bank response per dollar of bank lending is $\lambda_{nb,D}/\lambda^D$. Because the non-bank responds only to the bank loan rate, this applies to every round of induced bank lending, so the cumulative non-bank response is $(\lambda_{nb,D}/\lambda^D)dL_{\text{bank}}$, and the total-credit response is

$$dL_{\text{total}} = \left[1 + \frac{\lambda_{nb,D}}{\lambda}\right] \frac{\lambda^0}{1 - \theta\lambda} dD_0. \tag{41}$$

[17] In the baseline this also holds for a depositor preference shock: the accompanying outflow from market savings affects no pricing condition, so $\lambda_{\xi_D} = \lambda_D$. In the extension of Section 6, the two differ, as the outflow raises non-bank funding costs.

Because $\lambda_{nb,D} < 0$, total credit responds by less than bank credit: the deposit inflow reaches non-banks only through the fall in the bank loan rate, which crowds out non-bank lending. Section 6 introduces an additional channel by allowing non-bank costs to depend on short-term funding conditions, which are held fixed here.

One may ask whether movements in the residual safe claims B_M and B_P could affect the total-credit response. They could matter along two margins. First, trades in these assets could generate payment flows that settle as bank deposits; this margin is absorbed by measurement, since θ is disciplined by the observed propagation of deposit shocks and thus captures deposit creation through any channel. Second, changes in these claims could affect non-bank funding conditions; our assumption that non-bank funding is perfectly elastic shuts this channel, and the spillover exercise of Section 6 quantifies how large such effects would need to be to alter the aggregate conclusion.

5 Model Estimation

We estimate the model using aggregate U.S. banking data to recover the structural objects that govern the credit-deposit multiplier: the primitives underlying banks' marginal lending response λ, and the deposit recycling rate θ. The estimation combines capital, liquidity, and demand parameters measured directly from balance sheet and pricing data with deterministic minimum-distance moment conditions. The banking sector's deposit and loan pricing conditions discipline the shadow costs of illiquidity and leverage, while the non-bank pricing condition disciplines the average marginal cost of non-bank credit. We jointly target these pricing relationships and the reduced-form multiplier estimates from Section 3, which pin down the recycling rate of loan proceeds into stable deposits. The reduced-form bank-credit responses are not targeted and serve as a validation of the estimated model.

A useful feature of this approach is that the observed balance sheet quantities serve as inputs to the estimation rather than as targets of it. The model does not need to endogenously match the time-series levels of deposits, bank credit, liquid assets, money market fund assets, or non-bank credit. Instead, these balance sheet quantities are observed state variables that determine market shares, capital and liquidity slacks, and pricing wedges. Conditional on those observed states, the pricing moments discipline the shadow-cost parameters, while the reduced-

form multiplier moments discipline deposit recycling. The structural model is therefore used to interpret marginal pricing and propagation around observed balance sheets, rather than to explain the full evolution of financial-sector quantities.

5.1 Data and Inputs to Estimation

We assemble an aggregate quarterly panel for the U.S. commercial banking sector combining Federal Reserve balance sheet quantities with FDIC Quarterly Banking Profile interest flows. All quantities are expressed relative to nominal GDP.[18] The sample spans 1999Q1–2025Q4. We measure core deposits, loans, wholesale funding, and equity using aggregate balance sheet data. Deposit and loan rates are constructed from interest flows and stocks. The policy rate is the effective federal funds rate. On the depositor side, money market funds and currency serve as liquid outside options. On the borrower side, nonbank credit is measured using corporate debt securities from the Financial Accounts. Details of variable construction are provided in Appendix A.1.

Two groups of parameters serve as inputs to the estimation, as reported in Table 3. First, the balance-sheet-implied capital and liquidity ratios κ and χ define the pre- and post-reform regimes. Rather than imposing statutory values, we infer the banking sector's desired capital and liquidity ratios from observed balance sheet positions. In each regime $j \in \{pre, post\}$, let $b = 0.01$ denote a one-percentage-point ratio buffer. We set the capital ratio to the minimum observed capital ratio less this buffer,[19]

$$\kappa_j = \min_{t \in \mathcal{T}_j} \frac{E_t}{L_t + R_t} - b. \tag{42}$$

We set the liquidity ratio analogously, using the minimum observed reserve-to-wholesale-funding

[18] Scaling by nominal GDP renders quantities stationary and comparable across the sample. It also embed an assumption: τ and ξ are constant in GDP units, so a given slack provides the same comfort whenever it is the same share of the economy. This is natural, since the shocks that buffers absorb scale with the banking system.

[19] This follows the spirit of Begenau (2020), who disciplines regulatory requirements using observed bank balance-sheet ratios. The one-percentage-point buffer prevents the tightest observed balance sheet in a regime from having exactly zero slack; because it is expressed in ratio points rather than as a fixed level, the implied slack at the binding observation scales with the balance sheet, $b(L_t + R_t)$ for capital and bW_t for liquidity.

Table 3: Model Inputs

This table reports the inputs to model estimation. Panel A reports capital and liquidity ratios inferred from observed balance sheets in the early and late regimes. Panel B reports deposit and loan rate sensitivities.

Panel A: Balance Sheet-Implied Parameters

Symbol	Description	Pre-Reform	Post-Reform
κ	Capital requirement	0.027	0.083
χ	Liquidity ratio	0.368	0.704

Panel B: Rate Sensitivities

Symbol	Description	Value
α_d	Depositor rate sensitivity	1.270
α_ℓ	Borrower rate sensitivity	5.560

ratio less the same one-percentage-point buffer,

$$\chi_j = \min_{t \in \mathcal{T}_j} \frac{R_t}{W_t} - b. \tag{43}$$

This procedure pins the ratios to the tightest observed balance-sheet positions in each regime while allowing a small positive buffer at the binding observation. We plot the full quarterly series of the quantities K_t and Λ_t in Appendix Figure A.5.

Because the ratios are inferred from revealed balance-sheet behavior rather than imposed from statute, the regime shift in κ and χ absorbs not only regulatory reform but any structural change in the banking sector's desired buffers over the period. For example, consolidation that improves the average bank's access to wholesale funding would lower the sector's desired liquidity ratio and be reflected in χ_{post}. The estimated regimes should therefore be read as summarizing the banking sector's effective balance-sheet preferences in each era, whatever their source.

Second, the deposit and loan demand elasticities determine the demand markups that enter the pricing first-order conditions. These elasticities govern the responsiveness of deposits and loans to interest rates and therefore shape equilibrium funding pass-through and lending competition. We take these elasticities from the literature (Xiao, 2019; Diamond et al., 2024).[20]

[20]The exact values of these elasticities do not qualitatively affect our results.

Table 4: Targeted Moments and Parameter Discipline

This table summarizes the moments and variation that discipline each estimated parameter.

Object	Moment or variation	Economic role
τ, ξ	Bank deposit and loan pricing residuals, together with time-series variation in Λ_t, K_t, χ_t, and κ_t	Liquidity and leverage shadow-cost scales
$\theta_{pre}, \theta_{post}$	Early and late reduced-form deposit responses in excess of the initial inflow	Recycling of loan proceeds into stable deposits
μ_{nb}	Average residual non-bank pricing wedge after the demand markup	Non-bank marginal-cost component

5.2 Minimum-Distance Estimation and Parameter Estimates

Given these inputs, we estimate the parameter vector

$$\Theta = \{\, \tau,\ \xi,\ \theta_{pre},\ \theta_{post},\ \mu_{nb} \,\}, \tag{44}$$

by deterministic minimum distance, choosing parameters so that the model jointly matches the pricing first-order conditions and the reduced-form credit-deposit multipliers in the early and late periods. This follows the logic of structural banking estimation in which demand elasticities are disciplined outside the supply-side estimation, and supply-side primitives are recovered from equilibrium pricing and quantity moments (Wang et al., 2022); In our aggregate setting, identification comes from time-series variation in observed balance-sheet slacks, ratios, and spreads, together with the external multiplier moments, rather than from cross-sectional bank heterogeneity.

The shadow-cost parameters (τ, ξ) and the non-bank cost μ_{nb} are common across regimes; the recycling rates θ_{pre} and θ_{post} are regime-specific. Time variation in the liquidity and capital wedges, $\chi_t\tau/\Lambda_t$ and $\kappa_t\xi/K_t$, therefore comes only from measured sources: the slacks Λ_t and K_t, which evolve quarterly with bank balance sheets, and the ratios χ_t and κ_t, which vary in the pre- and post-period as inferred from balance sheet positions. Because the same (τ, ξ) must price both regimes, the regime-split pricing moments act as overidentifying restrictions: the estimated tightening of the wedges is identified by measured movements in slacks and ratios rather than by freely shifting parameters.

The estimation targets two sets of moments, summarized in Table 4: the banking and

non-bank pricing conditions, and the reduced-form multipliers. The banking sector's loan and deposit pricing first-order conditions discipline the shadow costs of liquidity τ and leverage ξ. The deposit pricing residual follows from the model's first order condition (25) directly,

$$\varepsilon_t^d(\Theta) = (r_t^d - r_t^S) + \mu_d + \frac{1}{\alpha_d(1 - s_t^d)} - \chi_t \frac{\tau}{\Lambda_t}. \tag{45}$$

The loan pricing residual adjusts the model's condition (24) for two features of the data. First, measured loan yields are contractual rates on risky credit, so we net out expected losses using the loan-loss provision rate p_t^ℓ; this measures the time-varying component of the per-unit lending cost μ_ℓ directly so that its cyclicality does not load onto the wedges. Second, measured bank credit L_t comprises loans and securities; we assume that securities are competitively priced and carry no default risk, thus both the markup and the expected-loss adjustment apply only to the loan component and are scaled by the loan share ω_t^ℓ. The loan pricing residual is thus,

$$\varepsilon_t^\ell(\Theta) = (r_t^\ell - r_t^S) - \omega_t^\ell p_t^\ell - \mu_\ell - \frac{\omega_t^\ell}{\alpha_\ell(1 - s_t^L)} - \chi_t \frac{\tau}{\Lambda_t} - \kappa_t \frac{\xi}{K_t}. \tag{46}$$

The baseline estimation then imposes $\mu_d = \mu_\ell = 0$ for the remaining operating-cost components, so that the pricing moments discipline τ and ξ directly. Any constant cost level absorbed by the shadow-cost parameters affects both regimes symmetrically and largely differences out of the cross-regime comparison that is our focus.

The baseline non-bank pricing condition disciplines μ_{nb}, the average per-unit cost of non-bank credit,

$$\varepsilon_t^{nb}(\Theta) = (r_t^{nb} - r_t^S) - \mu_{nb} - \frac{1}{\alpha_\ell(1 - s_t^{nb})}. \tag{47}$$

The identification is the mirror image of the bank side. Because the non-bank bears no balance-sheet wedges, its spread net of the demand markup identifies the cost level μ_{nb} directly. We later extend the model to consider how μ_{nb} may be affected by funding market conditions in Section 6.

The reduced-form multiplier estimates from Section 3 discipline the deposit recycling rates θ_{pre} and θ_{post}. Both moments are instances of the deposit multiplier, with the first-round response λ^0 set to match each period's Treasury settlement institutions. In the pre-period, settlements

move deposits without reserves, so $\lambda^0 = \lambda_t$ and the multiplier moment is,

$$g^m_{pre}(\Theta) = \hat{m}^{pre} - \frac{1}{T_{pre}} \sum_{t \in \mathcal{T}pre} \frac{1}{1 - \theta pre\, \lambda_t(\Theta)}. \tag{48}$$

In the post-period, after the Treasury General Account became a Federal Reserve account, settlements increase deposits and reserves together. The first round therefore propagates with the reserve-backed response λ^R_t, while the subsequent rounds of lending-generated deposits propagate with the recycled-deposit response λ_t,

$$g^m_{post}(\Theta) = \hat{m}^{post} - \frac{1}{T_{post}} \sum_{t \in \mathcal{T}post} \left(1 + \frac{\theta post, \lambda^R_t(\Theta)}{1 - \theta_{post}, \lambda_t(\Theta)}\right). \tag{49}$$

The minimum-distance objective stacks eight moment gaps: the pre- and post-period averages of the deposit, bank-loan, and non-bank pricing residuals, and the pre- and post-period multiplier gaps,

$$g(\Theta) = \left(\bar{\varepsilon}^d_{pre}, \bar{\varepsilon}^d_{post}, \bar{\varepsilon}^\ell_{pre}, \bar{\varepsilon}^\ell_{post}, \bar{\varepsilon}^{nb}_{pre}, \bar{\varepsilon}^{nb}_{post}, g^m_{pre}, g^m_{post}\right)'. \tag{50}$$

The estimator solves

$$\hat{\Theta} = \arg\min_{\Theta} g(\Theta)' W g(\Theta), \tag{51}$$

where we weight the moments with a matrix W that is fixed before optimization and block diagonal, with one block for the pricing moments and one for the deposit responses. Each pricing moment is weighted by the inverse of its long-run variance, estimated from preliminary residuals with a Newey-West correction for four quarterly lags. Each deposit target is weighted by the inverse of its variance, which we estimate from the influence functions of the IV local projections with a Newey-West correction for six weekly lags. This variance accounts for the correlation between the horizon-five and horizon-six coefficients that enter each target. Standard errors use the minimum-distance sandwich formula. Because the pricing moments and the deposit responses are estimated separately, we set the covariance between the two blocks to zero.

Table 5 compares each targeted moment with its model counterpart; the t-statistics divide each model-data gap by the moment's standard error. All eight targeted moments are fit within

Table 5: Targeted Moment Fit

This table compares the targeted moments and untargeted bank-credit responses with their model-implied counterparts. The deposit targets, m^D_{pre} and m^R_{post}, average horizons five and six; the superscript R denotes a reserve-backed initial inflow. The bank-credit responses, m^C_{pre} and m^C_{post}, provide untargeted validation. Each t-statistic is the model-data difference divided by the corresponding moment standard error. Pricing-moment standard errors use Newey-West covariances with four quarterly lags; response-moment standard errors use the covariance of the IV horizon average.

Moment	Actual	Model	t-stat
Pricing moments (annual percentage points)			
Deposit spread, pre	0.886	0.166	-0.819
Deposit spread, post	0.885	1.128	0.526
Bank-loan spread, pre	2.456	2.384	-0.095
Bank-loan spread, post	2.944	2.953	0.017
Non-bank spread, pre	3.155	2.697	-0.860
Non-bank spread, post	2.479	2.742	0.651
Reduced-form response moments (average horizons 5–6)			
m^D_{pre}	2.43	2.43	-0.000
m^R_{post}	1.04	1.04	0.000
Bank credit (untargeted)			
m^C_{pre}	1.91	2.14	0.276
m^C_{post}	0.15	0.48	1.038

one standard error.

We do not use the bank-credit responses in estimation, so they provide a validation check. The model counterpart of the bank-credit response to a one-dollar deposit inflow is the first-round lending response amplified by recycling,

$$m^C_j(\Theta) = \frac{1}{T_j} \sum_{t \in \mathcal{T}_j} \frac{\lambda^0_t(\Theta)}{1 - \theta_j \, \lambda_t(\Theta)}, \tag{52}$$

which we compare with the average bank-credit local-projection coefficient at horizons five and six: 1.91 in the early period and 0.15 in the late period. The comparison is informative because deposits and credit are linked by an accounting identity. Deposits rise by the initial inflow plus the recycled share of the induced credit, so the model deposit response equals $1+\theta_j \, m^C_j$. Because the deposit response is targeted, the credit response tests how the model divides the multiplier between banks' lending response and recycling. The model performs well on this untargeted margin. It implies bank-credit responses of 2.14 in the early period and 0.48 in the late period,

Table 6: Model Estimates

This table reports parameter estimates and model-implied responses for the early (2005–2010) and late (2011–2025) periods. Panel A reports the estimated parameters and the average lending response to recycled deposits. Parentheses report standard errors. Panel B reports the targeted deposit responses, and effects of an exogenous deposit inflow on bank and total credit. Total credit includes the induced non-bank response.

Panel A: Estimated Parameters

	Estimate	SE
Shadow costs		
τ Illiquidity	0.071	(0.022)
ξ Leverage	0.318	(0.076)
Deposit recycling		
θ_{pre}	0.665	(0.122)
θ_{post}	0.088	(0.937)
Recycled-deposit lending response		
λ^{rec}_{pre}	0.859	(0.031)
λ^{rec}_{post}	0.580	(0.072)

Panel B: Deposit and Credit Responses

	Model	Data
Credit-Deposit Multiplier		
m^D_{pre}	2.43	2.43
m^R_{post}	1.04	1.04
Recycled-Deposit Bank Credit		
Pre period	2.145	
Post period	0.618	
Effect on Total Credit		
Pre period	1.143	
Post period	0.349	

and it reproduces the sharp decline between the two.

5.3 The Total Bank-Credit Effect of Deposit Flows

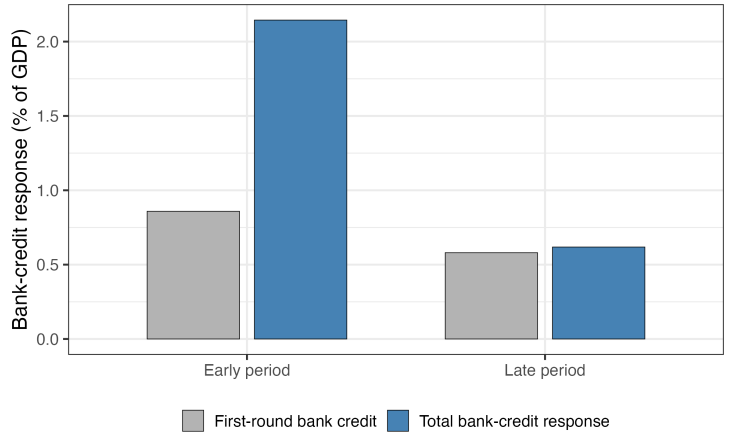
The estimated model decomposes the credit-deposit multiplier into its two constituent forces: banks' marginal lending response λ and the deposit recycling rate θ. With both in hand, we can now answer the central question of the paper: how much credit does a deposit inflow ultimately create, and how has this changed over the past two decades?

Table 6 reports the model-implied values for λ and θ. Banks' marginal lending response λ has declined from 0.86 to 0.58. This represents a decline of about a third in the share of each marginal deposit dollar that is channeled into new lending. The remainder is absorbed by reductions in wholesale funding. Deposits have moved from being primarily a source of credit creation to also substantially being a tool of balance sheet management.

The deposit recycling rate θ falls from 0.67 to 0.09, a large decline in the share of loan-generated payment flows that settle as stable deposits within the banking system. Section 5.4 uses the microfoundation of Section 4.7.2 to interpret this decline, separating a shift in who

Figure 4: First-Round and Total Bank-Credit Effects

This figure reports the first-round and total bank-credit effects of a deposit inflow equal to 1% of GDP in the early and late periods. The first-round effect corresponds to banks' initial lending response, $\lambda^D \Delta D_0$. The total effect incorporates full amplification through endogenous deposit re-depositing and multiplier dynamics, $\lambda^D \Delta D_0/(1 - \theta\lambda^D)$. Here λ^D is the lending response to recycled deposits, θ is the recycling rate, and ΔD_0 is the initial inflow.



receives credit proceeds, as bank credit moves from loans toward securities, from a decline in the deposit shares of the recipients themselves.

