THE DECLINING ROLE OF DEPOSITS IN CREDIT CREATION*

Naz Koont
Stanford University

Stefan Walz
Boston College

May 12, 2026

Abstract

Does deposit funding still drive credit creation? Deposit inflows affect credit both directly, through banks' lending 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 credit impact of a $1 deposit inflow has fallen from $1.73 to $0.42 over the past two decades. The decline reflects a weakening of the credit–deposit feedback loop, consistent with a declining specialness of bank deposits and greater use of non-bank alternatives. Regulation and policy also shape banks' credit response: capital requirements dampen it while liquidity regulation partially offsets, and both QE and QT reduce credit in the post-crisis environment. To offset the lost credit creation, non-bank credit must expand threefold beyond what banks' diminished lending response alone implies. 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, Zhiguo He, Arvind Krishnamurthy, Wenhao Li, Rodney Ramcharan, Martin Schneider, Amit Seru, Phil Strahan, Adi Sunderam, Christopher Tonetti and seminar participants at the USC Macro-Finance Workshop and Boston College 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). Beyond this direct effect, the act of lending itself can generate new stable deposits, generating a feedback loop that amplifies the overall 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 in the modern financial system, or about the forces that determine its strength and how it evolves as the structure of the financial sector changes.

In this paper, we decompose the total effect of deposit inflows on credit into a direct lending 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 tighter regulation 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 provide empirical evidence on the magnitude and evolution of banks' credit-deposit multiplier using a high-frequency identification strategy over the period 2005-2019. The credit–deposit multiplier embeds both banks' marginal lending response and the recycling of credit into new deposits, capturing how an initial deposit inflow propagates through successive rounds of credit creation. Exploiting plausibly exogenous innovations in the timing of U.S. Treasury cash settlements, which mechanically credit and debit private bank accounts but are orthogonal to economic fundamentals, we instrument for contemporaneous deposit inflows and trace their propagation through the banking system using local projections. This approach allows us to recover a causal estimate of how an initial deposit inflow translates into cumulative deposit growth over subsequent weeks. Our identification strategy allows us to circumvent two key endogeneity concerns related to the estimation of the credit-deposit multiplier. The first is correlated deposit flows driven by unobserved economic conditions, and the second is real-side feedback from lending to income and deposit demand. Exploiting high-frequency innovations in Treasury settlement timing provides an exogenous source of deposit inflows, and our focus on short horizons limits the scope for real-side adjustment.

We find that the credit-deposit multiplier is sizable: a $1 deposit inflow ultimately raises cumulative deposits by $2.05 through balance sheet propagation. However, the strength of this amplification has weakened markedly over time, with the multiplier declining by roughly half from $2.6 in the first half of our sample to around $1.3 in the second half. This decline indicates a weakening of one or both of the underlying margins that govern the feedback loop: either a reduced propensity of banks to expand credit in response to deposit inflows, or a lower tendency for credit creation to generate additional stable deposit funding. Such changes may be a result of post-crisis regulatory reforms or the growing role of non-bank intermediaries in credit and liquidity provision. To disentangle these mechanisms and recover the total effect on credit, we next develop a structural model that decomposes the decline in the credit-deposit multiplier into its underlying drivers.

Our structural model features depositors and borrowers who choose between banks and non-bank outside options, alongside a banking system that sets deposit and loan rates. Banks' balance sheet expansion is disciplined by capital and liquidity requirements, while non-bank intermediaries compete with banks in credit markets and thereby shape the elasticity of bank credit demand. Within this environment, the credit-deposit multiplier emerges endogenously: banks' lending response is determined by underlying primitives, while the recycling of loan proceeds into stable deposits governs the evolution of the deposit base.

We calibrate the model by matching model-implied optimality conditions to their empirical counterparts using aggregate U.S. banking data. The banking sector's deposit and loan first-order conditions jointly identify the shadow costs of illiquidity and leverage. The non-bank first-order condition identifies the cost of expanding non-bank credit supply, which determines how strongly non-bank lending offsets changes in bank lending. Finally, we use the reduced-form estimates of the credit-deposit multiplier in the early and late periods to pin down the recycling rate of loan proceeds into stable deposit funding.

The estimated model delivers the two sufficient statistics that govern the credit-deposit multiplier: banks' marginal lending response to deposit inflows, and the recycling rate of loan proceeds into stable deposits. We find that both have declined sharply. Per dollar of deposit inflow, banks' lending response has declined from .66 to .32. In the earlier period, banks channeled roughly two-thirds of each marginal deposit dollar into new lending; in the later period, less than one-third supports credit expansion. Most deposit inflows are now absorbed by reductions in wholesale funding or by increases in liquid asset buffers. The shift is stark: deposits have shifted from being primarily a source of credit creation to primarily a tool of balance sheet management. Moreover, the recycling of loan proceeds into deposit inflows has fallen from .92 to .76. In the earlier period, over 90 cents of every dollar of loan-generated payment flows ultimately settled as stable deposits within the banking system, so that nearly all of the funds created by lending returned in a form that relaxed banks' liquidity constraints and could support further credit expansion. In the later period, only about 76 cents returns as stable deposits. This is consistent with a reduced specialness of bank deposits as a savings instrument, as households and firms increasingly allocate funds toward non-bank alternatives such as money market funds and other savings vehicles, diminishing the share of payment flows that returns to the banking system as the kind of stable, low-cost funding that supports further credit expansion.

Turning to the central question of the paper, we find that the total effect of a deposit inflow on bank credit has declined from 1.73 to 0.42: each dollar of deposits today generates less than a quarter of the credit it once did. This total effect on credit is the product of banks' lending response and the amplification of deposits through the credit-deposit multiplier, so both forces contribute substantially to the decline. The weaker lending response reduces the credit generated in each round, while diminished deposit recycling 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 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 decompose the decline in the credit impact of deposit flows along two dimensions: shifts in structural primitives governing bank behavior, and shifts in balance-sheet quantities driven by post-crisis monetary policy. On the primitives side, we find that tighter capital regulation more than accounts for the decline in banks' lending response, while stricter liquidity regulation partially offsets it: by reducing effective liquidity slack, it makes stable deposits more valuable at the margin and raises the sensitivity of lending to deposit inflows. Diminished deposit recycling contributes a further, quantitatively important drag, reflecting a reduced tendency for loan proceeds to return to the banking system as stable funding.

Turning to quantities, we examine how Quantitative Easing (QE), which flooded the banking system with deposits and reserves between 2008 and 2014, and Quantitative Tightening (QT), which began draining reserves in mid-2022, affect the banking sector's capacity to intermediate deposits into credit. We find that both policies dampen this capacity despite operating in opposite directions on the balance sheet, though QT's effect is largely transitional as banks shrink their loan books to restore regulatory slack. Together, the two exercises reveal that the decline in credit creation reflects both a tightening of banks' regulatory environment and the diminished specialness of deposits, alongside a transformation of bank balance sheets themselves.

This raises a natural question: can the non-bank sector 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 it is small. Second, we consider whether independent growth in non-bank lending capacity could offset the decline. We find that replacing banks' first-round lending response requires a 4 percentage point reduction in non-bank costs, while replacing the full decline once the credit-deposit multiplier is taken into account requires a reduction nearly three times as large. Moreover, such cost improvements would need to recur with each episode of deposit-driven amplification. This exercise highlights that assessing the effects on total credit requires accounting not only for changes in banks' marginal lending response, but also for the strength of the underlying funding multiplier that governs amplification and that non-banks cannot easily replicate.

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 do not 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, 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 lending is a classic question in banking. It connects to seminal theories of why banks jointly provide deposit-taking and lending services (Diamond and Dybvig, 1983; Kashyap et al., 2002a), 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 (Kashyap et al., 2002b; Bernanke and Blinder, 1988; 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 lending to deposit 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 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

[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.

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 emphasize the importance of accounting for the recycling of loan 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., 2021a,b; Li et al., 2023; DeMarzo et al., 2024; Blickle et al., 2025; Egan et al., 2025; Lu et al., 2024; Koont et al., 2024). Recent work studies how central bank reserve creation expands bank balance sheets but may fail to increase usable liquidity because banks hoard reserves in anticipation of state-contingent liquidity demand (Acharya and Rajan, 2024). In particular, deposits can become a problem if the bank expect depositors redeem jointly, which shows up in our system as deposit specialness relative to wholesale funding. While these liquidity risks also influence banks' lending incentives in our framework, we focus on a distinct amplification channel: the feedback loop between lending and deposit creation. Although all payments ultimately settle through the banking system and non-bank liabilities are backed by claims on bank deposits or reserves, our framework highlights that only stable, low-margin deposits sustain the feedback loop between lending and deposit creation. The banking sector's share of deposits can therefore be interpreted as the share of stable funding in the system, relative to more institutional or flightier deposits that do not support this amplification mechanism. Our contribution is to incorporate this recycling of deposits into the measurement of the credit–deposit multiplier and to show how changes in the composition and stability of deposits affect the strength of credit creation in the banking system.

Finally, the decline we document in the role of bank deposits in credit creation relates 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). As non-banks increasingly disintermediate banks on both sides of their balance sheets, a smaller share of funding flows through the banking system's balance sheet amplification mechanism. This raises the question of whether the expansion of non-bank intermediation can compensate for the weakening of bank-based credit creation. We show that although non-bank credit can offset part of the decline in banks' marginal lending response, fully replacing the amplification generated by bank deposits is substantially more difficult once the credit–deposit multiplier is taken into account. Our contribution is therefore to highlight that assessing the aggregate effects of financial intermediation requires accounting not only for banks' marginal lending responses, but also for the strength of the underlying funding multiplier that governs how credit expands through the banking system.

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 lending 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 lending are jointly determined within the banking system. When banks extend credit, the associated payments generate deposits elsewhere in the financial system, some of which return to banks. These newly created deposits, which we denote by θ, can in turn support further balance sheet expansion. As a result, the ultimate increase in credit following a deposit inflow depends both on banks' initial lending response and on how lending activity generates new deposits within the financial system.