As a result, the total bank-credit response to a recycled deposit inflow, $\Delta L^{bank}_{total} = \lambda/(1 - \theta\lambda)$ per dollar of initial deposits, has declined from \$2.14 to \$0.62. Each dollar of deposits today generates less than a third of the bank credit it once did. This decline is driven both by the weaker lending response and diminished deposit recycling, which weakens the amplification from one round to the next. Figure 4 illustrates this by comparing the effect from banks' initial lending response to the total bank-credit effect following a one percent of GDP deposit inflow in the early and late periods. In the earlier period, the multiplier more than doubles the initial credit response. By the later period, amplification is much smaller. Taking the multiplier into account is crucial to understanding the magnitude of the decline in credit creation in the late period relative to the early period. Comparing only banks' marginal lending responses suggests a roughly one-third decline. But the total bank-credit effect has fallen by over 70 percent. The difference reflects the compounding nature of the multiplier: when θ weakens, the amplification mechanism contracts disproportionately.

5.4 What Drives the Decline in Recycling?

The microfoundation of Section 4.7.2 says where a decline in the recycling rate can come from. The recycling rate is the deposit share of the households who end up holding credit proceeds, $\theta = \nu s_H + (1 - \nu)s_I$, and (39) separates its change into a composition term, a shift in who receives the proceeds, and a deposit-specialness term, a change in the deposit shares of the recipients. Figure A.1 in Appendix A.1 reports evidence on each.

Panel A plots securities as a share of commercial bank credit. The share is higher in the late period than in the early period, consistent with post-crisis regulation pushing banks toward securities and away from loans (Stulz et al., 2022; Walz, 2026). In the benchmark of Section 4.7.2, in which loan proceeds settle with households and firms and securities proceeds with institutional investors, ν is the loan share of bank credit, so this shift lowers ν and, through the first term of (39), lowers θ in proportion to the gap $\bar{s}_H - \bar{s}_I$.

Panel B plots deposit holdings relative to financial assets for the real sector (households, nonprofits, and nonfinancial businesses) and for the investor sector (domestic financial institutions other than banks and the Federal Reserve). The two sectors differ sharply: the real sector holds several times as large a share of its financial assets in deposits as the investor sector, so s_I is well below s_H and routing proceeds toward institutional investors lowers θ. The real sector's deposit share has also declined in recent years, which lowers θ further through the second term of (39).

With s_I small, $\theta \approx \nu s_H$, so the two shifts compound. Both panels likely understate the relevant shifts, because they measure averages whereas θ depends on margins. Panel A measures the securities share of the stock of bank credit, whereas ν depends on the composition of the marginal expansion, which can be more securities-heavy than the stock (Appendix A.6). Panel B measures the deposit share of each sector as a whole, whereas θ depends on the deposit share of the households and firms within each sector that receive the proceeds, such as sellers of real assets, large corporate borrowers, and the non-bank financial institutions that banks lend to. A shift of the marginal dollar toward such recipients lowers θ.

6 Decomposing the Decline in Credit Creation

6.1 Drivers of the Decline in Bank-Credit Creation

We next use the model to study the drivers of the decline in the credit impact of deposit flows. We decompose the decline into two channels. The first varies the structural primitives governing bank behavior, including costs of capital and liquidity, and the extent of deposit recycling. The second varies balance-sheet quantities, asking how Quantitative Easing and Tightening affects the banking sector's capacity to intermediate deposits into credit. Together, the two exercises reveal that the decline reflects both an increase in banks' balance sheet costs and a decline in the specialness of deposits, as well as a transformation of the balance sheets themselves.

6.1.1 Changing Primitives: Capital, Liquidity, and Deposit Recycling

We decompose the early-to-late decline in the total bank-credit response, $\Delta L_{total}^{bank} = \lambda/(1-\theta\lambda)$, using regime-average counterfactuals in Figure 5. Starting from the late-period economy, we return each driver to its pre-period value on its own, holding the others at their late-period values: the capital wedge $\xi\kappa^2/K^2$, the liquidity wedge $\tau\chi^2/\Lambda^2$, and the recycling rate θ.[21] The late-period balance sheet is held fixed throughout.

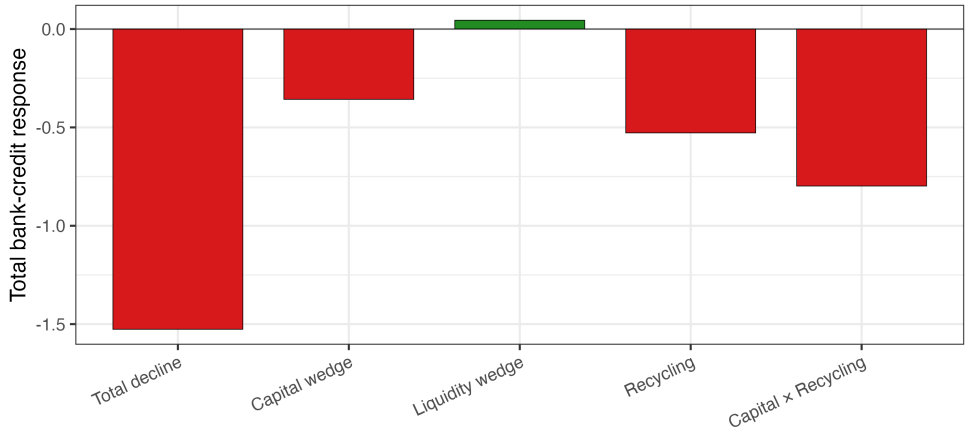
Because the total bank-credit response compounds the lending response and recycling, returning the capital wedge and the recycling rate together raises the response by more than the sum of returning each alone. Figure 5 reports this interaction as a separate bar, so that the bars sum to the effect of returning all three primitives to pre-period values.[22] The gap between this combined effect and the total decline is a residual reflecting the change in the balance sheet's own composition and scale, which we explore in the next subsection. Reduced recycling and the higher capital wedge are the two first-order forces. On their own, reduced recycling lowers the total bank-credit response by about \$0.53 per deposit dollar and the capital wedge by about

[21] Each late-period quarter's wedge is rescaled proportionally so that its geometric mean across quarters matches the pre-period geometric mean, preserving the late period's quarter-to-quarter variation. Matching geometric rather than arithmetic means keeps the quarters with the thinnest slack, whose wedges are largest, from dominating the comparison. We focus on the wedge terms rather than the scale parameters τ and ξ which are not interpretable in isolation: the economically relevant objects are the wedges, which combine the estimates with the observed ratios and slacks.

[22] The liquidity bar measures the further change from returning the liquidity wedge once capital and recycling have been returned, and so includes the small interactions of liquidity with the other two drivers.

Figure 5: Drivers of the Decline in Bank-Credit Creation

This figure reports the early-to-late change in the total bank-credit response, $\lambda^D/(1-\theta\lambda^D)$, where λ^D is the lending response to recycled deposits and θ is the recycling rate. The first bar shows the total change. The capital wedge and recycling bars hold the late-period balance sheet fixed and return one driver at a time to its early-period level, with the other drivers at late-period values. The capital $\times$ recycling bar is the additional change from returning both together, and the liquidity bar is the further change from then returning the liquidity wedge. These four bars sum to the effect of returning all three drivers.



$0.36, and their interaction lowers it by a further $0.80, the largest single component. The two forces compound: a weaker lending response leaves less credit to recycle, and weaker recycling leaves less funding to lend. The liquidity wedge works in the opposite direction, slightly offsetting the decline by increasing the liquidity benefit of the marginal deposit dollar.

6.1.2 Changing Quantities: Quantitative Easing and Tightening

We next turn from primitives to quantities. Two major post-crisis policies transformed the size and composition of bank balance sheets: Quantitative Easing (QE), which flooded the banking system with reserves, and Quantitative Tightening (QT), which drained reserves from the banking system. The exercises in this section ask how these balance-sheet changes alter the marginal bank-credit response to a new deposit dollar. They should therefore be distinguished from estimates of the direct effect of QE on the level of bank lending, such as Diamond et al. (2024). In the model, QE and QT operate on the same balance-sheet frictions in opposite directions. QE adds reserves and deposits together, relaxing the liquidity position but expanding the balance sheet against fixed equity, which tightens the capital position and lowers the marginal value of a deposit. QT can instead tighten the liquidity constraint by removing reserves without

a commensurate decline in deposit claims, raising the marginal value of stable funding.

We ask how QE changed the transmission of deposit inflows into bank credit. To answer this, we construct a counterfactual banking system in which QE-created reserves and the associated funding are removed, and compare the marginal bank-credit response to a dollar inflow across the observed and counterfactual balance sheets. Because balance-sheet policy injects money in reserve-backed form, we evaluate the response to such an inflow, with first-round response λ^R.

Constructing the counterfactual requires taking a stand on how QE-created reserves are funded. When the Federal Reserve purchases Treasury securities and MBS from non-bank counterparties, it creates reserves on bank balance sheets, and the sellers' bond holdings are replaced by claims on the banking system.[23] These QE-created balances land disproportionately with institutional holders, asset managers, pension funds, and dealers' clients, whose bank claims are rate-sensitive and functionally closer to wholesale funding than to the stable core deposits that relax banks' liquidity management (Blickle et al., 2025). In the model, the parameter ζ_{QE} governs this composition. In the baseline, we set $\zeta_{QE} = 0.5$, which means that half of the created wealth is held as core deposits, while the remainder appears on bank balance sheets as wholesale funding. QE therefore supplies reserves while diluting the funding mix toward flightier claims. The counterfactual removes QE-created reserves and the associated deposit and wholesale funding balances, holding bank and non-bank credit fixed, as detailed in Appendix A.1.

Figure 6 reports the result, which nets opposing balance-sheet margins. On the liquidity side, QE's reserves relax the liquidity position, though the wholesale funding QE arrives with demands liquid backing of its own, absorbing part of the relief. On the capital side, the balance-sheet expansion against fixed equity tightens the capital position. In our calibration the capital effect dominates: QE reduces the bank-credit response to a reserve-backed marginal dollar inflow by \$1.25.

We next ask how QT changes the transmission of deposit inflows into bank credit. Beginning in mid-2022, the Federal Reserve allowed its securities holdings to run off; as Treasury and MBS holdings mature without reinvestment, the Treasury reissues the debt to the market. Acharya and Rajan (2024) emphasize that this shrinkage can reduce bank-held reserves with-

[23] Household bond holdings lie outside the depositor menu by design: the menu $\{D, M, C\}$ models the allocation of liquid wealth, taking the division of portfolios between liquid and longer-duration assets as given. QE moves wealth across this boundary, converting bonds into money-like claims, and thus enters as an expansion of W^d.

out a commensurate reduction in deposit claims. This observation disciplines our exercises: the relevant question is how a smaller stock of liquid assets, relative to deposit claims, changes the marginal value of new stable deposit funding. We emphasize that these exercises do not measure the effect of QT on the level of bank lending; they ask whether, after the QT balance-sheet adjustment, an additional deposit dollar has greater multiplier power.

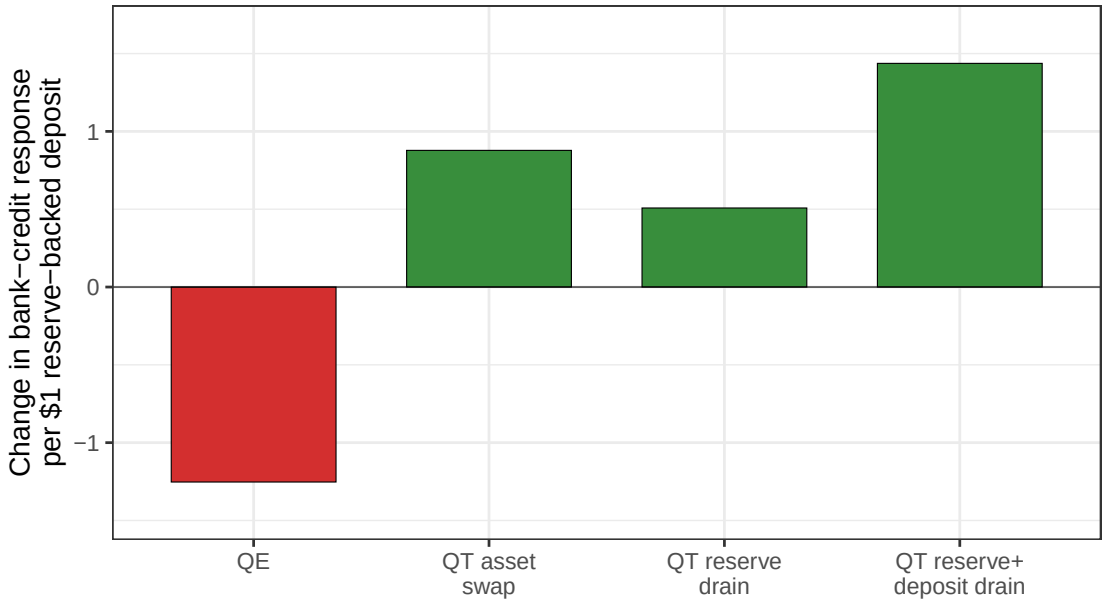
Because how QT's runoff lands on bank balance sheets is uncertain, we report three short-run counterfactuals, as detailed in Appendix A.1. The first treats QT as an asset swap: banks absorb one half of the runoff by exchanging reserves for less liquid securities, so total assets, deposits, wholesale funding, and credit are unchanged and only the liquidity buffer falls; isolating a pure liquidity tightening. The second is a reserve drain: bank credit and deposits are held fixed, reserves fall, and wholesale funding adjusts residually; in this case the reserve reduction deteriorates liquidity while the diminished reliance on wholesale funding reduces the liquidity burden of funding. The third is a reserve-and-deposit drain: reserves and deposits fall one-for-one with bank credit fixed. The second and third counterfactuals also shrink capital-relevant assets, so they combine the liquidity tightening with a loosening of the capital position.

Figure 6 reports the results. All three variants raise the marginal bank credit response: the asset swap by \$0.88, the short-run reserve drain by \$0.51, and the reserve-and-deposit drain by \$1.44. The ordering follows the mechanics. The asset swap is a pure liquidity tightening on the asset side. The reserve drain adds capital relief, but its liquidity tightening is partially offset by the accompanying decline in wholesale funding, which reduces the liquid backing the banking requires; on net it delivers the smallest effect. The reserve-and-deposit drain combines the full liquidity tightening, since wholesale funding is unchanged and no offset arises, with the capital relief, and therefore delivers the largest effect. Together with the QE result, the exercises show balance-sheet policy moving the same frictions in opposite directions: QE dampens the transmission of new deposits into credit, while QT amplifies it.

These exercises highlight that QE and QT operate through the same balance-sheet frictions in opposite directions. This carries a direct implication for the deposits channel of monetary policy in Drechsler et al. (2017): policy actions that operate through the supply or attractiveness of bank deposits have smaller credit effects in a QE-expanded banking system, where an additional deposit dollar commands less multiplier power. More broadly, the deposit multiplier is not a fixed structural constant but a policy-dependent object, lowered by balance-sheet expansion and partially restored by runoff.

Figure 6: Effects of QE and QT

This figure reports the change in bank credit per dollar of reserve-backed deposit inflow, $\lambda^0/(1-\theta\lambda^D)$, induced by QE and QT. Here λ^0 is the first-round lending response, λ^D is the response to recycled deposits, and θ is the recycling rate. The QE bar compares the observed late-period balance sheet with a no-QE counterfactual. The QT bars apply one half of the decline in Federal Reserve Treasury and MBS holdings from 2022Q2 to 2025Q4 to bank balance sheets. The asset swap holds total assets, credit, deposits, and wholesale funding fixed while reducing liquid assets. The reserve drain holds bank credit and deposits fixed; liquid assets other than bank credit fall, and wholesale funding adjusts residually. The reserve-and-deposit drain reduces reserves and deposits one-for-one, holding bank credit fixed.

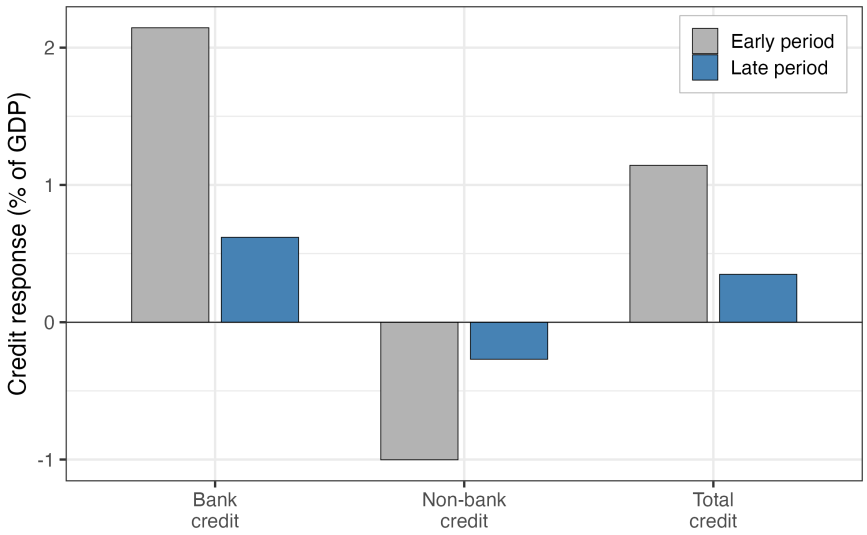


6.2 Implications for Aggregate Credit

The total effect of deposit inflows on bank credit has declined sharply, from \$2.14 to \$0.62. A natural question is whether non-bank credit intermediation, through bond markets, private credit, and other market-based channels, can compensate for this weakened deposit channel. We approach this question in two steps. We first examine the baseline equilibrium response of bank, non-bank, and aggregate credit to a deposit shock when non-bank marginal costs are unrelated to funding market conditions. We then ask how large the spillover from funds flowing into short-term funding markets to non-bank intermediation costs would need to be to restore the aggregate credit response to its early-period level.

Figure 7: Baseline Total Effect on Aggregate Credit

This figure reports the bank, non-bank, and total-credit responses to an ordinary deposit inflow equal to 1% of GDP in the early and late periods. Non-bank marginal costs are held fixed with respect to money-market depth, so the non-bank response reflects borrower substitution. Bars show response levels in each period.



6.2.1 Non-Bank Competitive Response

We begin with the direct margin: the equilibrium credit response to a deposit inflow when non-bank costs do not depend on funding market conditions. Figure 7 reports the baseline responses of bank, non-bank, and total credit to a deposit inflow of one percent of GDP in each period. Bank credit responses far less in the late period, and the total credit response falls by more than half. Non-bank credit contracts in both periods, the deposit inflow reaches non-banks only through the fall in bank loan rates as banks expand, which pulls borrowers away from non-bank credit.

This contraction clarifies why the competitive channel cannot close the gap between the early and late periods. Our motivating question is whether a more developed non-bank sector can deliver the credit creation that the bank channel no longer provides. Through borrower substitution alone, the answer is no, and the reason is one of sign: non-banks respond to a deposit inflow by contracting, so a more elastic or more substitutable non-bank sector deepens the contraction and widens the gap rather than narrowing it. For non-banks to compensate for the lost amplification, they must operate through a different margin, one that lowers their cost of intermediation as funds flow into short-term funding markets. We turn to that margin next.