To formalize this idea, let D_0 denote an exogenous inflow of deposits into the banking system and let L denote bank credit. The total effect of deposit inflows on credit can be summarized by the response dL/dD_0. Understanding this relationship requires examining how deposits and lending interact in the modern banking system and how balance sheet adjustments following deposit inflows can generate additional deposits. Section 2.2 discusses how deposits and lending are jointly determined in a fractional-reserve banking system. Section 2.3 illustrates the mechanism through a simple example. Finally, Section 2.4 formally defines the total credit effect by decomposing it into banks' marginal lending response and the amplification generated by deposit creation within the financial system.

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 lending 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 lending activity 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. Under narrow banking, the aggregate quantity of deposits is fixed by reserves, so that $Q^* = R$, where R denotes aggregate 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.

Narrow Banking

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

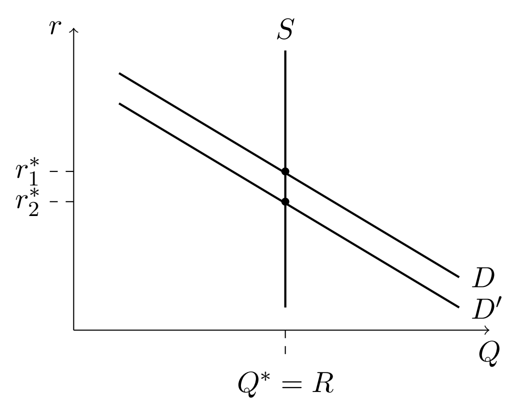


Figure 1: Narrow banking: deposits are backed one-for-one by reserves, so deposit quantity is pinned down by aggregate reserves. Demand shifts change equilibrium rates but not quantities.

In contrast, in a fractional reserve banking system as in Figure 2, deposits are jointly determined by (i) households' willingness to hold deposits and (ii) banks' willingness to expand

Fractional Reserve Banking

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

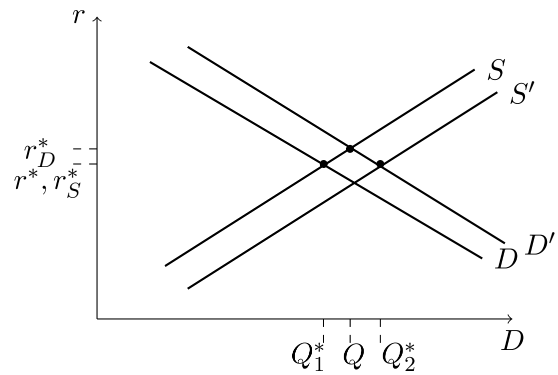


Figure 2: Non-narrow banking: deposits are jointly determined by an upward-sloping deposit supply curve (banks' willingness to expand balance sheets) and a downward-sloping deposit demand curve (households' willingness to hold deposits). Shifts in either demand or supply move both equilibrium rates and quantities.

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]

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. In Figure 3, we plot the evolution of M1/M0 from 1960 through 2025. The fact that the ratio fluctuates over time, and generally exceeds 1, is consistent with banks creating deposits

[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.

through lending.[3]

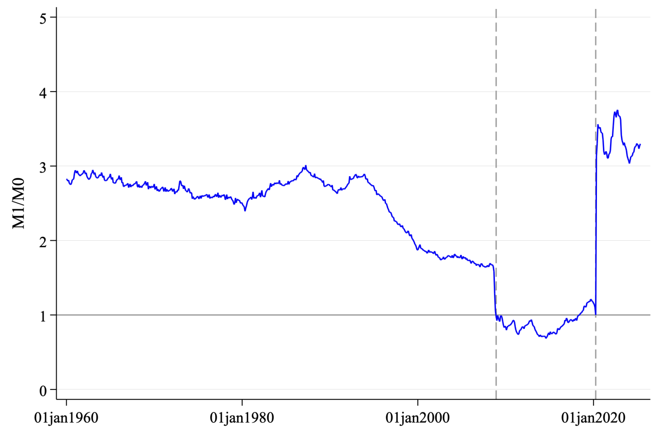


Figure 3: Evolution of M1/M0. Monthly data is from FRED (1960–2025). Vertical lines denote the onset of QE after the 2008 financial crisis and pandemic-related stimulus. In May 2020, savings deposits were reclassified as transaction accounts and added to M1.

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 lending 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 and thereby support additional balance sheet expansion. As a result, deposits created through lending can return to the banking system as funding and support further balance sheet expansion, so that an initial deposit inflow may generate successive rounds of balance sheet expansion within the banking system. The ultimate increase in credit can therefore exceed banks' initial lending response to deposits.

2.3 A Simple Example of the Credit-Deposit Feedback Loop

Beginning from an equilibrium quantity of deposits, an initial change in deposits—whether originating from a shift in demand or supply—can affect banks' balance sheets through its impact on funding conditions. If the deposits created through lending remain within the banking system as stable funding, they may relax banks' liquidity constraints at the margin and enable further

[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.

Table 1: Simple Example: Credit-Deposit Multiplier

(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	W	\$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	W	\$51
	\$214		\$214

(f) Deposit Recycling

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

In this numerical example the banking sector then expands lending by $\lambda = 0.7$ per dollar of deposit inflow, creating flighty funding, denoted as wholesale liabilities, alongside the new loans. A fraction $\theta = 0.9$ of flighty balances is recycled into deposits, initiating the next round.

balance sheet expansion. In this case, an initial change in deposits can propagate through the banking system as lending activity generates additional deposits that support further lending. The following example illustrates this mechanism in a simple setting.

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. The banking system begins with \$100 in reserves R and \$100 in wholesale funding W, as in panel (a). Suppose that a shock leads to increased demand for deposits by households. In response, banks issue \$100 in deposits D while retiring the more costly wholesale funding W, as in panel (b). The deposit inflow leaves total assets unchanged but improves the banking sector's funding composition, relaxing banks' liquidity management and raising their willingness to extend credit. As a result, suppose that banks extend new credit of \$70, as in panel (c). The loan proceeds initially appear as flighty wholesale funding at the issuing bank, and banks' balance sheets grow by \$70. The intermediary view of banks stops here, whereby an increase in deposit funding leads to an effect on banks' willingness to lend.

However, the quantity of deposits is endogenously determined, and this is what sets the feedback loop in motion. As borrowers spend the loan proceeds, payments circulate through the economy. Over time, some fraction of these proceeds settle as deposit funding in the banking system—say \$63—as in panel (d). This results in an increase in deposits and a reduction

Table 2: Aggregate Banking System: Incremental Expansion by Round

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**

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 lending equals $\lambda = 0.7$ times newly retained deposits, which in turn equal $\theta = 0.9$ times the previous round's lending. Successive rounds therefore shrink by the factor $\lambda\theta = 0.63$, and the final row reports the limiting value of the resulting geometric series.

in wholesale funding, driven again by deposit demand and banks' re-optimization of funding composition. This constitutes a further improvement in the banking sector's funding composition, again relaxing banks' liquidity management and raising their willingness to extend credit, as shown in panel (e). Some of the resulting loan proceeds again return as deposit funding in panel (f), continuing the feedback loop. After two rounds, deposits have increased to \$163 and loans to \$114.

This process continues until it converges, as shown in Table 2. Total deposits converge to \$270 and loans to \$189. Amplification arises because deposits are superior to wholesale funding: they relax the banking sector's liquidity management, 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

The simple example above illustrates how an initial deposit inflow can propagate through the banking system via successive rounds of lending and redepositing. As discussed in Section 2.1, a natural starting point for thinking about the effect of deposit inflows on credit is banks' marginal lending response to a deposit inflow, λ. However, because a portion of loan proceeds may return to the banking system as deposits, an initial deposit inflow can generate additional rounds of balance sheet expansion. This mechanism implies that the ultimate increase in credit may exceed banks' initial lending response to deposits. We formalize this mechanism using a simple propagation framework that separates banks' lending response from the recycling of deposits within the financial system.

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, which corresponds to banks' marginal lending response to deposit inflows. The second is θ, which reflects the marginal propensity for credit-financed expenditures to be redeposited within the banking system as stable bank deposits. Together, these parameters determine how an initial deposit inflow propagates through the system.

Given a deposit inflow ΔD_0, banks extend new loans 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 lending 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:

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

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 lending and redepositing. To recover the total effect of deposit inflows on bank credit, we combine this amplification mechanism with banks' marginal lending response to deposits. Specifically, if banks expand lending by λ for each additional dollar of deposits, then the total increase in credit generated by an initial deposit inflow ΔD_0 is given by

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

This expression highlights that the total credit effect of deposit inflows depends both on banks' initial lending response to deposits and on the amplification generated by the recycling of deposits within the banking system.

In the Appendix we consider whether non-bank institutions can generate a similar funding multiplier. While non-bank intermediaries share some features with banks, in that funding composition shapes their willingness to extend credit and their liabilities are determined in equilibrium by supply and demand, they lack a key ingredient that sustains the credit-deposit feedback loop. Bank deposits serve as the economy-wide settlement asset, so lending-generated payment flows return to the banking system as stable funding. Non-bank liabilities function as settlement media only within narrow subsystems such as institutional cash management or crypto markets, 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 credit-deposit 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 credit-deposit 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.[4]

We next purge this weekly series of predictable variation by residualizing on quarter-year fixed effects α_q, calendar week-of-year fixed effects α_w, and six lags of the weekly cash-flow series as in Equation (6), and use the resulting innovation, Z_t, as our instrument.[5]

$$CF_t = \alpha_q + \alpha_w + \sum_{k=1}^{6} \rho_k \, CF_{t-k} + Z_t. \tag{6}$$

This innovation captures unexpected timing variation in Treasury settlement flows. Indeed, the Ljung-Box portmanteau test fails to reject the null of no serial correlation in the innovation up to six weekly lags, consistent with the interpretation of the shock as high-frequency timing noise at the horizons that we use to estimate the deposit multiplier.