6.2.2 How Large Must the Funding-Market Spillover Be?

We now ask how strong the spillover margin would need to be to matter. To do so, we extend the baseline model to let non-bank funding conditions depend specifically on money-market depth M. Money market funds use these balances to hold short-term non-bank funding claims N and safe assets B_M, so a deeper pool of money-market balances may lower the marginal cost of non-bank intermediation. We capture this effect with the cost schedule

$$\mu_{nb}(M) = \bar{\mu}_{nb} - \kappa_M M. \tag{53}$$

Here κ_M measures the sensitivity of non-bank marginal costs to money-market depth. When lending-generated payment flows do not recycle into stable deposits, let ω_M denote the share that enters money market funds. Evaluated at the no-spillover total-credit response, the resulting inflow is

$$dM^{sel} = \omega_M(1-\theta)dL^0_{\text{total}}, \tag{54}$$

where the superscript 0 denotes the baseline with $\kappa_M = 0$. Appendix A.2 derives the total-credit response with this leakage channel, combining the bank response, borrower substitution, and the effect of money-market depth on non-bank funding costs.

We use the extension to ask what cost sensitivity would restore the late-period aggregate credit response to its early-period level. The baseline total-credit response falls from \$1.14 in the early period to \$0.35 in the late period, leaving a gap of \$0.79 per dollar of initial deposits. Under full leakage, $\omega_M = 1$, closing this gap requires κ_M=2.3. Thus, a one-percent-of-GDP increase in money-market depth must lower non-bank marginal intermediation costs by about 2.3 basis points. With half leakage, $\omega_M = 0.5$, the required loading doubles to about 4.5 basis points. More generally, the required cost sensitivity scales inversely with the share of non-recycled proceeds that reaches money markets.

The full-leakage case is the most favorable configuration for the non-bank funding channel and therefore provides the smallest required value of κ_M. Even in this case, borrower substitution and improved funding conditions do not replace the lost bank amplification unless money-market inflows pass through materially to non-bank intermediation costs. More broadly, because non-bank intermediation chains are longer, neither recycled non-bank proceeds nor leaked bank flows

are likely to compensate for the loss in bank recycling. On the direct margin, non-banks contract rather than substitute; on the funding margin, they do not replace the multiplier.

7 Conclusion

The literature on banking has long debated whether deposits fund loans or loans create deposits, but has paid relatively little attention to the feedback loop between the two. This paper shows that this feedback loop is central to understanding the declining role of deposits in credit provision. A dollar of deposits today generates less than half of the credit it did two decades ago, and accounting for the amplification channel is crucial: focusing on banks' marginal lending response alone understates the decline in total credit creation. The diminished specialness of bank deposits means that fewer credit proceeds return as stable funding to support further credit creation. As recycling weakens, the compounding nature of the multiplier causes the amplification mechanism to contract.

Our findings point to an implicit narrowing of banking, driven not by institutional separation but by regulation and non-bank competition. As a result, both money and credit have become less responsive to deposit inflows, dampening the amplification of the credit cycle. These changes have implications for the transmission of monetary and fiscal policy, the cyclicality of credit, and the consequences of further migration of funds from banks to non-bank alternatives such as money market funds and stablecoins. Concerns about such migration may have more muted credit effects today than they would have had in earlier periods, but the weakening of the bank channel itself represents a meaningful and difficult-to-replace loss of amplification in the financial system.

References

Acharya, Viral V. and Raghuram Rajan, "Liquidity, Liquidity Everywhere, Not a Drop to Use: Why Flooding Banks with Central Bank Reserves May Not Expand Liquidity," *The Journal of Finance*, 2024, *79* (5), 2943–2991.

Begenau, Juliane, "Capital requirements, risk choice, and liquidity provision in a business-cycle model," *Journal of Financial Economics*, 2020, *136* (2), 355–378.

_ **and Tim Landvoigt**, "Financial regulation in a quantitative model of the modern banking system," *The Review of Economic Studies*, 2022, *89* (4), 1748–1784.

Bernanke, Ben S. and Alan S. Blinder, "Credit, Money, and Aggregate Demand," *The American Economic Review*, 1988, *78* (2), 435–439.

Bernanke, Ben S, Mark Gertler, and Simon Gilchrist, "The financial accelerator in a quantitative business cycle framework," *Handbook of macroeconomics*, 1999, *1*, 1341–1393.

Blickle, Kristian, Jian Li, Xu Lu, and Yiming Ma, "The dynamics of deposit flightiness and its impact on financial stability," Technical Report, National Bureau of Economic Research 2025.

Brunnermeier, Markus K and Yuliy Sannikov, "A macroeconomic model with a financial sector," *American Economic Review*, 2014, *104* (2), 379–421.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru, "The secular decline of bank balance sheet lending," Technical Report, National Bureau of Economic Research 2024.

Corbae, Dean and Pablo D'Erasmo, "Capital buffers in a quantitative model of banking industry dynamics," *Econometrica*, 2021, *89* (6), 2975–3023.

DeMarzo, Peter M, Arvind Krishnamurthy, and Stefan Nagel, "Interest rate risk in banking," Technical Report, National Bureau of Economic Research 2024.

Denbee, Edward, Christian Julliard, Ye Li, and Kathy Yuan, "Network risk and key players: A structural analysis of interbank liquidity," *Journal of Financial Economics*, 2021, *141* (3), 831–859.

Diamond, Douglas W and Philip H Dybvig, "Bank runs, deposit insurance, and liquidity," *Journal of political economy*, 1983, *91* (3), 401–419.

Diamond, William, Zhengyang Jiang, and Yiming Ma, "The reserve supply channel of unconventional monetary policy," *Journal of Financial Economics*, 2024, *159* (C).

Donaldson, Jason Roderick, Giorgia Piacentino, and Anjan Thakor, "Warehouse banking," *Journal of Financial Economics*, 2018, *129* (2), 250–267.

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl, "The Deposits Channel of Monetary Policy*," *The Quarterly Journal of Economics*, 05 2017, *132* (4), 1819–1876.

_ , _ , and _ , "Banking on deposits: Maturity transformation without interest rate risk," *The Journal of Finance*, 2021, *76* (3), 1091–1143.

Egan, Mark, Ali Hortaçsu, Nathan Kaplan, Aditya Sunderam, and Vincent Yao, "Dynamic Competition for Sleepy Deposits," *NBER Working Paper*, 2025, (w34267).

Friedman, Milton and Anna J. Schwartz, "Money and Business Cycles," *The Review of Economics and Statistics*, 1963, *45* (1), 32–64.

Gertler, Mark and Nobuhiro Kiyotaki, "Chapter 11 - Financial Intermediation and Credit Policy in Business Cycle Analysis," in Benjamin M. Friedman and Michael Woodford, eds., *Benjamin M. Friedman and Michael Woodford, eds.*, Vol. 3 of *Handbook of Monetary Economics*, Elsevier, 2010, pp. 547–599.

_ and Peter Karadi, "A model of unconventional monetary policy," *Journal of Monetary Economics*, 2011, *58* (1), 17–34. Carnegie-Rochester Conference Series on Public Policy: The Future of Central Banking April 16-17, 2010.

Greenwood, Robin, Samuel G. Hanson, and Jeremy C. Stein, "A Comparative-Advantage Approach to Government Debt Maturity," *Journal of Finance*, 2015, *70* (4), 1683–1722.

He, Zhiguo and Arvind Krishnamurthy, "Intermediary Asset Pricing," *American Economic Review*, April 2013, *103* (2), 732–70.

Holmstrom, Bengt and Jean Tirole, "Financial intermediation, loanable funds, and the real sector," *the Quarterly Journal of economics*, 1997, *112* (3), 663–691.

Jamilov, Rustam and Tommaso Monacelli, "Bewley banks," *Review of Economic Studies*, 2025, p. rdaf062.

Kashyap, Anil K., Raghuram Rajan, and Jeremy C. Stein, "Banks as Liquidity Providers: An Explanation for the Coexistence of Lending and Deposit-taking," *The Journal of Finance*, 2002, *57* (1), 33–73.

Kashyap, Anil K, Raghuram Rajan, and Jeremy C Stein, "Banks as liquidity providers: An explanation for the coexistence of lending and deposit-taking," *The Journal of finance*, 2002, *57* (1), 33–73.

Kiyotaki, Nobuhiro and John Moore, "Credit Cycles," *Journal of Political Economy*, 1997, *105* (2), 211–248.

Koont, Naz, Tano Santos, and Luigi Zingales, "Destabilizing digital" bank walks"," Technical Report, National Bureau of Economic Research 2024.

Kundu, Shohini, Seongjin Park, and Nishant Vats, "The Geography of Bank Deposits and the Origins of Aggregate Fluctuations," 2021. SSRN Working Paper. Available at SSRN: https://ssrn.com/abstract=3883605 or http://dx.doi.org/10.2139/ssrn.3883605.

Li, Lei, Elena Loutskina, and Philip E. Strahan, "Deposit market power, funding stability and long-term credit," *Journal of Monetary Economics*, 2023, *138*, 14–30.

Li, Ye and Yi Li, "Payment Risk and Bank Lending," *Working Paper*, 2022.

_ , _ , Huijun Sun, and Sergey Sarkisyan, "The Network Structure of Money Multiplier," *Working Paper*, 2026.

Lu, Xu, Yang Song, and Yao Zeng, "Tracing the impact of payment convenience on deposits: Evidence from depositor activeness," *Jacobs Levy Equity Management Center for Quantitative Financial Research Paper, The Wharton School Research Paper*, 2024.

Piazzesi, Monika and Martin Schneider, "Payments, credit and asset prices," 2018.

Schneider, Thomas, Philip Strahan, and Jun Yang, *Banking in the United States* 04

Stulz, René M, Alvaro G Taboada, and Mathijs A van Dijk, "Bank Liquid Assets, the Portfolio Motive, and Capital Requirements," Working Paper 30340, National Bureau of Economic Research August 2022.

Thakor, Anjan and G Yu Edison, "Funding liquidity creation by banks," *Journal of Financial Stability*, 2024, *73*, 101295.

Tobin, James, "Commercial banks as creators of 'money," 1963.

Walz, Stefan, "Bank Regulation as Interest Rate Policy," *Columbia Business School Research Paper*, 2026, (4741062).

Wang, Yifei, Toni M. Whited, Yufeng Wu, and Kairong Xiao, "Bank Market Power and Monetary Policy Transmission: Evidence from a Structural Estimation," *Journal of Finance*, 2022, *77* (4), 2093–2141.

Werner, Richard A, "Can banks individually create money out of nothing?—The theories and the empirical evidence," *International review of financial analysis*, 2014, *36*, 1–19.

Whited, Toni M, Yufeng Wu, and Kairong Xiao, "Will central bank digital currency disintermediate banks?," *Available at SSRN 4112644*, 2022.

Xiao, Kairong, "Monetary Transmission through Shadow Banks," *The Review of Financial Studies*, 10 2019, *33* (6), 2379–2420.

Appendix

Contents

A.1 Institutional Details and Additional Empirical Results **2**

A.1.1 Non-Bank Multiplier 2

A.2 The Household Ledger in the Simple Example 4

A.3 First-Stage Estimates 6

A.4 Full-Sample Estimates 7

A.5 Split-Sample Estimates 8

A.6 Credit Responses to Deposit Inflows 9

A.7 Bank Asset Composition and Sectoral Deposit Holdings 10

A.8 Instrument Local Projections 11

A.9 Full-Sample and Split-Sample Responses 12

A.2 Model Derivations **13**

A.1 Bank First Order Conditions 13

A.2 Local Pass-Through Derivations 14

A.3 Spillovers from Money Markets to Non-banks 20

A.4 Profits and Credit Creation at Longer Horizons 25

A.5 Microfoundation of the Recycling Rate: Derivations and Extensions 26

A.6 Asset Composition, Liquidity, and Deposit Recycling 35

A.3 Details on Model Calibration **40**

A.1 Data Construction 40

A.2 Demand System Details 41

A.3 Implied Balance Sheet Slack Variables 42

A.4 Additional Counterfactual Results **43**

A.1 Counterfactual Construction: QE and QT 43

A.2 Money-Market-to-Non-bank Spillover Calculation 44

A.1 Institutional Details and Additional Empirical Results

A.1.1 Non-Bank Multiplier

Having established the link between deposits and bank credit, and the forces that give rise to the credit-deposit multiplier, it is useful to consider whether an analogous multiplier operates for non-bank intermediaries.

Funding composition and credit provision. Like a bank, a non-bank intermediary's funding composition can shape its willingness to extend credit. More stable funding (longer-maturity debt, committed facilities, or equity-like claims) relaxes rollover and liquidity management frictions and supports additional lending at the margin. In this sense, the marginal lending response to a funding inflow is positive for non-banks as well, so that $\lambda_{nb} > 0$. The microfoundation differs from the bank case: non-banks typically do not face LCR-style liquidity regulation, and the relevant friction operates through funding costs and rollover risk rather than through a regulatory liquidity wedge, but the qualitative channel is the same.

Equilibrium determination of liabilities. Like a bank, a non-bank's liabilities are jointly determined in equilibrium by the supply of claims the intermediary is willing to issue and the demand for those claims from investors. Shifts on either side move both the quantity of non-bank funding and its price, just as in the bank deposit market. A non-bank sector facing an inflow of investor demand for its liabilities expands its balance sheet and extends additional credit, through the same logic that governs bank balance sheet expansion.

The settlement asset and deposit recycling. The key distinction between banks and non-banks lies in the recycling parameter θ. The credit-deposit feedback loop requires that loan-generated payment flows return to the issuing sector as stable funding. For banks, this recycling is substantial because bank deposits serve as the economy-wide settlement asset: nearly all payments in the real economy clear through deposits, so a large fraction of lending-generated spending mechanically returns to the banking system as deposits. Non-bank liabilities, by contrast, function as settlement media only within narrow subsystems (e.g., money market fund shares in institutional cash management or stablecoins in crypto markets). Payment flows that cross the

boundary of these subsystems leak into bank deposits and do not return as non-bank funding.

As a result, non-bank recycling rates are small in absolute terms, and the limited recycling that does occur within non-bank subsystems typically does not support long-term credit provision to the real economy. Money market funds hold short-dated instruments such as commercial paper and Treasuries rather than extending long-term credit. Stablecoin-based lending finances leveraged positions in other crypto assets rather than real investment, and tri-party repo channels credit to dealers and hedge funds for securities financing rather than to households and firms. Each non-bank subsystem either fails to extend the relevant type of credit, or channels it to leveraged financial intermediaries rather than to the real economy.

Non-bank Multiplier. For long-term credit provision to the real economy, the object of interest in this paper, we therefore treat $\theta_{nb} \approx 0$, so that the non-bank multiplier $m_{nb} = 1/(1 - \lambda_{nb}\theta_{nb}) \approx 1$. The credit-deposit feedback loop is a feature specific to the banking system's combination of settlement-asset liabilities and long-term lending to households and firms, and has no close non-bank analog along the dimensions that matter for aggregate credit creation. We therefore focus on banks' credit-deposit multiplier in the main text and abstract from non-bank multipliers.

A.2 The Household Ledger in the Simple Example

Table A.1 tracks the household side of the example in Table 1: the balance sheet of the household sector, which holds the bank's liabilities and owes its loans, and its decomposition into the borrower and the seller of the assets that the borrower purchases. Household liquid wealth consists of deposits D, payment balances PB, and money market fund shares M; the money market fund invests in the bank's wholesale funding, so that $M = W$, as is implicit in Table 1. Households also hold illiquid assets $\mathcal{A}$, in fixed supply ($300) and traded among households, and owe bank loans $\mathcal{B}$, so that net worth is $\mathcal{N} = D + PB + M + \mathcal{A} - \mathcal{B}$, as in Section 4.2.1. The bank has no equity, as in Table 1; with equity, bank shares would appear on the household side and cancel on consolidation.

In Table A.1, household net worth is $400 in every panel, equal to real assets plus bank reserves, because each claim created through lending is matched by a liability. Liquid wealth equals the size of the bank's balance sheet and rises with credit, from $100 to $170 and then to $214, and household debt rises with bank loans.

Table A.1: The Simple Example with Household Ledgers

The bank panels reproduce Table 1. Below each is the balance sheet of the household sector, which holds the bank's liabilities and owes its loans. Bank: R reserves, L loans, D stable deposits, PB payment balances, W wholesale funding. Households: D, PB, M money market fund shares, $\mathcal{A}$ illiquid assets, $\mathcal{B}$ bank loans, $\mathcal{N}$ net worth. Money market funds invest in the bank's wholesale funding, so that $M = W$; illiquid assets are in fixed supply and are traded among households. Round-2 figures are rounded.

(a) Initial

Banking system

Assets		**Liabilities**	
R	\$100	W	\$100
	\$100		\$100

Household sector

Assets		**Liabilities**	
M	\$100	$\mathcal{N}$	\$400
$\mathcal{A}$	\$300		
	\$400		\$400

(b) Deposit Inflow

Banking system

Assets		**Liabilities**	
R	\$100	D	\$100
		W	\$0
	\$100		\$100

Household sector

Assets		**Liabilities**	
D	\$100	$\mathcal{N}$	\$400
M	\$0		
$\mathcal{A}$	\$300		
	\$400		\$400

(c) Credit Creation

Banking system

Assets		**Liabilities**	
R	\$100	D	\$100
L	\$70	PB	\$70
	\$170		\$170

Household sector

Assets		**Liabilities**	
D	\$100	$\mathcal{B}$	\$70
PB	\$70	$\mathcal{N}$	\$400
$\mathcal{A}$	\$300		
	\$470		\$470

(d) Deposit Recycling

Banking system

Assets		**Liabilities**	
R	\$100	D	\$163
L	\$70	W	\$7
	\$170		\$170

Household sector

Assets		**Liabilities**	
D	\$163	$\mathcal{B}$	\$70
M	\$7	$\mathcal{N}$	\$400
$\mathcal{A}$	\$300		
	\$470		\$470

(e) Credit Creation

Banking system

Assets		**Liabilities**	
R	\$100	D	\$163
L	\$114	PB	\$44
		W	\$7
	\$214		\$214

Household sector

Assets		**Liabilities**	
D	\$163	$\mathcal{B}$	\$114
PB	\$44	$\mathcal{N}$	\$400
M	\$7		
$\mathcal{A}$	\$300		
	\$514		\$514

(f) Deposit Recycling

Banking system

Assets		**Liabilities**	
R	\$100	D	\$203
L	\$114	W	\$11
	\$214		\$214

Household sector

Assets		**Liabilities**	
D	\$203	$\mathcal{B}$	\$114
M	\$11	$\mathcal{N}$	\$400
$\mathcal{A}$	\$300		
	\$514		\$514

A.3 First-Stage Estimates

Table A.2: First Stage Estimates

This table reports first-stage estimates for weekly log growth of seasonally adjusted core deposits, defined as total deposits less large time deposits. Column (1) uses contemporaneous and lagged Treasury cash-flow innovations in the full sample (2005–2025), residualized over the full sample. Columns (2) and (3) report early-period (2005–2010) and late-period (2011–2025) estimates using both innovations, with Treasury dollar flows residualized separately within each period. Residualization includes week-of-year and quarter-year effects and lags 1–5. Winsorization cutoffs for outcomes and controls are computed over the full 2005–2025 sample. Treasury innovations are normalized by lagged core deposits; weekly deposit growth is winsorized at the 1st and 99th percentiles; Treasury innovations are retained without winsorization. All columns include contemporaneous changes in the weekly average effective federal funds rate and four deposit-growth lags. January weeks and October 2008 are excluded. All columns use regression origins with six observed outcome horizons; the last origin is October 29, 2025. Parentheses report conventional OLS standard errors; stars indicate significance based on these standard errors. The first-stage F-statistic jointly tests the excluded instruments in each column.