[4]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.

[5]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.

Table 3: First Stage Estimates

The table reports first-stage estimates from the instrumental variables specification linking national core deposit growth to Treasury cash-flow timing innovations. The sample covers 2005–2019. Standard errors are reported in parentheses. Coefficients on the instrument and its lag are scaled by 10,000. One, two, and three stars indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1) 2005-2019	(2) 2005-2011	(3) 2012-2019
L.Instrument	0.17***	0.27***	0.09
	(0.05)	(0.09)	(0.06)
Instrument	0.18***	0.20**	0.15**
	(0.05)	(0.09)	(0.06)
L.Deposit Growth	-0.28***	-0.25***	-0.31***
	(0.04)	(0.05)	(0.05)
Observations	731	314	417
Adjusted R^2	0.094	0.089	0.106
F	26.36	11.20	17.36

We implement this strategy in a 2SLS framework. To accommodate potential short-run timing dispersion between Treasury settlement and the recording of deposit balances, we instrument contemporaneous deposit growth using both the contemporaneous and one-period lag of the Treasury cash-flow timing innovation, allowing the first stage to load on delayed effects without imposing serial correlation on the shock itself. We measure deposits using aggregate core deposits from FRED, defined as total deposits held by all U.S. commercial banks excluding large time deposits. Our first stage specification is as follows,

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

In the second stage, we regress cumulative deposit growth at horizon h on instrumented deposit growth at time t within the local projection framework. In the vector of controls X_t we include the lag of deposit growth ΔD_{t-1}. For horizon h, the second stage regression equation is,

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

Instrument validity requires that the relevance and exclusion assumptions be satisfied. The relevance condition is supported by our first-stage F-statistics in Table 3. 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 consider two potential threats to identification. First, in the post-crisis period, Treasury cash flows may affect reserves through movements in the Treasury's account at the Federal Reserve. This could generate co-movement between deposits and reserves that is not driven by deposit inflows. In Appendix A.1, we show that our results are not sensitive to controlling for contemporaneous changes in aggregate reserves, indicating that the estimates primarily reflect deposit inflows rather than reserve co-movement at weekly frequency. Second, deposit inflows could affect credit through real-side channels rather than balance sheet propagation. For example, deposit-funded lending may stimulate economic activity, raising incomes and wealth and thereby shifting out deposit demand. However, such real-side adjustments operate over longer horizons, and the six-week window we study is likely too short for these mechanisms to play a meaningful role. As a result, real-side feedback effects are unlikely to drive our estimated deposit multiplier.

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

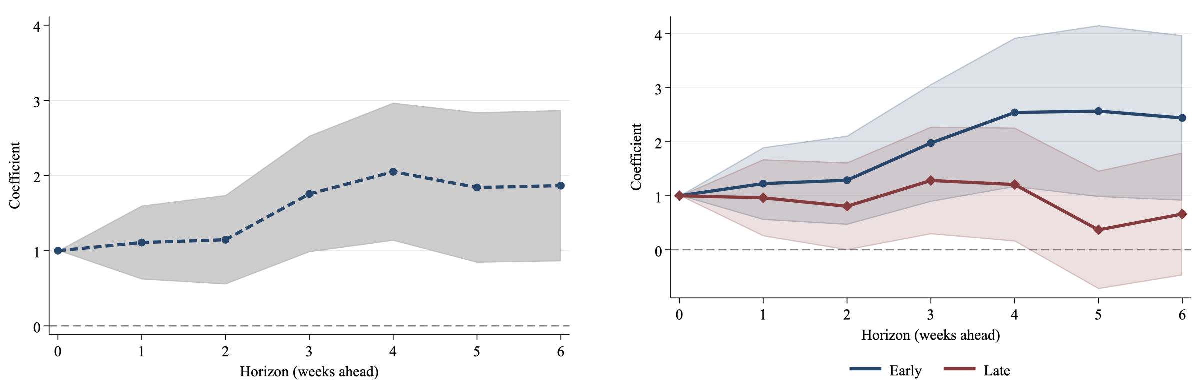
Figure 4 visualizes the coefficient estimates for our IV specification in Equation (8) at horizons h from 0 to 6, with shaded bands indicating 95% confidence intervals. Panel (a) captures the estimation for the complete sample. This specification captures how an initial funding inflow propagates through the banking system over time, including through flows of funds between banks and across institutions. We define the long-run multiplier m_D^h as the peak cumulative response, resulting in a value of 2.05. This implies that a 1 percent deposit inflow ultimately leads to a 2.05 percent increase in total deposits in the banking system. 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 an early and late period, estimating the credit-deposit multiplier over the years 2005-2011 and 2012-2019. Panel (b) compares the dynamics of the deposit multiplier across two sample periods. We estimate, for each period, our IV specification in Equation (8) at horizons $h = 0$ to 6. Comparing the magnitudes shows a marked decrease in the latter sample, falling from 2.57 in the earlier period to only 1.28.