	(1) Full	(2) Early	(3) Late
Current shock	0.12^{***}	0.08	0.17^{***}
	(0.03)	(0.05)	(0.04)
Lagged shock	0.12^{***}	0.23^{***}	0.08^{**}
	(0.03)	(0.05)	(0.04)
Deposit lags	4	4	4
Rate change	Yes	Yes	Yes
F-Statistic	19.06	10.78	13.32
R^2	0.17	0.31	0.16
Num. obs.	948	240	707

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors in parentheses.

A.4 Full-Sample Estimates

Table A.3: Full-Sample Responses to Deposit Inflows

This table reports IV local projections for 2005–2025 at horizons $h = 1, \ldots, 8$ weeks. Panel A reports cumulative log core-deposit growth; Panel B reports cumulative bank-credit changes divided by lagged core deposits. Core deposits are seasonally adjusted total deposits less large time deposits. Weekly deposit growth is instrumented with contemporaneous and one-week-lagged Treasury cash-flow innovations normalized by lagged core deposits. Weekly deposit growth is winsorized at the 1st and 99th percentiles; Treasury innovations are retained without winsorization. Both equations control for contemporaneous changes in the weekly average effective federal funds rate and four deposit-growth lags. Bank credit additionally includes four lags of weekly credit changes divided by lagged core deposits; credit changes are adjusted by subtracting ten reporting-break amounts. January weeks and October 2008 are excluded. The last regression origin is October 29, 2025 (October 22 and October 15, 2025, at horizons seven and eight). Parentheses report Newey-West standard errors with lag h. The first-stage F-statistic jointly tests the two excluded instruments.

Panel A: Core Deposits

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$
Deposit Growth	1.12^{***}	1.08^{***}	1.45^{***}	1.78^{***}	1.69^{***}	1.68^{***}	1.76^{***}	1.79^{***}
	(0.20)	(0.31)	(0.36)	(0.41)	(0.55)	(0.57)	(0.60)	(0.63)
Deposit lags	4	4	4	4	4	4	4	4
Rate change	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Statistic	19.06	19.06	19.06	19.06	19.06	19.06	18.85	18.78
Num. obs.	948	948	948	948	948	948	947	946

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors in parentheses.

Panel B: Bank Credit

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$
Deposit Growth	0.23	0.17	0.42	0.90^{**}	1.28^{**}	1.34^{**}	1.43^{**}	1.85^{**}
	(0.21)	(0.26)	(0.32)	(0.43)	(0.56)	(0.60)	(0.68)	(0.73)
Deposit lags	4	4	4	4	4	4	4	4
Credit lags	4	4	4	4	4	4	4	4
Rate change	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Statistic	19.00	19.00	19.00	19.00	19.00	19.00	18.80	18.73
Num. obs.	948	948	948	948	948	948	947	946

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors in parentheses.

A.5 Split-Sample Estimates

Table A.4: Responses to Deposit Inflows: Core Deposits by Period

This table reports separate early-period (2005–2010) and late-period (2011–2025) IV local projections of cumulative log core-deposit growth at horizons $h = 1, \ldots, 8$ weeks. Core deposits are seasonally adjusted total deposits less large time deposits. Weekly deposit growth is instrumented with contemporaneous and lagged Treasury cash-flow innovations in both periods; innovations are normalized by lagged core deposits. Weekly deposit growth is winsorized at the 1st and 99th percentiles; Treasury innovations are retained without winsorization. Both periods control for four deposit-growth lags and contemporaneous changes in the weekly average effective federal funds rate. January weeks and October 2008 are excluded. The last late-period regression origin is October 29, 2025 (October 22 and October 15, 2025, at horizons seven and eight). Parentheses report Newey-West standard errors with lag h. The first-stage F-statistic jointly tests the excluded instruments in each period. Treasury cash-flow innovations are constructed separately within 2005–2010 and 2011–2025: dollar flows are residualized on week-of-year and quarter-year effects and lags 1–5, and residuals are retained without winsorization within each period before division by lagged core deposits. Winsorization cutoffs for outcomes and controls are computed over the full 2005–2025 sample. Calendar lags and leads are formed before regression-origin exclusions.

Panel A: Early Period (2005–2010)

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$
Deposit Growth	1.01^{***}	1.29^{***}	1.68^{***}	2.00^{***}	2.29^{***}	2.57^{***}	2.22^{***}	2.16^{***}
	(0.27)	(0.39)	(0.40)	(0.48)	(0.69)	(0.65)	(0.64)	(0.74)
Deposit lags	4	4	4	4	4	4	4	4
Rate change	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Statistic	10.78	10.78	10.78	10.78	10.78	10.78	10.78	10.78
Num. obs.	240	240	240	240	240	240	240	240

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors in parentheses.

Panel B: Late Period (2011–2025)

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$
Deposit Growth	1.00^{***}	0.90^{***}	1.15^{***}	1.36^{***}	1.01^{**}	1.07^{**}	1.40^{**}	1.25^{**}
	(0.25)	(0.31)	(0.38)	(0.43)	(0.47)	(0.53)	(0.62)	(0.62)
Deposit lags	4	4	4	4	4	4	4	4
Rate change	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Statistic	13.32	13.32	13.32	13.32	13.32	13.32	13.01	12.94
Num. obs.	707	707	707	707	707	707	706	705

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors in parentheses.

A.6 Credit Responses to Deposit Inflows

Table A.5: Responses to Deposit Inflows: Bank Credit by Period

This table reports separate early-period (2005–2010) and late-period (2011–2025) IV local projections of cumulative seasonally adjusted bank-credit changes divided by lagged core deposits at horizons $h = 1, \ldots, 8$ weeks. Core deposits are seasonally adjusted total deposits less large time deposits. Credit changes subtract ten reporting-break amounts, and controls include four lags of weekly credit changes divided by lagged core deposits. Weekly deposit growth is instrumented with contemporaneous and lagged Treasury cash-flow innovations in both periods; innovations are normalized by lagged core deposits. Weekly deposit growth is winsorized at the 1st and 99th percentiles; Treasury innovations are retained without winsorization. Both periods control for four deposit-growth lags and contemporaneous changes in the weekly average effective federal funds rate. January weeks and October 2008 are excluded. The last late-period regression origin is October 29, 2025 (October 22 and October 15, 2025, at horizons seven and eight). Parentheses report Newey-West standard errors with lag h. The first-stage F-statistic jointly tests the excluded instruments in each period. Treasury cash-flow innovations are constructed separately within 2005–2010 and 2011–2025: dollar flows are residualized on week-of-year and quarter-year effects and lags 1–5, and residuals are retained without winsorization within each period before division by lagged core deposits. Winsorization cutoffs for outcomes and controls are computed over the full 2005–2025 sample. Calendar lags and leads are formed before regression-origin exclusions.

Panel A: Early Period (2005–2010)

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$
Deposit Growth	0.58*	0.38	0.60	1.17*	1.76**	2.07**	2.16**	2.59***
	(0.31)	(0.36)	(0.44)	(0.61)	(0.83)	(0.86)	(0.84)	(0.95)
Deposit lags	4	4	4	4	4	4	4	4
Credit lags	4	4	4	4	4	4	4	4
Rate change	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Statistic	10.75	10.75	10.75	10.75	10.75	10.75	10.75	10.75
Num. obs.	240	240	240	240	240	240	240	240

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors in parentheses.

Panel B: Late Period (2011–2025)

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$
Deposit Growth	0.00	−0.10	−0.04	−0.01	0.13	0.17	0.07	0.05
	(0.18)	(0.23)	(0.25)	(0.28)	(0.30)	(0.33)	(0.37)	(0.41)
Deposit lags	4	4	4	4	4	4	4	4
Credit lags	4	4	4	4	4	4	4	4
Rate change	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Statistic	12.47	12.47	12.47	12.47	12.47	12.47	12.17	12.10
Num. obs.	707	707	707	707	707	707	706	705

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors in parentheses.

A.7 Bank Asset Composition and Sectoral Deposit Holdings

Figure A.1: Bank Asset Composition and Sectoral Deposit Holdings

Panel A shows securities as a share of commercial bank credit using weekly, seasonally adjusted H.8 data. Dashed lines mark period averages. Panel B shows deposit holdings relative to financial assets using quarterly Financial Accounts data. The real sector comprises households, nonprofit organizations, and nonfinancial businesses; the investor sector comprises domestic financial institutions other than the Federal Reserve and depository institutions. Deposit holdings include checkable deposits and currency plus time and savings deposits; money market fund checkable deposits are omitted. Shading marks 2005–2010 and 2011–2025.



Panel A: Securities Share of Bank Credit

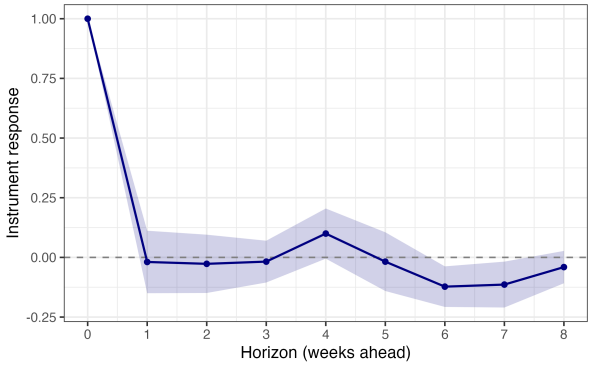


Panel B: Deposit Holdings Relative to Financial Assets by Sector

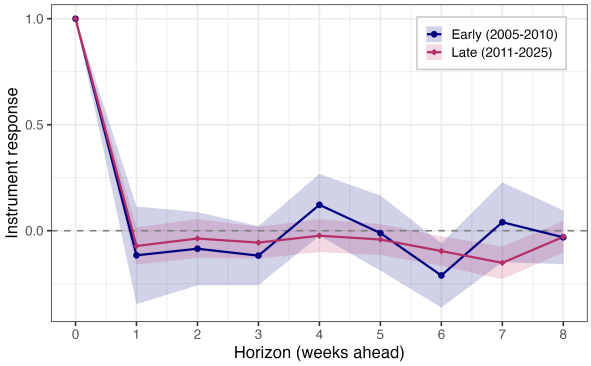
A.8 Instrument Local Projections

Figure A.2: Treasury Cash-Flow Instrument Local Projections

This figure reports OLS local projections of the Treasury cash-flow instrument Z_{t+h} on its contemporaneous value Z_t at horizons $h = 0, \ldots, 8$ weeks. Panel A shows the full sample (2005–2025); Panel B compares the early (2005–2010) and late (2011–2025) periods. Treasury dollar flows are residualized on week-of-year and quarter-year effects and lags 1–5; residuals are retained without winsorization and normalized by lagged core deposits. Full-sample innovations are residualized over the full sample; early and late innovations are constructed separately within each period. Regressions include the one-week-lagged instrument, four deposit-growth lags, and contemporaneous changes in the weekly average effective federal funds rate. Winsorization cutoffs for the deposit-growth controls are computed over the full 2005–2025 sample. Future instruments are taken from the uninterrupted weekly calendar before excluding January and October 2008 regression origins. Each sample uses common origins with all six future instrument observations available; horizons seven and eight use these origins with the longer lead observed. Horizon zero equals one by construction. Shaded bands show pointwise 95% confidence intervals based on Newey-West standard errors with lag h.



Panel A: Full Sample

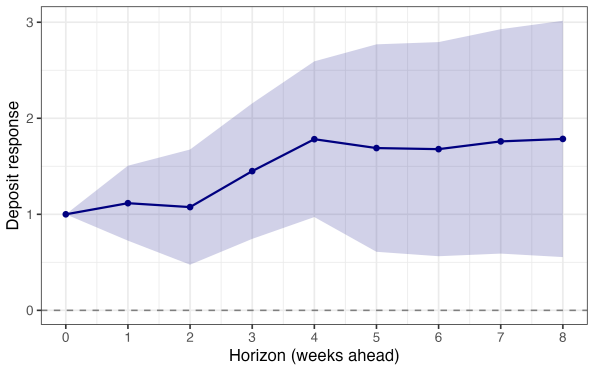


Panel B: Early and Late Periods

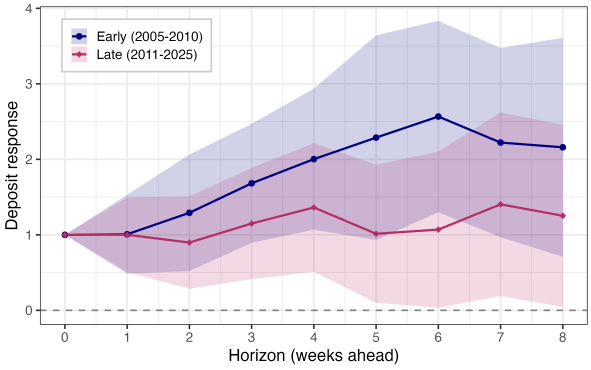
A.9 Full-Sample and Split-Sample Responses

Figure A.3: Responses to Deposit Inflows: Full and Split Samples

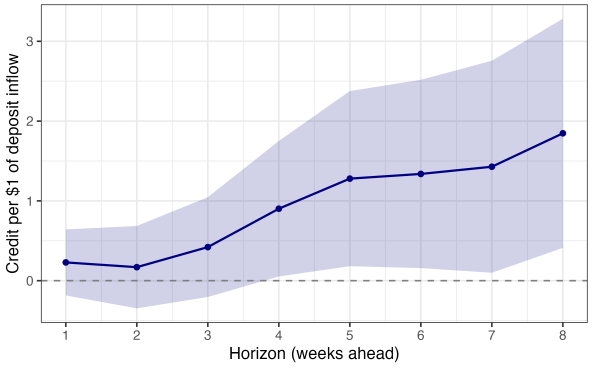
This figure reports IV local projections of deposit and bank-credit responses to a deposit inflow. Panels A and C show full-sample estimates for 2005–2025; panels B and D show separate early-period (2005–2010) and late-period (2011–2025) estimates. All regressions use contemporaneous and one-week-lagged Treasury cash-flow innovations normalized by lagged core deposits. Full-sample innovations are residualized over the full sample. Split-sample innovations are residualized separately within each period, matching the split-sample tables and model targets. Residualization includes week-of-year and quarter-year effects and lags 1–5; residuals are retained without winsorization. Deposits are cumulative log growth of seasonally adjusted core deposits (total deposits less large time deposits); bank credit is cumulative seasonally adjusted credit changes divided by lagged core deposits, with ten reporting-break amounts subtracted. Winsorization cutoffs for outcomes and controls are computed over the full 2005–2025 sample. All regressions control for four deposit-growth lags and contemporaneous changes in the weekly average effective federal funds rate; credit adds four lags of weekly credit changes divided by lagged core deposits. January weeks and October 2008 are excluded, and the last regression origin is October 29, 2025 (October 22 and October 15, 2025, at horizons seven and eight). Deposit responses at horizon zero equal one by normalization. Shaded bands show pointwise 95% confidence intervals based on Newey-West standard errors with lag h.



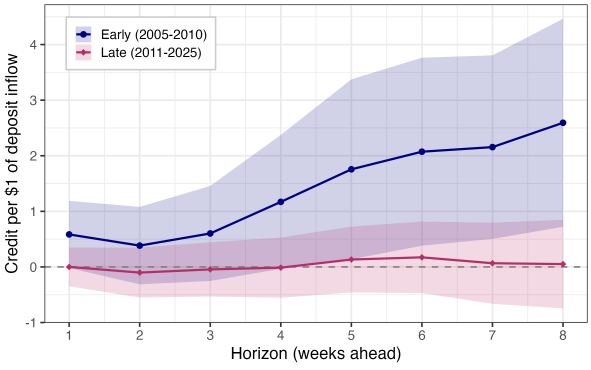
Panel A: Core Deposits, Full Sample



Panel B: Core Deposits, Early and Late Periods



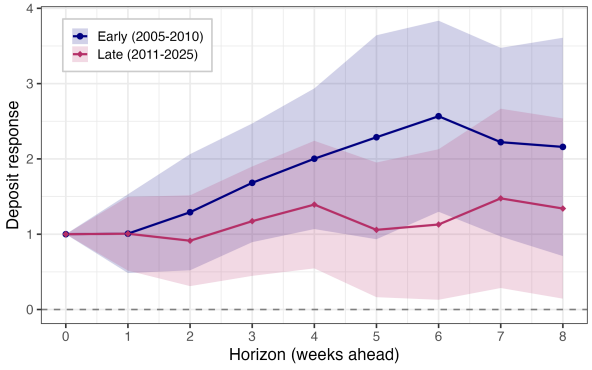
Panel C: Bank Credit, Full Sample



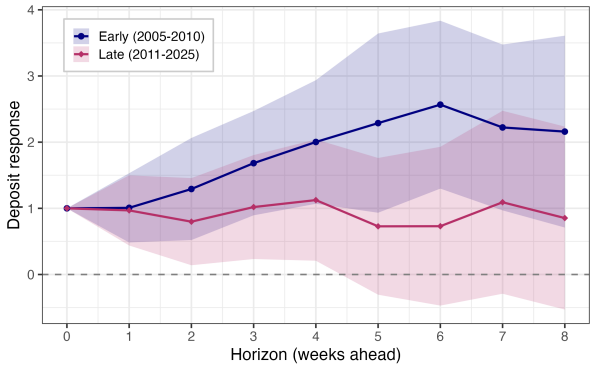
Panel D: Bank Credit, Early and Late Periods

Figure A.4: Deposit Multiplier Estimates: Robustness to Changes in Reserves

This figure reports early-period (2005–2010) and late-period (2011–2025) IV local projections of cumulative log core-deposit growth at horizons $h = 0, \ldots, 8$ weeks. The early-period estimates match Panel B of Figure 3. The late-period estimates additionally control for the one-week-lagged log level of aggregate reserve balances in Panel A and for contemporaneous weekly log growth of aggregate reserve balances, winsorized at the 1st and 99th percentiles, in Panel B. Weekly deposit growth is instrumented with contemporaneous and lagged Treasury cash-flow innovations residualized separately within each period and normalized by lagged core deposits. All regressions control for four deposit-growth lags and contemporaneous changes in the weekly average effective federal funds rate. January weeks and October 2008 are excluded. The last regression origin is October 29, 2025 (October 22 and October 15, 2025, at horizons seven and eight). Deposit responses at horizon zero equal one by normalization. Shaded bands show pointwise 95% confidence intervals based on Newey-West standard errors with lag h.



Panel A: Level of Reserves



Panel B: Changes in Reserves

A.2 Model Derivations

A.1 Bank First Order Conditions

The bank chooses loan and deposit rates taking reserves R and equity E as state variables. With wholesale funding determined residually by $W = L + R - E - D$, profits are

$$\Pi = (r^{\ell} - r^{S} - \mu_{\ell})L + (r^{S} - r^{d} - \mu_{d})D + \xi \log K + \tau \log \Lambda, \tag{1}$$

where

$$K = E - \kappa(L + R), \qquad \Lambda = R - \chi W = \chi E + (1 - \chi)R - \chi L + \chi D.$$

Using the logit demand derivatives

$$\frac{dL}{dr^{\ell}} = -\alpha^{\ell}(1 - s^{\ell})L, \qquad \frac{dD}{dr^{d}} = \alpha^{d}(1 - s^{d})D,$$

the loan and deposit first-order conditions are

$$r^{\ell} = r^{S} + \mu_{\ell} + \frac{1}{\alpha^{\ell}(1 - s^{\ell})} + \chi\frac{\tau}{\Lambda} + \kappa\frac{\xi}{K}, \tag{2}$$

$$r^{d} = r^{S} - \mu_{d} - \frac{1}{\alpha^{d}(1 - s^{d})} + \chi\frac{\tau}{\Lambda}. \tag{3}$$

These are the pricing equations used in estimation. The pass-through derivations below differentiate them around an interior equilibrium.