The reduced-form multiplier embeds two distinct economic forces: a direct effect, captured by banks' marginal propensity to lend out of deposit inflows, and an indirect effect, captured by the share of loan proceeds that recycle back into deposit funding. The observed decline could therefore reflect changes in either channel, driven by post-crisis regulatory reforms and policy, the expanding role of non-bank intermediaries, or shifts in household behavior. Disentangling these forces and quantifying the total effect of deposit flows on credit requires additional structure, which we turn to next.

Figure 4: 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 2019. Panel (b) compares estimates for the early (2005-2011) versus late (2012-2019) samples. Shaded bands denote 95% confidence intervals.



(a) Full Sample

(b) Early vs. Late

4 Model

In this section we develop a model of the U.S. banking sector that connects the credit-deposit multiplier to structural primitives governing bank and household optimization, allowing us to decompose it into its constituent parts and quantify the total effect of an exogenous deposit shock on credit.

4.1 Overview

The model features four types of agents. There are depositors, borrowers, a representative banking system, and a representative non-bank credit intermediary. The key exogenous variables in the model are the current policy rate r^S and the equity capital of the banking sector E, which characterize the state of the world. Given these variables, banks set deposit and loan rates to maximize profits, subject to capital and liquidity constraints. The non-bank intermediary supplies credit to borrowers and funds itself elastically at the policy rate, facing convex intermediation costs. Depositors allocate their wealth across bank deposits, a non-bank savings vehicle, and cash, while borrowers choose among bank credit, non-bank credit, and not borrowing.

4.2 Depositors

There is a mass W^d of depositors. Each depositor i makes a discrete choice over a menu $k \in \{D, M, C\}$ that includes a bank deposit D, a non-bank savings vehicle M, and cash C, which differ in the yields they offer and the liquidity services they provide. Depositor i's utility from option k is

$$u_i^k \;=\; \alpha_d r^k + \xi_k + \varepsilon_i^k, \tag{9}$$

where α_d is the deposit rate sensitivity of depositors and ε_i^k is the idiosyncratic taste shock for depositor i over option k. Bank deposits pay rate r^d and provide liquidity services ξ_d; the non-bank savings vehicle pays 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 market share of option $k \in \{d, M, C\}$ is

$$s_k(r^d) \;=\; \frac{e^{\xi_k + \alpha_d r^k}}{e^{\xi_D + \alpha_d r^d} \;+\; e^{\alpha_d r^S} \;+\; e^{\xi_C}}, \tag{10}$$

and the demand for deposits can be written as,

$$D(r^d) \;=\; s_D(r^d) \cdot W^d. \tag{11}$$

4.3 Borrowers

There is a mass of borrowers with wealth W^b. Each borrower j makes a discrete choice over a menu $k \in \{\ell, nb, \emptyset\}$ that includes bank credit ℓ, non-bank credit nb, and not borrowing $\emptyset$. Borrower j's utility from option k is

$$u_j^k \;=\; -\alpha_\ell r^k + \xi_k + \eta_j^k, \tag{12}$$

where α_ℓ is the rate sensitivity of borrowers and η_j^k is the idiosyncratic taste shock for borrower j 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_j^k is i.i.d. Type-1 extreme value. The market share of option $k \in \{\ell, nb, \emptyset\}$ is thus

$$s_k(r^k) \;=\; \frac{e^{\xi_k - \alpha_\ell r^k}}{1 + e^{\xi_L - \alpha_\ell r^\ell} + e^{\xi_{nb} - \alpha_\ell r^{nb}}}, \tag{13}$$

and loan quantities are given by[6]

$$L_k(r^k) \;=\; W^b \, s_k(r^k), \qquad k \in \{\ell, nb\}. \tag{14}$$

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

4.4 Banking System

4.4.1 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 banking sector also holds central bank reserves R, which we treat as an exogenous, system-level quantity.[7] The balance sheet constraint is

$$L(r^\ell) + R = E + D(r^d) + W, \tag{15}$$

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

$$W = L(r^\ell) + R - E - D(r^d). \tag{16}$$

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

$$\max_{r^d, r^\ell} \,(r^\ell - r^S)L \;+\; (r^S - r^d)D \;-\; \frac{\kappa_W}{2}W^2 \;+\; \xi \ln\big(K\big) \;+\; \tau \ln\big(\Lambda\big). \tag{17}$$

In (17), the first two terms capture profits from loan and deposit markets respectively.[8] The remaining terms represent costs: a quadratic wholesale funding cost κ_W, reflecting the expense of external financing outside the bank's deposit base (Kashyap and Stein, 1995), and penalties from capital and liquidity requirements, which we describe in detail below. These frictions link deposit-taking to loan supply: without them, deposit inflows would have no effect on lending, and the credit-deposit multiplier would reduce to one.