A.2 Local Pass-Through Derivations

This section derives the local pass-through objects used in the model by totally differentiating the equilibrium system around an interior optimum. We hold equity and policy rates fixed and determine wholesale funding residually from the bank balance sheet. Note that while the main text writes the recycled-deposit pass-through as λ, here we write it as λ_D to distinguish it from the first-round responses to other shocks.

We study small shocks around an interior optimum $x^* = (r_d^*, r_\ell^*, r_{nb}^*, W^*, D^*, L^*, L_{nb}^*, M^*)$. Given our functional forms, the objective and demand maps are C^1 on the domain $\{K > 0, \Lambda > 0\}$, and we assume two regularity conditions at x^*:

(i) **Interiority:** the optimum lies in the smooth region, with $K > 0$, $\Lambda > 0$, $W > 0$, and all shares strictly in $(0, 1)$;

(ii) **Nonsingular Jacobian:** the Jacobian of the stacked equilibrium system with respect to x is invertible.

Under (i)–(ii), the Implicit Function Theorem delivers a unique, differentiable local equilibrium map. Its C^1 smoothness permits total differentiation of each equilibrium identity, which we carry out in turn for the model's balance-sheet definitions and optimality conditions to obtain the linearized relations among $(d\Lambda, dK, dL, dD, dW, dL_{nb}, dM)$.

The equations to be differentiated are

$$L + R = E + D + W, \qquad \text{(balance sheet)} \tag{4}$$

$$\Lambda = R - \chi W, \qquad \text{(liquidity slack)} \tag{5}$$

$$K = E - \kappa(L + R), \qquad \text{(capital slack)} \tag{6}$$

$$r^{\ell} = r^{S} + \mu_{\ell} + \frac{1}{\alpha^{\ell}(1 - s_{\ell})} + \chi\frac{\tau}{\Lambda} + \kappa\frac{\xi}{K}, \qquad \text{(bank loan pricing)} \tag{7}$$

$$r^{nb} = r^{S} + \mu_{nb} + \frac{1}{\alpha^{\ell}(1 - s_{nb})}, \qquad \text{(non-bank pricing)} \tag{8}$$

(a) Balance sheet identity. Differentiating (4) gives

$$dW = dL + dR - dD. \tag{9}$$

Thus, a loan expansion must be financed by deposits, reserves, or wholesale funding through the balance sheet.

(b) Regulatory slacks. Differentiating (5) and (6), and substituting (9), gives

$$d\Lambda = (1 - \chi)dR - \chi dL + \chi dD, \qquad dK = -\kappa(dL + dR). \tag{10}$$

The corresponding differential of the regulatory part of the loan-pricing condition is

$$\begin{aligned} d\left(\chi\frac{\tau}{\Lambda} + \kappa\frac{\xi}{K}\right) = &\left(\frac{\tau\chi^2}{\Lambda^2} + \frac{\xi\kappa^2}{K^2}\right) dL - \frac{\tau\chi^2}{\Lambda^2} dD \\ &+ \left(\frac{\xi\kappa^2}{K^2} - \frac{\tau\chi(1 - \chi)}{\Lambda^2}\right) dR. \end{aligned} \tag{11}$$

The dL term is the marginal tightening from balance-sheet expansion. The dD term is the liquidity relief from replacing wholesale funding with deposits. The dR term is the net reserve effect: reserves relax the liquidity constraint but expand assets relevant for the capital constraint.

(c) Bank loan pricing. Differentiating the bank loan-pricing condition (7), holding r^S and μ_ℓ fixed, gives

$$dr^{\ell} = \frac{1}{\alpha^{\ell}(1 - s_{\ell})^2} ds_{\ell} + d\left(\chi\frac{\tau}{\Lambda} + \kappa\frac{\xi}{K}\right). \tag{12}$$

With primitive borrower demand shifters fixed, the logit share derivatives are

$$ds_\ell = -\alpha^\ell s_\ell(1 - s_\ell)dr^\ell + \alpha^\ell s_\ell s_{nb} dr^{nb}, \tag{13}$$

The corresponding bank loan quantity derivative is

$$dL = -\alpha^\ell L\big[(1 - s_\ell)dr^\ell - s_{nb} dr^{nb}\big]. \tag{14}$$

Substituting (13) into (12) yields

$$dr^\ell = -\frac{s_\ell}{1 - s_\ell} dr^\ell + \frac{s_\ell s_{nb}}{(1 - s_\ell)^2} dr^{nb} + d\left(\chi\frac{\tau}{\Lambda} + \kappa\frac{\xi}{K}\right), \tag{15}$$

or, collecting the dr^ℓ terms,

$$dr^\ell = \frac{s_\ell s_{nb}}{1 - s_\ell} dr^{nb} + (1 - s_\ell)d\left(\chi\frac{\tau}{\Lambda} + \kappa\frac{\xi}{K}\right). \tag{16}$$

(d) Non-bank pricing. Differentiating the non-bank pricing condition (8), holding r^S and μ_{nb} fixed, gives

$$dr^{nb} = \frac{1}{\alpha^\ell(1 - s_{nb})^2} ds_{nb}. \tag{17}$$

With primitive borrower demand shifters fixed, the non-bank share derivative is

$$ds_{nb} = \alpha^\ell s_{nb}\big[s_\ell dr^\ell - (1 - s_{nb})dr^{nb}\big]. \tag{18}$$

Substituting (18) into (17) gives

$$dr^{nb} = \frac{s_{nb}}{(1 - s_{nb})^2}\big[s_\ell dr^\ell - (1 - s_{nb})dr^{nb}\big].$$

Solving for the non-bank rate response,

$$\frac{dr^{nb}}{dr^\ell} = \frac{s_\ell s_{nb}}{1 - s_{nb}}. \tag{19}$$

(e) Residual loan-demand curvature. Combining bank loan demand (14) with (19) gives

$$dr^{\ell} = -\frac{1}{\alpha^{\ell} L\left[(1-s_{\ell}) - s_{nb}\frac{dr^{nb}}{dr^{\ell}}\right]} dL.$$

The bank loan-pricing condition equates the loan rate to the demand markup plus regulatory wedges. To isolate the movement in those wedges, we subtract the change in the demand markup from dr^{ℓ}. Since the markup is $1/[\alpha^{\ell}(1-s_{\ell})]$, its derivative along the same loan-demand curve is

$$d\left[\frac{1}{\alpha^{\ell}(1-s_{\ell})}\right] = \frac{s_{\ell}}{\alpha^{\ell} L(1-s_{\ell})^2} dL.$$

The left-hand side of the differentiated bank loan-pricing condition is therefore

$$\begin{aligned} dr^{\ell} - d\left[\frac{1}{\alpha^{\ell}(1-s_{\ell})}\right] &= -\left[\frac{1}{\alpha^{\ell} L\left((1-s_{\ell}) - s_{nb}\frac{dr^{nb}}{dr^{\ell}}\right)} + \frac{s_{\ell}}{\alpha^{\ell} L(1-s_{\ell})^2}\right] dL \\ &= -\frac{1}{\tilde{\kappa}_{\ell}^{nb}} dL. \end{aligned}$$

Equivalently,

$$\tilde{\kappa}_{\ell}^{nb} = \alpha^{\ell} L(1-s_{\ell})^2 \frac{(1-s_{\ell}) - s_{nb}\frac{dr^{nb}}{dr^{\ell}}}{(1-s_{\ell}) - s_{\ell} s_{nb}\frac{dr^{nb}}{dr^{\ell}}}. \tag{20}$$

$$dr^{\ell} - d\left[\frac{1}{\alpha^{\ell}(1-s_{\ell})}\right] = -\frac{1}{\tilde{\kappa}_{\ell}^{nb}} dL. \tag{21}$$

This is the effective residual loan-demand curvature facing banks after borrower substitution and non-bank repricing.

Using (21), the primitive loan-pricing condition becomes

$$-\frac{1}{\tilde{\kappa}_{\ell}^{nb}} dL = d\left(\chi\frac{\tau}{\Lambda} + \kappa\frac{\xi}{K}\right). \tag{22}$$

Substituting (11) into (22) gives

$$\begin{aligned} \left[1 + \tilde{\kappa}_{\ell}^{nb}\left(\frac{\tau\chi^2}{\Lambda^2} + \frac{\xi\kappa^2}{K^2}\right)\right] dL = \tilde{\kappa}_{\ell}^{nb}\frac{\tau\chi^2}{\Lambda^2} dD \\ + \tilde{\kappa}_{\ell}^{nb}\left(\frac{\tau\chi(1-\chi)}{\Lambda^2} - \frac{\xi\kappa^2}{K^2}\right) dR. \end{aligned} \tag{23}$$

The dD term is the liquidity relief from deposit funding. The dR term is the net reserve effect.

Recycled-deposit pass-through. For a recycled deposit inflow, reserves are fixed. Setting $dR = 0$ in (23) gives

$$\left[1 + \tilde{\kappa}_\ell^{nb}\left(\frac{\tau\chi^2}{\Lambda^2} + \frac{\xi\kappa^2}{K^2}\right)\right] dL = \tilde{\kappa}_\ell^{nb}\frac{\tau\chi^2}{\Lambda^2} dD.$$

Hence

$$\lambda_D \equiv \frac{dL}{dD} = \frac{\tilde{\kappa}_\ell^{nb}\frac{\tau\chi^2}{\Lambda^2}}{1 + \tilde{\kappa}_\ell^{nb}\left(\frac{\tau\chi^2}{\Lambda^2} + \frac{\xi\kappa^2}{K^2}\right)} = \frac{\frac{\tau\chi^2}{\Lambda^2}}{1/\tilde{\kappa}_\ell^{nb} + \frac{\tau\chi^2}{\Lambda^2} + \frac{\xi\kappa^2}{K^2}}. \tag{24}$$

With $dR = 0$, wholesale funding moves one-for-one with the part of the deposit inflow not used for bank credit:

$$\frac{dW}{dD} = \lambda_D - 1. \tag{25}$$

The associated non-bank credit response through borrower substitution follows from the same borrower-demand derivatives. Holding borrower demand shifters fixed,

$$dL = -\alpha^\ell L\left[(1 - s_\ell)dr^\ell - s_{nb}dr^{nb}\right],$$
$$dL_{nb} = \alpha^\ell L_{nb}\left[s_\ell dr^\ell - (1 - s_{nb})dr^{nb}\right].$$

Using the local non-bank repricing response (19), the bank-loan derivative implies

$$dr^\ell = -\frac{1}{\alpha^\ell L\left[(1 - s_\ell) - s_{nb}\frac{dr^{nb}}{dr^\ell}\right]} dL.$$

Substituting this expression into the non-bank quantity derivative gives

$$dL_{nb} = -\frac{L_{nb}}{L}\frac{s_\ell - (1 - s_{nb})\frac{dr^{nb}}{dr^\ell}}{(1 - s_\ell) - s_{nb}\frac{dr^{nb}}{dr^\ell}} dL.$$

For a marginal recycled-deposit inflow, $dL = \lambda_D dD$. Therefore

$$\lambda_{nb,D} \equiv \frac{\partial L_{nb}}{\partial D} = -\lambda_D\frac{L_{nb}}{L}\frac{s_\ell - (1 - s_{nb})\frac{dr^{nb}}{dr^\ell}}{(1 - s_\ell) - s_{nb}\frac{dr^{nb}}{dr^\ell}}. \tag{26}$$

This is the local non-bank response per dollar of realized recycled deposits. The total response

to an initial deposit impulse is obtained below after distinguishing the shock-specific first-round lending response from the recycled-deposit response that governs subsequent rounds.

Reserve-backed deposit shock. For a reserve-backed deposit injection, reserves move with deposits. Returning to (23) and dividing by the coefficient on dL gives

$$dL = \lambda_D dD - \frac{\frac{\xi\kappa^2}{K^2} - \frac{\tau\chi(1-\chi)}{\Lambda^2}}{1/\tilde{\kappa}_\ell^{nb} + \frac{\tau\chi^2}{\Lambda^2} + \frac{\xi\kappa^2}{K^2}} dR. \tag{27}$$

The second term is the net balance-sheet-expansion cost of reserves: the accompanying asset expansion tightens the capital constraint, while reserves partly offset this by relaxing the liquidity constraint. A reserve-backed deposit injection sets $dR = dD$, so the first-round reserve-backed lending response is

$$\lambda^R \equiv \left.\frac{dL}{dD}\right|_{dR=dD} = \lambda_D - \frac{\frac{\xi\kappa^2}{K^2} - \frac{\tau\chi(1-\chi)}{\Lambda^2}}{1/\tilde{\kappa}_\ell^{nb} + \frac{\tau\chi^2}{\Lambda^2} + \frac{\xi\kappa^2}{K^2}} = \frac{\frac{\tau\chi}{\Lambda^2} - \frac{\xi\kappa^2}{K^2}}{1/\tilde{\kappa}_\ell^{nb} + \frac{\tau\chi^2}{\Lambda^2} + \frac{\xi\kappa^2}{K^2}}. \tag{28}$$

This object is a first-round lending response to a deposit inflow that arrives with reserves. Unlike λ_D, it need not be positive: if the capital cost from the accompanying asset expansion is sufficiently large, $\xi\kappa^2/K^2 > \tau\chi/\Lambda^2$, then the numerator in (28) is negative and reserve-backed deposits reduce bank lending at the margin. It is distinct from the recycled-deposit response λ_D, which applies to later loan-generated deposits that do not come with new reserves.

First-round and recycled-round accounting. Let dD_0 denote the initiating deposit impulse, and let λ^0 denote the bank lending response that applies to that first impulse. For a pure recycled-deposit shock, $\lambda^0 = \lambda_D$. For a reserve-backed deposit shock, $\lambda^0 = \lambda^R$. Subsequent loan-generated deposits are ordinary recycled deposits, so those rounds propagate with λ_D, not necessarily with λ^0.

The bank-credit recursion is therefore

$$dL_{\text{tot}}^{\text{bank}} = \lambda^0 dD_0 + \lambda_D \theta dL_{\text{tot}}^{\text{bank}}. \tag{29}$$

Solving,

$$\frac{dL_{\text{tot}}^{\text{bank}}}{dD_0} = \frac{\lambda^0}{1 - \theta\lambda_D}. \tag{30}$$

The corresponding deposit multiplier is

$$\frac{dD_{\text{tot}}}{dD_0} = 1 + \theta \frac{dL_{\text{tot}}^{\text{bank}}}{dD_0} = 1 + \frac{\theta \lambda^0}{1 - \theta \lambda_D}. \tag{31}$$

When $\lambda^0 = \lambda_D$, this collapses to the usual $(1 - \theta\lambda_D)^{-1}$ multiplier. When the first round is reserve-backed, $\lambda^0 = \lambda^R$, giving the post-period multiplier target used in (49).

The non-bank borrower-substitution response scales with total bank credit, not directly with the initial deposit impulse. Using (26),

$$dL_{nb,\text{tot}}^{\text{sub}} = \frac{\lambda_{nb,D}}{\lambda_D} dL_{\text{tot}}^{\text{bank}} = \frac{\lambda_{nb,D}}{\lambda_D} \frac{\lambda^0}{1 - \theta \lambda_D} dD_0. \tag{32}$$

Thus, holding money-market depth fixed, the total aggregate-credit response to the initial deposit impulse is

$$\frac{dL_{\text{total}}}{dD_0} = \left(1 + \frac{\lambda_{nb,D}}{\lambda_D}\right) \frac{\lambda^0}{1 - \theta \lambda_D}. \tag{33}$$

A.3 Spillovers from Money Markets to Non-banks

The local pass-through derivation above uses the baseline non-bank pricing condition in (8). This subsection adds the spillover from money markets to non-banks by allowing short-term funding depth M to shift non-bank marginal intermediation costs directly. Let

$$\mu_{nb}(M) = \bar{\mu}_{nb} - \kappa_M M, \qquad \kappa_M > 0,$$

so that deeper money markets lower non-bank marginal costs. Taking M as given when setting the loan rate, the extended non-bank profit problem is

$$\max_{r^{nb}} \ \Pi_{nb}^M = (r^{nb} - r^S - \mu_{nb}(M)) L_{nb}(r^{nb}). \tag{34}$$

The first-order condition is

$$0 = L_{nb} + \left(r^{nb} - r^S - \mu_{nb}(M)\right) \frac{\partial L_{nb}}{\partial r^{nb}}. \tag{35}$$

Using the logit own-rate derivative

$$\frac{\partial L_{nb}}{\partial r^{nb}} = -\alpha^{\ell}(1 - s_{nb})L_{nb},$$

(35) gives the extended non-bank pricing condition:

$$r^{nb} = r^{S} + \mu_{nb}(M) + \frac{1}{\alpha^{\ell}(1 - s_{nb})} \tag{36}$$

Its differential is

$$dr^{nb} = \frac{1}{\alpha^{\ell}(1 - s_{nb})^{2}} ds_{nb} + \mu'_{nb}(M)dM. \tag{37}$$

The same borrower-substitution term moves both the non-bank share and non-bank loan quantity:

$$ds_{nb} = \alpha^{\ell} s_{nb}\big[s_{\ell} dr^{\ell} - (1 - s_{nb})dr^{nb}\big],$$
$$dL_{nb} = \alpha^{\ell} L_{nb}\big[s_{\ell} dr^{\ell} - (1 - s_{nb})dr^{nb}\big].$$

Substituting the share derivative into (37) gives

$$dr^{nb} = \frac{s_{nb}}{(1 - s_{nb})^{2}}\big[s_{\ell} dr^{\ell} - (1 - s_{nb})dr^{nb}\big] + \mu'_{nb}(M)dM.$$

Solving for the non-bank rate response,

$$dr^{nb} = \frac{s_{\ell} s_{nb}}{1 - s_{nb}} dr^{\ell} + (1 - s_{nb})\mu'_{nb}(M)dM = \frac{s_{\ell} s_{nb}}{1 - s_{nb}} dr^{\ell} - (1 - s_{nb})\kappa_{M} dM. \tag{38}$$

Thus the money-market spillover does not change the local non-bank repricing response used in the recycled-deposit pass-through derivation: when $dM = 0$,

$$\frac{dr^{nb}}{dr^{\ell}} = \frac{s_{\ell} s_{nb}}{1 - s_{nb}},$$

the same expression as in (19). The additional term in (38) matters only when the shock moves money-market depth.