We model banks' regulatory capital and liquidity constraints with log penalties, which reduce profits by raising marginal costs as the capital or liquidity slack approaches zero. First, the bank faces a risk-invariant capital constraint captured by $\xi \ln\big(K\big)$. Letting $\kappa > 0$ denote the

[7] In the Appendix and in the quantitative calibration, we generalize the model to allow the banking sector's liquid asset position to adjust endogenously as a residual of its balance sheet optimization. Empirically, banks manage not only central bank reserves but also interbank claims such as fed funds sold and reverse repo, which provide effective liquidity beyond the stock of reserves. We capture this margin by replacing fixed reserves R with a broader liquid asset residual that adjusts to clear the balance sheet identity. This extended specification yields a quantitatively richer lending response λ while preserving the same qualitative forces emphasized in the main text.

[8] Reserves and wholesale funds both earn the policy rate r^S. With equity E fixed, their returns affect the level of profits but not the bank's profit maximization.

level of the capital requirement, let the capital slack K be defined as

$$K \;=\; E - \kappa\,(L + R) \;=\; (1-\kappa)E \;-\; \kappa(D + W) \geq 0. \tag{18}$$

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 is similar to the Supplementary Leverage Ratio (SLR) implemented in the Basel III regulatory framework. 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 requirement captured by $\tau \ln\bigl(\Lambda\bigr)$. Let $\chi \in [0,1]$ be the required ratio of liquid asset holdings relative to runnable liabilities, and $\psi \in [0,1]$ governs the runnable fraction of deposits in the liquidity rule. Define the liquidity slack Λ as,

$$\Lambda = R - \chi(\psi D + W) \geq 0 \tag{19}$$

Here, the liquidity slack Λ is the excess of reserves R over a required fraction χ of runnable liabilities, given by the runnable share ψ of deposits D and all wholesale funding W. This formulation is in the spirit of the Liquidity Coverage Ratio (LCR) in Basel III, which requires banks to hold sufficient high-quality liquid assets to cover a fraction of their runnable short-term liabilities. From (19), higher runnable liabilities such as wholesale funding W or runnable deposits ψD tighten the constraint, while greater reserves R relax it. As runnable 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.

4.4.2 First Order Conditions

Given the profit maximization problem above, we next derive the corresponding first order conditions. We begin with the condition for the loan rate $r^{\ell\star}$,

$$r^{\ell\star} = r^S + \kappa_W W + \frac{1}{\alpha_\ell(1 - s_\ell(r^{\ell\star}))} + \chi\,\frac{\tau}{\Lambda^\star} + \kappa\,\frac{\xi}{K^\star}. \tag{20}$$

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 marginal wholesale funding cost, which reflects the fact that with reserves fixed, additional lending must be funded by expanding the balance sheet, raising reliance on wholesale borrowing. 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 runnable liabilities and therefore tightens the liquidity requirement when reserves are fixed. The fifth is a capital wedge, which reflects the shadow cost of balance sheet expansion under the leverage constraint: 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.

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

$$r^{d\star} \;=\; r^S + \kappa_W W \;-\; \frac{1}{\alpha_d\left(1 - s_d(r^{d\star} \mid r^S)\right)} \;+\; \chi(1-\psi)\frac{\tau}{\Lambda^\star}. \tag{21}$$

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 marginal wholesale funding cost, which raises the deposit rate banks are willing to pay because each additional deposit reduces wholesale borrowing one-for-one, saving $\kappa_W W$ at the margin. 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 because deposits reduce reliance on runnable wholesale funding in the liquidity requirement when $\psi < 1$; as the runnable share of deposits increases, this effect is dampened.

Taken together, equations (20) and (21) reveal a direct link between deposit taking and credit provision. Deposits reduce reliance on costly wholesale funding and, when $\psi < 1$, relax the liquidity requirement—both of which lower the optimal loan rate. 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.

4.5 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 the demand function of borrowers $L_{nb}(r^{nb})$. The credit intermediary earns a spread over the policy rate r^S and faces convex intermediation costs governed by $\kappa_{nb} > 0$, which can capture, for example, costs related to funding, monitoring, liquidity management, or regulation. The credit intermediary's profit maximization problem is given by,

$$\max_{r_{nb}} \ \Pi_{nb} = (r_{nb} - r_S)\, L_{nb}(r_{nb}) - \frac{\kappa_{nb}}{2} L_{nb}^2(r_{nb}). \tag{22}$$

The optimal non-bank loan rate satisfies

$$r_{nb} = r_S + \frac{1}{\alpha_\ell(1 - s_{nb})} + \kappa_{nb} \cdot L_{nb}(r_{nb}). \tag{23}$$

This pricing condition is analogous to the bank loan pricing equation in Equation (20). The first term is again the policy rate, which provides the baseline cost of funds in the economy. The second 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. The third is a cost wedge that is increasing in the scale of credit extended; while banks face distinct wedges arising from liquidity and capital constraints, this term captures non-bank intermediation costs in reduced form.