Substituting (38) into the non-bank quantity derivative gives

$$dL_{nb} = \alpha^{\ell} L_{nb}(1 - s_{nb}) s_{\ell} dr^{\ell} - \alpha^{\ell} L_{nb}(1 - s_{nb})^{2} \mu'_{nb}(M)dM. \tag{39}$$

Holding the bank loan rate fixed, the direct non-bank response to money-market depth is there-

fore

$$\lambda_{nb,M}^{direct} \equiv \left.\frac{\partial L_{nb}}{\partial M}\right|_{dr^\ell=0} = -\alpha^\ell L_{nb}(1-s_{nb})^2\mu'_{nb}(M) = \alpha^\ell L_{nb}(1-s_{nb})^2\kappa_M. \tag{40}$$

Money-market depth also shifts bank credit through borrower substitution. Bank loan demand and (38) imply

$$dr^\ell = -\frac{1}{\alpha^\ell L\left[(1-s_\ell) - s_{nb}\frac{s_\ell s_{nb}}{1-s_{nb}}\right]}dL + \frac{s_{nb}(1-s_{nb})\mu'_{nb}(M)}{(1-s_\ell) - s_{nb}\frac{s_\ell s_{nb}}{1-s_{nb}}}dM.$$

Subtracting the bank loan-markup derivative, as in (21), and substituting into the differentiated loan-pricing condition yields

$$\begin{aligned}\left[1+\tilde{\kappa}_\ell^{nb}\left(\frac{\tau\chi^2}{\Lambda^2}+\frac{\xi\kappa^2}{K^2}\right)\right]dL = \tilde{\kappa}_\ell^{nb}\frac{\tau\chi^2}{\Lambda^2}dD \\ + \tilde{\kappa}_\ell^{nb}\left(\frac{\tau\chi(1-\chi)}{\Lambda^2}-\frac{\xi\kappa^2}{K^2}\right)dR \\ + \tilde{\kappa}_\ell^{nb}\frac{s_{nb}(1-s_{nb})\mu'_{nb}(M)}{(1-s_\ell) - s_{nb}\frac{s_\ell s_{nb}}{1-s_{nb}}}dM.\end{aligned} \tag{41}$$

The first two terms are the baseline recycled-deposit and reserve effects. The final term is the money-market spillover. Since $\mu'_{nb}(M) < 0$, an increase in M lowers non-bank loan rates and shifts some credit away from banks:

$$\lambda_{L,M} \equiv \frac{\partial L}{\partial M} = \frac{\tilde{\kappa}_\ell^{nb}\dfrac{s_{nb}(1-s_{nb})\mu'_{nb}(M)}{(1-s_\ell) - s_{nb}\dfrac{s_\ell s_{nb}}{1-s_{nb}}}}{1+\tilde{\kappa}_\ell^{nb}\left(\dfrac{\tau\chi^2}{\Lambda^2}+\dfrac{\xi\kappa^2}{K^2}\right)} = -\frac{\tilde{\kappa}_\ell^{nb}\dfrac{s_{nb}(1-s_{nb})\kappa_M}{(1-s_\ell) - s_{nb}\dfrac{s_\ell s_{nb}}{1-s_{nb}}}}{1+\tilde{\kappa}_\ell^{nb}\left(\dfrac{\tau\chi^2}{\Lambda^2}+\dfrac{\xi\kappa^2}{K^2}\right)}. \tag{42}$$

Aggregate-credit accounting. The bank response in (41) is only one part of the total-credit response. Substituting the bank loan-demand derivative into (39) gives the non-bank quantity response as a function of the bank-loan quantity and money-market depth:

$$dL_{nb} = \underbrace{-\frac{L_{nb}}{L}\frac{s_\ell(1-s_{nb})}{(1-s_\ell) - s_{nb}\frac{s_\ell s_{nb}}{1-s_{nb}}}}_{\text{borrower substitution per dollar of } dL}dL + \lambda_{nb,M}^{cost}dM, \tag{43}$$

The cost-channel coefficient is the coefficient on dM in (39) after holding fixed the bank-loan quantity dL. It has a direct non-bank marginal-cost piece and an indirect bank-rate piece:

$$\begin{aligned}\lambda^{cost}_{nb,M} &= -\alpha^{\ell} L_{nb}(1-s_{nb})^2 \mu'_{nb}(M) \quad + \alpha^{\ell} L_{nb}(1-s_{nb}) s_{\ell} \frac{s_{nb}(1-s_{nb})\mu'_{nb}(M)}{(1-s_{\ell}) - s_{nb}\frac{s_{\ell}s_{nb}}{1-s_{nb}}} \\ &= -\alpha^{\ell} L_{nb}(1-s_{nb})^2 \mu'_{nb}(M) \left[1 - \frac{s_{\ell}s_{nb}}{(1-s_{\ell}) - s_{nb}\frac{s_{\ell}s_{nb}}{1-s_{nb}}}\right].\end{aligned} \tag{44}$$

The first addend is the direct dM term in (39); the second substitutes the dM-component of dr^{ℓ} from the bank loan-demand derivative into the dr^{ℓ} term in (39). The first term in (43) is ordinary borrower substitution: when bank lending rises and bank loan rates fall, non-bank credit falls. The second term is the remaining money-market cost channel after this borrower-substitution response to dL has been separated out. With $\mu'_{nb}(M) = -\kappa_M$, this term is positive when deeper money markets lower non-bank marginal costs.

Combining (41) with (43), the local change in bank and non-bank credit, before closing the deposit-recycling loop, can be written as

$$\begin{aligned}dL + dL_{nb} &= \underbrace{dL}_{\text{bank credit}} + \underbrace{\frac{\lambda_{nb,D}}{\lambda_D} dL}_{\text{non-bank borrower substitution}} + \underbrace{\lambda^{cost}_{nb,M} dM}_{\text{money-market cost channel}} \\ &= \left(1 + \frac{\lambda_{nb,D}}{\lambda_D}\right)(\lambda_D dD + \lambda_{L,M} dM) + \lambda^{cost}_{nb,M} dM,\end{aligned} \tag{45}$$

where $\lambda_{nb,D}/\lambda_D$ is the borrower-substitution term from (26). The total-credit response to an initial deposit impulse is obtained only after adding the θ accounting below.

Loan-proceed leakage into money markets. The one-step expression in (45) takes dM as an external change in money-market depth. To keep the money-market spillover as a diagnostic extension rather than a second fixed point, we select the no-spillover total-credit response as the allocation that generates payment flows. From (33),

$$\frac{dL^0_{\text{total}}}{dD_0} = \left(1 + \frac{\lambda_{nb,D}}{\lambda_D}\right) \frac{\lambda^0}{1 - \theta\lambda_D}, \tag{46}$$

where the superscript 0 denotes $\kappa_M = 0$. Under full leakage of non-recycled proceeds into money-market funds, the selected money-market inflow is

$$dM^{sel} = (1-\theta)dL^0_{\text{total}}. \tag{47}$$

Evaluating the money-market spillover once around this selected allocation gives

$$\begin{aligned}\frac{dL^{spill}_{\text{total}}}{dD_0} &= \frac{dL^0_{\text{total}}}{dD_0} + \left[\left(1+\frac{\lambda_{nb,D}}{\lambda_D}\right)\lambda_{L,M} + \lambda^{cost}_{nb,M}\right]\frac{dM^{sel}}{dD_0} \\ &= \frac{dL^0_{\text{total}}}{dD_0}\left\{1+(1-\theta)\left[\left(1+\frac{\lambda_{nb,D}}{\lambda_D}\right)\lambda_{L,M} + \lambda^{cost}_{nb,M}\right]\right\}.\end{aligned} \tag{48}$$

This selected-allocation calculation accounts for the fact that non-bank credit also generates payment flows, while avoiding an additional money-market fixed point. If only a fraction of non-recycled proceeds enters money-market funds, replace $1-\theta$ in (47) and (48) by that fraction times $1-\theta$.

Deposit-preference shock. A pure preference shift from money market funds into deposits has $dR = 0$ and $dM = -dD$. From (41),

$$dL = \lambda_D dD + \lambda_{L,M} dM.$$

Setting $dM = -dD$ gives

$$\lambda_{\xi_D} \equiv \left.\frac{dL}{dD}\right|_{dR=0, dM=-dD} = \lambda_D - \lambda_{L,M}. \tag{49}$$

Because $\lambda_{L,M} < 0$, a shift out of money market funds and into deposits raises bank lending both through the usual deposit-liquidity channel and by increasing non-bank marginal costs. If reserves are fixed, $dW/dD = \lambda_{\xi_D} - 1$. For the corresponding local change in bank and non-bank credit, (45) gives

$$\left.\frac{dL + dL_{nb}}{dD}\right|_{dR=0, dM=-dD} = \left(1+\frac{\lambda_{nb,D}}{\lambda_D}\right)(\lambda_D - \lambda_{L,M}) - \lambda^{cost}_{nb,M}. \tag{50}$$

For a reserve injection that also changes depositor wealth, such as the QE counterfactual,

the imposed split determines dD, and money-market-fund assets absorb the remaining reserve-created wealth:

$$dM = dR - dD.$$

Equations (41) and (45) combine the local bank-credit response, non-bank borrower substitution, and the direct money-market cost channel at the resulting balance-sheet state. For loan-proceed spillovers, the leakage accounting in (48) additionally links the change in M to the non-recycled share of bank lending.

Wholesale funding. All wholesale-funding responses come from the balance sheet:

$$dW = dL + dR - dD.$$

Thus $dW/dD = \lambda_D - 1$ for recycled deposits, $dW/dD = \lambda_{\xi_D} - 1$ for deposit-preference shocks, and $dW/dD = \lambda^R$ for reserve-backed deposit injections with $dR = dD$.

A.4 Profits and Credit Creation at Longer Horizons

At longer horizons, the credit-creating capacity of deposit inflows is further strengthened by profits generated through lending activity. When borrowers use credit to finance production, they generate output Y and earn profits $\Pi = Y - wN - r^{\ell}L - r^{nb}L_{nb}$, while banks and nonbank lenders earn intermediation profits from interest spreads. These profits ultimately accrue to households as wages, dividends, and retained earnings, raising depositor wealth W_d and shifting out deposit demand. To the extent that bank profits are retained rather than distributed, they also raise bank equity E, relaxing the capital constraint and increasing the bank's willingness to expand its balance sheet.

The first channel operates through θ. Note that while the total pool of cash injected into the payment system by lending remains ΔL, higher household wealth W_d from positive profits shifts out deposit demand, increasing the fraction of circulating balances that households choose to hold as stable deposits, effectively raising θ. The second channel operates through λ. Retained bank profits raise equity E, which relaxes the capital constraint $K = E - \kappa(L + R)$ and lowers the shadow cost of balance sheet expansion, increasing the lending response λ in response to any given deposit inflow.

Both channels operate at business cycle frequency, since profits must be realized, distributed or retained, and reinvested before they affect deposit demand or bank equity, whereas the credit-deposit multiplier that we study propagates through bank balance sheets at a high frequency horizon over weeks. We therefore hold E fixed and evaluate θ at prevailing deposit demand conditions, so that the multiplier captures short-run balance sheet propagation in isolation.

The profit-driven amplification that we abstract from is closely related to the financial accelerator studied in the macro-finance literature (Bernanke et al., 1999; Gertler and Kiyotaki, 2010; He and Krishnamurthy, 2013), in which higher intermediary net worth relaxes borrowing constraints and enables further credit expansion. Our credit-deposit multiplier operates before and independently of this channel. At the same time, the two mechanisms would interact at longer horizons: profits raise depositor wealth and bank equity, which strengthen the credit-deposit feedback loop through both θ and λ, generating more lending and further profits.

A.5 Microfoundation of the Recycling Rate: Derivations and Extensions

This appendix states the household problem behind the liquid-wealth block of Section 4.2, derives the results of Section 4.7.2, and develops two extensions: adjustment of the deposit rate, and internalization of recycling by banks. Throughout, $j \in \{H, I\}$ indexes the two groups of households; the representative household of Section 4.2 is the case of a single group. We write θ for the baseline recycling rate, which holds the deposit rate fixed, and $\tilde{\theta}$ for the recycling rate when the deposit rate adjusts.

The household problem. A household of group j chooses both sides of its balance sheet (9), solving

$$\max_{W_j^d,\, \mathcal{A}_j,\, \mathcal{B}_j \geq 0} \; r_j^W W_j^d + v_j(W_j^d) + F_j(\mathcal{A}_j) - r^\ell \mathcal{B}_j \qquad \text{subject to } W_j^d + \mathcal{A}_j = \mathcal{N}_j + \mathcal{B}_j. \quad (51)$$

The first term is what liquid wealth earns once allocated across deposits, market savings, and cash according to (11). Its per-dollar return is the inclusive value of that allocation,

$$r_j^W \;\equiv\; \frac{1}{\alpha_d} \ln\left(e^{\xi_D^j + \alpha_d r^d} + e^{\alpha_d r^S} + e^{\xi_C}\right), \quad (52)$$

which combines the yields of the three instruments with their liquidity services and with the idiosyncratic tastes in (10) (Anderson, de Palma, and Thisse, 1988). Because it is linear in W_j^d, the choice of liquid wealth separates from its allocation, and the shares in (11) hold at any level of liquid wealth. The second term, v_j, is the value of liquid wealth as a buffer against payment needs, increasing and strictly concave, in the same way that the log penalty in (21) captures the declining value of the banking sector's liquidity slack. The third term, F_j, is the concave payoff from illiquid assets: a household values additional housing and durables less at the margin, and a firm funds its most productive projects first.

Let r_j^o denote the Lagrange multiplier on the balance-sheet constraint, the value to the household of a dollar of funds. The first-order conditions are

$$r_j^W + v_j'(W_j^d) = r_j^o, \qquad F_j'(\mathcal{A}_j) = r_j^o, \qquad r^\ell \geq r_j^o \text{ with equality if } \mathcal{B}_j > 0, \tag{53}$$

so r_j^o is the household's opportunity cost of funds: the loan rate for a household that borrows, and the marginal return on illiquid assets, which is then no higher than the loan rate, for one that does not. The first condition is the choice of liquid wealth,

$$v_j'(W_j^d) \;=\; r_j^o - r_j^W. \tag{54}$$

The right side is the cost of carrying a dollar of liquid wealth: what the dollar would earn or save if it were not kept liquid, less what it earns as liquid wealth. The household adds to its buffer until the value of the last dollar equals this cost. Because v_j is strictly concave, (54) pins down liquid wealth itself and not only net worth: a household keeps a buffer while carrying debt, and does not net the two against each other. Liquid wealth demand, $W_j^d = (v_j')^{-1}(r_j^o - r_j^W)$, falls in the cost of carry, with slope $1/v_j'' < 0$. Net worth enters only through the opportunity cost: for a household that does not borrow, higher net worth raises illiquid holdings, lowers their marginal return, and so raises desired liquid wealth; for a household that borrows, additional net worth retires debt and leaves liquid wealth unchanged.

Under the log forms $v_j(W) = b_j \ln W$ and $F_j(\mathcal{A}) = a_j \ln \mathcal{A}$, the problem is the portfolio counterpart of a Cobb-Douglas consumer. Substituting the balance sheet into (51), a household that borrows maximizes

$$b_j \ln W_j^d + a_j \ln \mathcal{A}_j \;-\; (r^\ell - r_j^W)\, W_j^d \;-\; r^\ell\, \mathcal{A}_j \;+\; r^\ell \mathcal{N}_j : \tag{55}$$

it values the services of its liquid and illiquid holdings with constant weights and pays for each at its cost of carry, with the budget supplied by borrowing at the loan rate. As under Cobb-Douglas demand, it spends a fixed amount of interest on each, $(r^\ell - r_j^W)\, W_j^d = b_j$ and $r^\ell \mathcal{A}_j = a_j$, so that $W_j^d = b_j/(r^\ell - r_j^W)$ and $\mathcal{A}_j = a_j/r^\ell$. For a household that does not borrow, r_j^o replaces the loan rate. General concave v_j and F_j relax the unit elasticities.

Market clearing and the absorption of credit-created liquid wealth. Consolidate the non-bank sector. Claims among non-banks, such as non-bank loans and the money market fund shares that finance them, net out; debt owed to banks equals bank loans ℓ; and the illiquid assets held by non-banks equal their supply $\bar{\mathcal{A}}$ less the securities S held by banks. Summing (9) across households therefore gives $\sum_j W_j^d = \sum_j \mathcal{N}_j + \ell + S - \bar{\mathcal{A}}$, and at given net worth and asset supply, $\sum_j dW_j^d = d\ell + dS = dL$: when the loan and asset markets clear, through the loan rate and asset returns, the market for liquid wealth clears by Walras' law. The share held by group H is $\nu = dW_H^d/dL$. In the simplest case, households and firms borrow from banks and trade real assets among themselves, while institutional investors sell the securities that banks purchase, so that ν is the loan share of marginal bank credit, $\pi = \lambda_\ell/\lambda$ in Appendix A.6; in general, proceeds are passed on across groups before they settle, and $\nu = \pi g_\ell + (1-\pi) g_s$ as in (37).

Households absorb the created liquid wealth through the cost of carry. Differentiating (54), with $\partial r_j^W/\partial r^d = s_j$,

$$dW_j^d \;=\; \frac{dr_j^o - s_j\, dr^d}{v_j''(W_j^d)}. \tag{56}$$

A credit expansion works through the first term, and through different households at each step. Banks expand lending by cutting the loan rate, which lowers r_j^o for the households that borrow; the borrowers keep part of the loan as liquid wealth and spend the rest on assets; and the sellers part with assets only at prices that make holding the liquid claims worthwhile, so asset prices rise and expected returns fall, which lowers r_j^o for the households that sell. A recipient who does not want the claim passes it on by buying an asset from someone else, or extinguishes it by repaying its own bank debt, which reduces dL. By Walras' law the increase in desired liquid wealth equals the net credit dL: in the aggregate, holding the liquid wealth that credit creates is the same thing as carrying the net debt. The deposit rate plays no part in this argument, and the argument says nothing about it. The level of liquid wealth is pinned down by the loan and asset markets; the deposit rate governs its allocation through (11), and whether it moves is a question

about bank pricing, taken up below.

The one place where the deposit rate could affect the level of liquid wealth is the second term in (56): a higher return on liquid wealth raises desired liquid wealth, which households would accommodate by borrowing more or by trading illiquid assets across groups. We abstract from this response, so that D_r below is the slope of deposit demand at given liquid wealth and λ is unchanged. The abstraction is immaterial when deposit rates do not adjust ($\psi = 0$) and is likely to be small otherwise: with ε_W the semi-elasticity of liquid wealth demand with respect to the cost of carry, the omitted response of deposits is at most $(\varepsilon_W/\alpha_d)\,\tilde{s}/(1-\tilde{s})$ times D_r, where $\tilde{s}$ is the deposit-weighted average deposit share, and this bound is reached only if households can expand liquid wealth at an unchanged opportunity cost of funds, which the weeks over which the propagation runs do not allow.

Credit creation and the recipients of proceeds. Equation (**??**) holds whatever the loan finances. If it finances consumption, the borrower's debt rises with no asset against it and the seller receives the proceeds as income, so net worth moves from the borrower to the seller while aggregate net worth and liquid wealth change as in (**??**); liquid wealth falls again only when the loan is repaid, which we abstract from because loans are long-maturity claims. In (33), only the final holder matters: a recipient may pass proceeds on, and ν records where they come to rest. Proceeds that leave both groups, into currency, tax payments, or the foreign sector, make the two weights sum to less than one. Two groups are the simplest case. With a continuum of households indexed by their deposit share s and credit proceeds distributed according to G, $\theta = \int s\, dG(s)$, and a decline in θ is a shift of G toward low-share households; the satiation extension below adds a dependence of θ on the level of liquid wealth.

Implied demand for credit. For a household that borrows, (53) gives $W_j^d = (v_j')^{-1}(r^\ell - r_j^W)$ and $\mathcal{A}_j = (F_j')^{-1}(r^\ell)$, so its debt is $\mathcal{B}_j = W_j^d + \mathcal{A}_j - \mathcal{N}_j$: a household borrows when its net worth falls short of the holdings it desires at the loan rate,

$$\mathcal{N}_j \;<\; \bar{\mathcal{N}}_j \;\equiv\; (v_j')^{-1}(r^\ell - r_j^W) + (F_j')^{-1}(r^\ell), \tag{57}$$

and does not borrow otherwise. Its debt falls in the loan rate, $\partial \mathcal{B}_j/\partial r^\ell = 1/v_j'' + 1/F_j'' < 0$, with a liquidity component and an investment component. In the borrower block of Section 4.2.3, the

semi-elasticity of bank credit demand with respect to the loan rate is $\alpha_\ell(1 - s_\ell)$; matching it to $-(1/v_j'' + 1/F_j'')/\mathcal{B}_j$ is a normalization on the curvature of v_j and F_j.

Marginal and average deposit shares. The recycling rate and the aggregate deposit share weight the same two group shares differently:

$$s_d = \frac{W_H^d}{W^d}\, s_H + \frac{W_I^d}{W^d}\, s_I, \qquad \theta = \nu\, s_H + (1-\nu)\, s_I, \qquad \theta - s_d = \Big(\nu - \frac{W_H^d}{W^d}\Big)\big(s_H - s_I\big). \tag{58}$$

The two weights are determined by different objects. Under the log liquidity value, $W_j^d = b_j/\phi_j$, where $\phi_j \equiv r_j^o - r_j^W$ is group j's cost of carry and b_j the scale of its demand for liquidity, which reflects its real wealth and transaction needs. The stock weight W_H^d/W^d depends on the levels b_j/ϕ_j, whereas changes in liquid wealth satisfy $dW_j^d/W_j^d = db_j/b_j - d\phi_j/\phi_j$. The propagation of credit after a deposit inflow leaves b_j unchanged and lowers the cost of carry of the households it reaches, so

$$\nu = \frac{W_H^d\,(-d\phi_H/\phi_H)}{W_H^d\,(-d\phi_H/\phi_H) + W_I^d\,(-d\phi_I/\phi_I)}, \tag{59}$$

which equals the stock weight only if the proportional fall in the cost of carry is the same for both groups. Loans lower the cost of carry of households and firms, and securities purchases that of institutional investors, so in general $\nu \neq W_H^d/W^d$. The difference persists: differentiating $s_d = D/W^d$ at a given deposit rate, with $dD = \theta\, dL$ and $dW^d = dL$,

$$\frac{ds_d}{dL} = \frac{\theta - s_d}{W^d}, \tag{60}$$

so a dollar of credit moves the average toward the margin only in proportion to its size relative to the stock, while growth in real wealth, which raises b_H and b_I together, expands liquid wealth in proportion to existing holdings and leaves s_d where it is.

Deposit-rate adjustment. In equilibrium the deposit rate satisfies (25), and lending moves the liquidity slack that enters it. Let $\psi \in [0, 1]$ denote the fraction of the implied adjustment that occurs while the propagation runs its course, with $\psi = 0$ for sticky deposit rates and $\psi = 1$ for full adjustment. Around the equilibrium, with reserves fixed, deposit demand and deposit

pricing satisfy

$$dD = dD_0 + \theta\, dL + D_r\, dr^d, \tag{61}$$

$$dr^d = -\,\psi\, \frac{\tau \chi^2}{\Lambda^2}\, \big(dD - dL\big), \tag{62}$$

where dD_0 is an exogenous shift in deposit demand at an unchanged deposit rate and $D_r \equiv \partial D/\partial r^d$ is the slope of deposit demand. Solving,

$$dD \;=\; \frac{dD_0}{1+\psi c} \;+\; \tilde{\theta}\, dL, \qquad \tilde{\theta} \;=\; \frac{\theta + \psi c}{1+\psi c}, \qquad c \;\equiv\; D_r\, \frac{\tau \chi^2}{\Lambda^2}. \tag{63}$$

This is the propagation equation (35), with $dD_0/(1+\psi c)$ the realized inflow and $\tilde{\theta}$ the equilibrium recycling rate, a weighted average of θ and one. Because the deposit rate does not enter the loan pricing condition (24), λ is unchanged, and the multiplier on the realized inflow is $1/(1-\lambda\tilde{\theta})$. The term c measures how strongly banks bid for deposits as they lend: the deposits attracted by the deposit-rate response per dollar of loan proceeds that would otherwise be funded wholesale. It combines the slope of deposit demand with the liquidity term $\tau\chi^2/\Lambda^2$ that also governs λ, so it introduces no new parameter.

Equation (61) follows from totally differentiating $D = s_H(r^d)\, W_H^d + s_I(r^d)\, W_I^d$,

$$dD \;=\; \sum_j s_j\, dW_j^d \;+\; \sum_j W_j^d\, \frac{\partial s_j}{\partial r^d}\, dr^d \;+\; \sum_j W_j^d\, \frac{\partial s_j}{\partial \xi_D^j}\, d\xi_D^j, \tag{64}$$

with $\partial s_j/\partial r^d = \alpha_d\, s_j(1-s_j)$ and $\partial s_j/\partial \xi_D^j = s_j(1-s_j)$. Each group's liquid wealth changes through its share of credit proceeds, (33), and through changes unrelated to lending, such as a reserve-backed expansion, which we denote $d\bar{W}_j^d$. Collecting terms gives (61) with

$$dD_0 \;\equiv\; \sum_j \big[\, s_j\, d\bar{W}_j^d + D_j(1-s_j)\, d\xi_D^j \,\big], \qquad\qquad D_r \;\equiv\; \alpha_d \sum_j D_j(1-s_j), \tag{65}$$

so that dD_0 collects the exogenous inflows considered in the paper, a preference shift toward deposits or an expansion of liquid wealth not generated by bank lending. Equation (62) follows from (25): holding the policy rate, the operating cost, and the markdown fixed, $dr^d =$

$-(\chi\tau/\Lambda^2)\,d\Lambda$, and with $\Lambda = R - \chi W$ and $W = L + R - E - D$,

$$d\Lambda \;=\; (1-\chi)\,dR + \chi\,(dD - dL), \tag{66}$$

which is (10) of Appendix A.2. A deposit inflow that is not lent out retires wholesale funding, raises the slack, and lowers the value of a marginal deposit to the bank; lending that is not matched by deposits does the reverse. Allowing a fraction ψ of this adjustment,

$$dr^d \;=\; -\,\psi\,\frac{\chi\tau}{\Lambda^2}\Big[(1-\chi)\,dR + \chi\,(dD - dL)\Big], \tag{67}$$

which with $dR = 0$ is (62). Substituting (67) into (61) for general dR,

$$dD \;=\; \widehat{dD}_0 + \tilde{\theta}\,dL, \qquad \widehat{dD}_0 \;\equiv\; \frac{dD_0 - \psi c\,\frac{1-\chi}{\chi}\,dR}{1+\psi c}, \tag{68}$$

and combining with $dL = \lambda\,dD$ gives $dD = \widehat{dD}_0/(1-\lambda\tilde{\theta})$. Relative to the underlying shift, $dD/dD_0 = 1/[\,1 - \lambda\theta + \psi c\,(1-\lambda)\,]$, which is smaller than $1/(1-\lambda\theta)$ because $\lambda < 1$.

Holding the deposit rate fixed would be restrictive if the deposit rate entered the marginal cost of lending directly, as it would if loans were funded at the margin by deposits rather than wholesale funding, if the penalties depended on profits, or if the stability of deposits depended on the rate paid on them, or if loan demand responded to the deposit rate. Under either departure λ would depend on the deposit rate and (61)–(62) would have to be solved jointly with the loan pricing condition. Neither is present here: wholesale funding is the marginal funding source, the penalties depend on balance-sheet quantities, and borrower demand does not depend on the deposit rate.

The pricing component and its bound. Writing (63) as

$$\tilde{\theta} \;=\; (1-w)\,\theta + w, \qquad w \;\equiv\; \frac{\psi c}{1+\psi c}, \tag{69}$$

w is the share of measured recycling that pricing accounts for. An inflow of deposits lowers wholesale reliance and the liquidity wedge, so the deposit rate falls and the realized inflow is a fraction $1/(1+\psi c)$ of the underlying shift; lending adds wholesale funding of which recycling retires only θ, so the deposit rate rises relative to no lending, which is why $\tilde{\theta}$ exceeds θ. On net

a deposit inflow leaves the deposit rate lower, because banks lend out only a fraction $\lambda < 1$ of it. Inverting (69), $\theta = \tilde{\theta} - \psi c\,(1 - \tilde{\theta})$, and because θ is an average of deposit shares it cannot be negative:

$$\psi c \;\leq\; \frac{\tilde{\theta}}{1 - \tilde{\theta}}, \tag{70}$$

equal to \$1.99 at the early-period estimate of \$0.67 and \$0.10 at the late-period estimate of \$0.09. The bound also limits how much of the decline in $\tilde{\theta}$ pricing can explain. If quantity recycling were the same in both periods and the pricing component vanished in the late period, matching the early estimate would require $\psi c_{\text{pre}} = (\tilde{\theta}_{\text{pre}} - \tilde{\theta}_{\text{post}})/(1 - \tilde{\theta}_{\text{pre}}) \approx 1.73$, or $w \approx 0.63$: deposit-rate adjustment would have to lift early-period recycling from 0.09 to 0.67 and contribute nothing in the late period. If instead ψc were similar in the two periods, $\Delta\theta = (1 + \psi c)\,\Delta\tilde{\theta}$, and the decline in quantity recycling would exceed the measured decline. Finally, c is decreasing in the liquidity slack Λ: when reserves are abundant, banks bid less for the marginal deposit and $\tilde{\theta}$ falls toward θ, so holding $\tilde{\theta}$ fixed in the balance-sheet counterfactuals of Section 6 is exact when $\psi = 0$ and, if anything, conservative otherwise.

Reserve-backed inflows. The recursion (68) holds with the same $\tilde{\theta}$ when reserves move with the initial inflow; only the first round differs. With $dL = \lambda\,dD + (\lambda^R - \lambda)\,dR$, where λ^R is the reserve-backed response (28),

$$dD \;=\; \frac{\widehat{dD}_0 + \tilde{\theta}\,(\lambda^R - \lambda)\,dR}{1 - \lambda\tilde{\theta}}, \tag{71}$$

which for $\psi = 0$ and $dR = dD_0$ reduces to the reserve-backed multiplier $1 + \tilde{\theta}\lambda^R/(1 - \tilde{\theta}\lambda)$ in (49).

Internalizing recycling. The baseline assumes that banks do not internalize recycling, because loan proceeds settle largely at other banks. Suppose instead that the banking sector perceives that a share $\iota \in [0, 1]$ of the deposits recycled by its lending returns to it, with $\iota = 1$ a single consolidated banking sector, so that it treats deposits as $D(r^d) + \iota\theta L$ when choosing the loan rate. Wholesale funding then rises by $1 - \iota\theta$ per dollar of lending, and the loan pricing condition (24) becomes

$$r^\ell \;=\; r^S + \mu_\ell + \frac{1}{\alpha_\ell(1 - s_\ell)} + (1 - \iota\theta)\,\chi\frac{\tau}{\Lambda} + \kappa\frac{\xi}{K} - \iota\theta\left(r^S - r^d - \mu_d\right): \tag{72}$$

the liquidity wedge is scaled by $1 - \iota\theta$, and the loan is credited with the deposit margin on the recycled share. Actual deposits still satisfy $dD = dD_0 + \theta\, dL$. Differentiating (72) at a fixed deposit rate, following the steps of Appendix A.2,

$$\left[\frac{1}{\tilde{\kappa}_\ell^{nb}} + (1-\iota\theta)(1-\theta)\,\frac{\tau\chi^2}{\Lambda^2} + \frac{\xi\kappa^2}{K^2}\right] dL \;=\; (1-\iota\theta)\,\frac{\tau\chi^2}{\Lambda^2}\, dD_0, \tag{73}$$

which is $dL/dD_0 = \lambda_\iota/(1-\theta\lambda_\iota)$ with

$$\lambda_\iota \;=\; \frac{(1-\iota\theta)\,\dfrac{\tau\chi^2}{\Lambda^2}}{\dfrac{1}{\tilde{\kappa}_\ell^{nb}} + (1-\iota\theta)\,\dfrac{\tau\chi^2}{\Lambda^2} + \dfrac{\xi\kappa^2}{K^2}}. \tag{74}$$

Internalization does not replace the multiplier; it changes the lending response that enters it. A bank that treats its loans as partly self-funding lends more in levels but is less sensitive to liquidity conditions at the margin, because its loan rate loads on the liquidity slack with weight $1-\iota\theta$, so λ_ι and the total response are decreasing in ι. Recycled deposits return to a bank roughly in proportion to its share of the deposit market, so ι is small for any individual bank. Because (72) is also the loan pricing condition used in estimation, a value of ι above zero would change the estimates of τ and ξ as well.

Satiation in deposits. The liquidity value v_j concerns total liquid wealth; the deposit share within it does not depend on the level of liquid wealth. Suppose instead that the liquidity value of deposits declines in a group's deposit holdings relative to its transaction needs $\bar{D}_j$, so that ξ_D^j is replaced by $\xi_D^j - \gamma \ln(D_j/\bar{D}_j)$ with $\gamma \geq 0$. Then

$$\frac{\partial D_j}{\partial W_j^d} \;=\; \frac{s_j}{1+\gamma\,(1-s_j)}, \qquad\qquad \frac{\partial D_j}{\partial r^d} \;=\; \frac{\alpha_d\, D_j\,(1-s_j)}{1+\gamma\,(1-s_j)}, \tag{75}$$

and $\theta = \nu\,\partial D_H/\partial W_H^d + (1-\nu)\,\partial D_I/\partial W_I^d$. For $\gamma > 0$ the marginal propensity to hold deposits lies below the deposit share, and both decline as liquid wealth rises relative to transaction needs, so an expansion of liquid wealth lowers θ directly. The baseline is $\gamma = 0$.

A.6 Asset Composition, Liquidity, and Deposit Recycling

The baseline analysis bundles loans and securities into total bank credit. This extension lets the bank choose its loan rate and securities holdings separately, derives the loan and securities responses to a deposit inflow, and maps them into the recycling rate of Section 4.7.2.

A.6.1 Asset Composition and Securities as Liquidity

We decompose total bank credit as $L = \ell + S$, where $\ell = \ell(r^\ell, r^{nb})$ denotes loans and S securities. The bank chooses the loan rate r^ℓ against borrower demand, as in the baseline, and chooses the securities quantity S. The balance sheet, capital slack, and liquidity slack are

$$\ell + S + R = E + D + W, \tag{76}$$

$$K = E - \kappa(\ell + S + R), \tag{77}$$

$$\Lambda = R + \chi_S S - \chi_W W - \chi_D D. \tag{78}$$

Securities provide liquidity, in the spirit of the Liquidity Coverage Ratio (LCR), while loans do not. Each dollar of securities contributes χ_S to the liquidity buffer, with $0 < \chi_S \leq 1$, while reserves contribute one dollar. We also allow core deposits to carry an outflow weight $\chi_D \geq 0$. A deposit dollar that replaces wholesale funding therefore relaxes liquidity by $\chi_W - \chi_D$. This accommodates differences in the stability of deposits within the core-deposit category. Substituting wholesale funding, $W = \ell + S + R - E - D$, gives

$$\Lambda = \chi_W E + (\chi_W - \chi_D)D - \chi_W \ell + (\chi_S - \chi_W)S + (1 - \chi_W)R. \tag{79}$$

A loan financed with wholesale funding requires χ_W of liquid backing. A security financed in the same way supplies χ_S of liquidity, so its net liquidity requirement is $\chi_W - \chi_S$. Reserves supply a full dollar of liquidity and therefore have a net requirement of $\chi_W - 1$. Loans, securities, and reserves enter the capital constraint with the same weight. The bank solves

$$\max_{r^d, r^\ell, S} (r^\ell - r^S - \mu_\ell)\ell(r^\ell, r^{nb}) + (r^{\text{sec}} - r^S)S + (r^S - r^d - \mu_d)D(r^d) + \xi \ln K + \tau \ln \Lambda, \tag{80}$$

taking the policy rate and securities yield as given. As in the baseline, an individual bank takes aggregate recycling as external when choosing its rates and asset holdings. The securities yield

r^{sec} is fixed. Retaining the baseline logit demand for deposits, the interior first-order conditions are

$$r^{\ell} = r^{S} + \mu_{\ell} + \frac{1}{\alpha^{\ell}(1 - s_{\ell})} + \kappa\frac{\xi}{K} + \chi_{W}\frac{\tau}{\Lambda}, \tag{81}$$

$$r^{\text{sec}} - r^{S} = \kappa\frac{\xi}{K} + (\chi_{W} - \chi_{S})\frac{\tau}{\Lambda}, \tag{82}$$

$$r^{d} = r^{S} - \mu_{d} - \frac{1}{\alpha_{d}(1 - s_{d})} + (\chi_{W} - \chi_{D})\frac{\tau}{\Lambda}. \tag{83}$$

Loans and securities face the same capital wedge, $\kappa\xi/K$. Each security dollar provides χ_S of liquid backing, reducing its liquidity wedge relative to loans. When $\chi_S < \chi_W$, wholesale-funded securities still consume liquidity on net. Deposits relax liquidity when their outflow weight is below that of the wholesale funding they replace.

A.6.2 Local Pass-Through to Loans and Securities

We follow the baseline derivation, holding reserves, equity, the policy rate, and the securities yield fixed. Total bank credit is $L = \ell + S$. Differentiating the balance sheet and regulatory slacks gives

$$dW = dL - dD, \tag{84}$$

$$dK = -\kappa dL, \tag{85}$$

$$\begin{aligned} d\Lambda &= (\chi_{W} - \chi_{D})dD - \chi_{W}dL + \chi_{S}dS \\ &= (\chi_{W} - \chi_{D})dD - (\chi_{W} - \chi_{S})dL - \chi_{S}d\ell. \end{aligned} \tag{86}$$

The liquidity response is the baseline expression augmented by the liquidity supplied by securities. The borrower-choice and non-bank pricing calculations are unchanged. Using the residual loan-demand curvature $\tilde{\kappa}_{\ell}^{nb}$ from (20), the differentiated loan-pricing condition is

$$-\frac{1}{\tilde{\kappa}_{\ell}^{nb}}d\ell = -\frac{\kappa\xi}{K^{2}}dK - \frac{\tau\chi_{W}}{\Lambda^{2}}d\Lambda. \tag{87}$$

Because the securities yield is fixed, its pricing condition gives

$$0 = -\frac{\kappa\xi}{K^{2}}dK - \frac{\tau(\chi_{W} - \chi_{S})}{\Lambda^{2}}d\Lambda. \tag{88}$$

Subtracting the securities condition from the loan condition, and using $dK = -\kappa dL$, yields

$$d\ell = \tilde{\kappa}_\ell^{nb} \frac{\tau \chi_S}{\Lambda^2} d\Lambda, \tag{89}$$

$$\frac{\xi \kappa^2}{K^2} dL = \frac{\tau(\chi_W - \chi_S)}{\Lambda^2} d\Lambda. \tag{90}$$

These conditions determine how additional liquidity translates into loans and total credit. Substituting them into the liquidity-slack response and collecting terms gives

$$\left[1 + \frac{\tau K^2 (\chi_W - \chi_S)^2}{\xi \kappa^2 \Lambda^2} + \tilde{\kappa}_\ell^{nb} \frac{\tau \chi_S^2}{\Lambda^2}\right] d\Lambda = (\chi_W - \chi_D) dD. \tag{91}$$

Solving (91) for the change in liquidity slack gives

$$d\Lambda = \frac{\chi_W - \chi_D}{1 + \frac{\tau K^2 (\chi_W - \chi_S)^2}{\xi \kappa^2 \Lambda^2} + \tilde{\kappa}_\ell^{nb} \frac{\tau \chi_S^2}{\Lambda^2}} dD. \tag{92}$$

The differentiated asset-pricing conditions imply

$$d\ell = \tilde{\kappa}_\ell^{nb} \frac{\tau \chi_S}{\Lambda^2} d\Lambda. \tag{93}$$

Substituting (92) into this expression yields

$$d\ell = \frac{\tilde{\kappa}_\ell^{nb} \frac{\tau \chi_S (\chi_W - \chi_D)}{\Lambda^2}}{1 + \frac{\tau K^2 (\chi_W - \chi_S)^2}{\xi \kappa^2 \Lambda^2} + \tilde{\kappa}_\ell^{nb} \frac{\tau \chi_S^2}{\Lambda^2}} dD. \tag{94}$$

Hence, dividing the numerator and denominator by $\tilde{\kappa}_\ell^{nb}$, the marginal loan response is

$$\lambda_\ell \equiv \frac{d\ell}{dD} = \frac{\frac{\tau \chi_S (\chi_W - \chi_D)}{\Lambda^2}}{\frac{1}{\tilde{\kappa}_\ell^{nb}} + \frac{\tau \chi_S^2}{\Lambda^2} + \frac{\tau K^2 (\chi_W - \chi_S)^2}{\xi \kappa^2 \Lambda^2 \tilde{\kappa}_\ell^{nb}}}. \tag{95}$$

The securities-pricing condition implies

$$dS = \left[\frac{K^2 (\chi_W - \chi_S)}{\xi \kappa^2 \chi_S \tilde{\kappa}_\ell^{nb}} - 1\right] d\ell. \tag{96}$$

Substituting the loan response therefore gives

$$\lambda_S \equiv \frac{dS}{dD} = \frac{\frac{\tau(\chi_W-\chi_D)}{\Lambda^2}\left[\frac{K^2(\chi_W-\chi_S)}{\xi\kappa^2\tilde{\kappa}_\ell^{nb}} - \chi_S\right]}{\frac{1}{\tilde{\kappa}_\ell^{nb}} + \frac{\tau\chi_S^2}{\Lambda^2} + \frac{\tau K^2(\chi_W-\chi_S)^2}{\xi\kappa^2\Lambda^2\tilde{\kappa}_\ell^{nb}}}. \tag{97}$$

For $\chi_D < \chi_W$, loans increase, while securities increase if

$$\chi_W - \chi_S > \tilde{\kappa}_\ell^{nb}\chi_S\frac{\xi\kappa^2}{K^2}. \tag{98}$$

Adding the loan and securities responses gives

$$\lambda \equiv \frac{dL}{dD} = \lambda_\ell + \lambda_S = \frac{\frac{\tau K^2(\chi_W-\chi_S)(\chi_W-\chi_D)}{\xi\kappa^2\Lambda^2\tilde{\kappa}_\ell^{nb}}}{\frac{1}{\tilde{\kappa}_\ell^{nb}} + \frac{\tau\chi_S^2}{\Lambda^2} + \frac{\tau K^2(\chi_W-\chi_S)^2}{\xi\kappa^2\Lambda^2\tilde{\kappa}_\ell^{nb}}}. \tag{99}$$

Thus, when $\chi_D < \chi_W$, total credit rises if $\chi_S < \chi_W$. Securities can decline even when total credit rises. More generally, both assets expand when (98) holds. Wholesale funding adjusts by $dW/dD = \lambda - 1$.

To recover the baseline, hold securities fixed, $dS = 0$, so that total credit adjusts through loans alone. The differentiated loan-pricing condition then gives

$$\left.\frac{d\ell}{dD}\right|_{dS=0} = \frac{\frac{\tau\chi_W(\chi_W-\chi_D)}{\Lambda^2}}{\frac{1}{\tilde{\kappa}_\ell^{nb}} + \frac{\tau\chi_W^2}{\Lambda^2} + \frac{\xi\kappa^2}{K^2}}. \tag{100}$$

With $\chi_D = 0$ and $\chi_W = \chi$, this is the baseline response λ_D in (24). With securities freely chosen, a zero securities response occurs at the equality case of (98).

A.6.3 Asset Composition and the Recycling Rate

The recycling rate of Section 4.7.2 depends on the composition of bank credit through the share ν of credit proceeds that come to rest with households and firms, $\nu = \pi g_\ell + (1-\pi)g_s$ in (37), where π is the loan share of marginal bank credit. In the baseline the capital constraint prices loans and securities alike and leaves π open. Here loans and securities carry different liquidity weights, respond differently to a deposit inflow, and π follows.

Deposits created by a marginal expansion are

$$dD \;=\; dD_0 + \theta_\ell \, d\ell + \theta_S \, dS, \tag{101}$$

with $\theta_\ell = g_\ell s_H + (1 - g_\ell) s_I$ and $\theta_S = g_s s_H + (1 - g_s) s_I$ the deposit shares of the recipients of loan and securities proceeds, as in (38). Substituting $d\ell = \lambda_\ell \, dD$ and $dS = \lambda_S \, dD$ from (95) and (97),

$$dD \;=\; \frac{dD_0}{1 - \lambda_\ell \theta_\ell - \lambda_S \theta_S}, \qquad\qquad dL \;=\; (\lambda_\ell + \lambda_S) \, dD. \tag{102}$$

Comparing with the baseline multiplier $1/(1 - \lambda\theta)$, with $\lambda = \lambda_\ell + \lambda_S$,

$$\lambda\theta \;=\; \lambda_\ell \, \theta_\ell + \lambda_S \, \theta_S, \qquad\qquad \theta \;=\; \frac{\lambda_\ell}{\lambda} \, \theta_\ell + \frac{\lambda_S}{\lambda} \, \theta_S, \tag{103}$$

which is (38) with $\pi = \lambda_\ell / \lambda$. Each asset contributes the product of its own lending response and its own recycling rate to the feedback between deposits and credit. The extension supplies the bank's side of this product, λ_ℓ and λ_S; the recycling rates θ_ℓ and θ_S depend on the pass-through shares g_ℓ and g_s, and the multiplier estimates discipline the combination $\lambda\theta$. With securities held fixed, as in (100), $\pi = 1$ and $\theta = \theta_\ell$.

The decomposition (39) can be read in these terms. Holding g_ℓ and g_s fixed, its composition term is $\Delta\pi \, (\bar{\theta}_\ell - \bar{\theta}_S)$. A fall in π moves θ toward θ_S and, through (103), changes the multiplier even when the total response λ is unchanged. Because π is the composition of the marginal expansion rather than of the stock of bank credit, it can move by more than the securities share of bank credit: when securities count toward the liquidity buffer, $\chi_S > 0$, a bank that meets a deposit inflow by adding securities consumes less of its liquidity slack than one that adds loans, which can make the marginal expansion more securities-heavy than the stock. Section 5.4 uses this reading of the composition term when interpreting the estimated decline in θ.

A.3 Details on Model Calibration

A.1 Data Construction

This appendix describes the construction of the aggregate quantity and price series used in the calibration and estimation of the model.

We assemble a quarterly panel for the U.S. commercial banking sector spanning 1999Q1 to 2024Q4. Balance sheet quantities are obtained from the Federal Reserve's weekly series for all commercial banks, including total assets, total liabilities, total loans, total deposits, and large time deposits. Weekly observations are converted to quarterly frequency by taking the last available observation within each calendar quarter. Core deposits are defined as total deposits net of large time deposits. Wholesale funding is defined as total liabilities net of core deposits, so large time deposits enter residual wholesale bank funding rather than the stable deposit measure. Equity is constructed residually as the difference between assets and liabilities. The model's liquid asset position R is measured as total bank assets net of bank credit, where bank credit includes loans and securities. This measure includes central bank reserves together with other liquid claims such as fed funds sold, reverse repo, and cash items, which provide effective liquidity alongside reserves. It is therefore larger than the stock of reserve balances alone, and the liquidity slack Λ and the QE and QT exercises should be read with this broader liquid asset measure in mind.

Deposit and loan rates are constructed using interest income and expense data from the FDIC Quarterly Banking Profile (QBP). Specifically, deposit and loan rates are measured as annualized ratios of trailing four-quarter interest flows to trailing four-quarter average stocks. Because the bank-credit quantity includes loans and securities, the baseline uses the QBP loan yield as the gross return proxy for bank credit and applies provision and demand-markup terms only to the loan share in the pricing equation. Using trailing windows smooths high-frequency variation in flows and ensures comparability with the quarterly balance sheet quantities. The policy rate is measured by the effective federal funds rate, averaged over the same trailing four-quarter window.

Market credit rates are measured using Moody's Aaa corporate bond yields. On the borrower side, nonbank credit is proxied by the stock of outstanding U.S. corporate debt securities from the Financial Accounts of the United States. On the depositor side, money market fund

assets and currency in circulation are used as alternatives to deposits. All quantities are scaled by nominal GDP to create consistent time series.

Regulatory quantities entering the model are constructed from aggregate balance sheet objects. The capital requirement in each regime is set to the minimum observed ratio $E_t/(L_t + R_t)$ less a one-percentage-point buffer. The liquidity requirement in each regime is set to the minimum observed ratio R_t/W_t less the same one-percentage-point buffer. The buffer is therefore measured in ratio points rather than as a common GDP-scaled level of slack. Capital slack and liquidity slack are computed as functions of equity, loans, reserves, core deposits, and wholesale funding, and therefore vary over time with the evolution of bank balance sheets.

A.2 Demand System Details

This appendix provides additional details on the construction of market sizes, market shares, and demand objects used in the calibration.

On the depositor side, the market size in quarter t is defined as

$$W_t^d \equiv D_t + M_t + C_t,$$

where D_t denotes core deposits, M_t denotes money market fund assets, and C_t denotes currency in circulation. Depositor market shares are given by $s_t^d \equiv D_t/W_t^d$, $s_t^M \equiv M_t/W_t^d$, and $s_t^C \equiv C_t/W_t^d$.

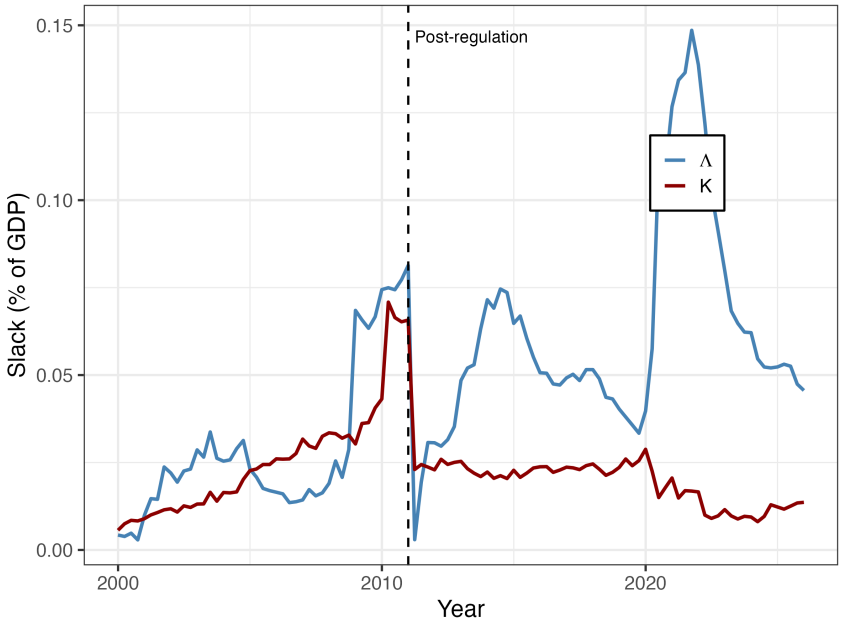
On the borrower side, bank credit is measured by aggregate bank credit L_t, and nonbank credit is measured by the stock of outstanding corporate debt securities $L_{nb,t}$. To allow for an outside option of not borrowing, total borrower market size is defined as

$$W_t^b \equiv \eta_b(L_t + L_{nb,t}),$$

where $\eta_b > 1$ pins down the average share of non-borrowers. Borrower market shares are $s_t^L \equiv L_t/W_t^b$ and $s_t^{nb} \equiv L_{nb,t}/W_t^b$, with the no-borrowing share given by $s_t^0 \equiv 1 - s_t^L - s_t^{nb}$. In the baseline calibration we set the no-borrowing share to $s^0 = 0.20$, so $\eta_b = 1/(1 - s^0) = 1.25$.

Figure A.5: Model-Implied Capital and Liquidity Slack

This figure plots the model-implied capital slack K_t and liquidity slack Λ_t, each expressed relative to GDP. The vertical line marks the start of the late regime. Lower values indicate tighter constraints.



A.3 Implied Balance Sheet Slack Variables

Next we provide additional details on the evolution of the capital and liquidity slack variables used in the model. Capital slack is defined as the difference between available equity and required equity under the balance-sheet-implied capital requirement. Liquidity slack is defined as the difference between reserves and wholesale bank liabilities scaled by the balance-sheet-implied liquidity requirement. Both slack variables are functions of balance sheet quantities and therefore vary over time with the evolution of deposits, loans, wholesale bank funding, reserves, and equity. Figure A.5 plots the implied quarterly series for capital slack and liquidity slack over the sample period. The figure illustrates a pronounced tightening of both constraints following the financial crisis and the implementation of post-crisis regulatory reforms. These movements play a central role in shaping pricing behavior and the propagation of deposit shocks in the model.

A.4 Additional Counterfactual Results

A.1 Counterfactual Construction: QE and QT

A.1.1 Quantitative Easing

We measure QE-created depositor wealth as the fraction of reserve growth attributable to the Fed's asset purchases. For each quarter t, we define

$$Q_{QE,t} = \phi \cdot \max(R_t - \bar{R}_{pre}, 0), \tag{104}$$

where $\bar{R}_{pre}$ is the average reserve-to-GDP ratio over 2005–2007, before QE began, and ϕ is the share of reserve growth attributed to asset purchases that create money-like wealth for the non-bank public. Our baseline sets $\phi = 1$. We assume QE-created wealth is initially split between deposits and short-term funding markets, with no expansion of cash. Let ζ_{QE} denote the deposit share of this initial split. The baseline sets $\zeta_{QE} = 0.5$:

$$D_{QE,t} = \zeta_{QE} Q_{QE,t}, \qquad M_{QE,t} = (1 - \zeta_{QE}) Q_{QE,t}, \qquad C_{QE,t} = 0. \tag{105}$$

We construct a counterfactual balance sheet by removing QE-created reserves and the associated depositor-wealth components:

$$\tilde{D}_t = D_t - D_{QE,t}, \qquad \tilde{M}_t = M_t - M_{QE,t}, \qquad \tilde{C}_t = C_t, \qquad \tilde{R}_t = R_t - Q_{QE,t}. \tag{106}$$

We evaluate the counterfactual at the observed 2021Q4 balance sheet, the last quarter before the QT window begins. Given the smaller deposit base $\tilde{D}$, money market fund depth $\tilde{M}$, and reserves $\tilde{R}$, we hold bank and non-bank credit fixed and let wholesale funding adjust residually to satisfy the bank balance sheet. We then evaluate the recycled-deposit response $\tilde{\lambda}_D$ and the reserve-backed first-round response $\tilde{\lambda}^R$ at the counterfactual balance sheet and compute the bank-credit response per dollar of reserve-backed deposit inflow, $\tilde{\lambda}^R/(1 - \theta_{post}\tilde{\lambda}_D)$, the object reported in Figure 6. Holding θ_{post} fixed in both the observed and no-QE economies isolates the balance-sheet effect of QE on the marginal bank-credit response, rather than adding a separate counterfactual change in the deposit-recycling rate.

A.1.2 Quantitative Tightening

We measure the total QT volume as the decline in the Fed's combined Treasury and MBS holdings from their peak in 2022Q2 to 2024Q4, expressed as a fraction of GDP. Let this quantity be X_{QT}. A pass-through parameter ω captures the share of the runoff absorbed outside bank balance sheets, for example through non-bank or money market fund adjustment. The bank balance-sheet shock is therefore $(1-\omega)X_{QT}$, and the baseline sets $\omega = 0.5$.

We consider three short-run implementations. In the asset-swap experiment, banks absorb $(1-\omega)X_{QT}$ of the runoff by replacing reserves with less liquid securities. Total bank assets other than bank credit, deposits, wholesale funding, and credit are held fixed, so the capital-relevant balance sheet is unchanged. The liquidity asset entering the liquidity constraint falls by $(1-\omega)X_{QT}$, capturing a tighter liquidity buffer without an immediate decline in deposit claims. In the reserve-drain experiment, liquid assets other than bank credit fall by $(1-\omega)X_{QT}$. Bank credit and deposits are held fixed, and wholesale funding adjusts residually to satisfy the bank balance-sheet identity,

$$W = L + R - E - D.$$

In the reserve-and-deposit-drain experiment, liquid assets other than bank credit and deposits both fall by $(1-\omega)X_{QT}$. Bank credit and non-bank credit are held fixed, and wholesale funding is again determined by the bank balance-sheet identity. Because assets and stable deposit liabilities fall one-for-one, this short-run experiment isolates the effect of lower liquid asset buffers without the Acharya-Rajan wedge between reserves and deposit claims.

A.2 Money-Market-to-Non-bank Spillover Calculation

This appendix records the simple calculation behind the non-bank offset exercise in Section 6. In the baseline figure, non-bank marginal costs are unrelated to money-market depth, so $\kappa_M = 0$. The money-market spillover calculation asks how large κ_M would have to be to restore the late-period aggregate credit response to its early-period level.

Let C_{pre} and C_{post} denote the baseline total-credit responses to a one percent of GDP deposit inflow in the early and late periods, combining the bank-credit response with non-bank borrower substitution. In the current calibration, $C_{pre} = 1.14$ and $C_{post} = 0.35$, so the aggregate

gap to be closed is

$$\Delta C = C_{pre} - C_{post} = 0.79.$$

For reference, the corresponding bank-credit responses are $B_{pre} = 2.14$ and $B_{post} = 0.62$. The aggregate gap is about half the bank-credit decline because non-bank crowding-out through borrower substitution was larger in the early period, when the bank response itself was stronger.

Let a share ω_M of loan proceeds that do not recycle into stable deposits flow into money market funds. Then money-market depth is evaluated at the selected no-spillover total-credit response:

$$dM^{sel} = \omega_M(1 - \theta_{post})dL^0_{\text{total},post}.$$

Here $dL^0_{\text{total},post}$ is the no-spillover total-credit response, including the deposit-recycling multiplier and baseline non-bank borrower substitution. Using (48), define $\Delta L^{spill}_{\text{total},post}(\kappa_M)$ as the selected-allocation total-credit response under a money-market spillover. The terms $\lambda_{L,M,post}(\kappa_M)$ and $\lambda^{cost}_{nb,M,post}(\kappa_M)$ are both linear in κ_M under the parameterization $\mu_{nb}(M) = \bar{\mu}_{nb} - \kappa_M M$, and the spillover increment is linear in the leakage share ω_M. The offset loading is the value that solves

$$\Delta L^{spill}_{\text{total},post}(\kappa_M^{offset}) - \Delta L^{spill}_{\text{total},post}(0) = \Delta C.$$

Under full leakage, $\omega_M = 1$, the required loading is $\kappa_M^{offset} = 2.27$. Under half leakage, $\omega_M = 0.5$, the required loading doubles to $\kappa_M^{offset} = 4.54$.