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

4.6 Equilibrium

An equilibrium consists of deposit and loan rates $\{r^{d\star}, r^{\ell\star}, r_{nb}^{\star}\}$, wholesale borrowing $W^\star$, and resulting allocations of deposits $D(r^{d\star})$, bank loans $L(r^{\ell\star})$, and non-bank loans $L_{nb}(r_{nb}^{\star})$, given an exogenous level of reserves R, such that:

1. Depositors maximize utility over bank deposits, cash, and the non-bank savings vehicle, yielding deposit demand $D(r^{d\star})$.

2. Borrowers maximize utility over bank loans, non-bank loans, and not borrowing, yielding bank and non-bank loan demands $L(r^{\ell\star})$ and $L_{nb}(r^{\star}_{nb})$.

3. The representative banking system maximizes its profits given in Equation (17), subject to its balance sheet identity in Equation (15) and the definitions of capital slack K in Equation (18) and liquidity slack Λ in Equation (19).

4. The non-bank credit intermediary maximizes its profits given in Equation (22).

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? Recall that for a given deposit inflow dD_0, the total credit effect is,

$$\Delta L_{total} = \lambda \cdot m_D \cdot \Delta D_0, \tag{24}$$

which depends on the credit-deposit multiplier

$$m_D = \frac{1}{1 - \lambda\theta}. \tag{25}$$

Equation (24) highlights that the total credit effect of deposit inflows depends on two distinct forces. The parameter λ governs how much credit is created per dollar of stable deposits, while θ determines the fraction of the resulting payment flows that settle back into the banking system as stable funding. Neither alone generates amplification: if $\theta = 0$, lending creates only flighty liabilities that never stabilize, so total stable deposits remain dD_0; if $\lambda = 0$, there is nothing to recycle. It is their interaction, summarized by the product $\lambda\theta$, that determines the strength of the multiplier and the degree to which deposit inflows are amplified into credit creation.

4.7.1 Equilibrium Lending Response λ

Banks' marginal lending response λ captures the direct effect, and is defined as the equilibrium derivative

$$\lambda = \frac{dL}{dD},$$

which we recover as an endogenous object from the bank's first order conditions. Intuitively, λ governs how strongly banks expand lending when deposits increase, taking into account the induced effects on wholesale funding and the capital and liquidity slacks.

We derive the analytical expression for λ in Appendix A.2. Under standard regularity conditions, the Implicit Function Theorem ensures that the local equilibrium mapping

$$x = (r_d, r_\ell, r_{nb}, W, D, L, L_{nb})$$

is differentiable with respect to any (small) shock, including shocks to bank funding conditions ($d\xi_b$ or dW_d). This differentiability allows us to take total differentials of the bank's balance sheet identity, liquidity constraint, capital constraint, loan-supply first-order condition, and the nonbank credit condition around the interior equilibrium, and to solve the resulting linear system for the local co-movements of dL, dD, dW, dL_{nb}, and $d\Lambda$. The resulting expression for the banking sector's equilibrium lending response is,

$$\lambda = \frac{dL}{dD} = \frac{\kappa_W + \frac{\tau\chi^2}{\Lambda^2}(1-\psi)}{\frac{1}{\tilde{\kappa}_\ell^{nb}} + \kappa_W + \frac{\tau\chi^2}{\Lambda^2} + \frac{\xi\kappa^2}{K^2}}. \tag{26}$$

Several economic forces shape this response. First, λ is larger when wholesale funding is expensive (high κ_W). In that case, deposits represent a relatively attractive source of funding, so an inflow translates more strongly into additional lending. Second, liquidity regulation plays a dual role. When deposits are more stable (low ψ), they provide greater liquidity relief compared to wholesale funding, which increases the lending response. At the same time, when liquidity is tight (low Λ or high τ), expanding loans raises the shadow cost of liquidity, dampening the response. The overall liquidity effect therefore depends on both deposit stability and how binding the liquidity constraint is.

Third, competition from nonbanks affects λ through $\tilde{\kappa}_\ell^{nb}$, which captures how responsive bank lending is to the bank loan rate once the nonbank sector adjusts endogenously.[9] When substitution between banks and nonbanks is strong and nonbank supply is relatively elastic, a given expansion in bank lending requires larger equilibrium movements in loan rates and induces greater crowding out of nonbank credit. In this case, the residual demand curve facing banks is flatter, $\tilde{\kappa}_{\ell}^{nb}$ is smaller, and the deposit–lending multiplier λ is lower. By contrast, when nonbank competition is weaker, banks can expand lending with smaller rate adjustments and less cross-sector reallocation, leading to a larger λ.

Finally, capital constraints discipline balance sheet expansion. When capital slack is limited (low K or high ξ and κ), additional lending tightens the leverage constraint and raises the marginal cost of credit supply, reducing λ. Overall, λ reflects the interaction of funding costs, regulatory constraints, and credit-market competition in determining how strongly deposit growth translates into bank lending.

4.7.2 Equilibrium Deposit Recycling θ

The recycling parameter θ captures the indirect effect through the fraction of payment flows generated by lending that are retained in the banking system as deposits. When the bank extends ΔL in new loans, borrowers spend the proceeds, creating new payment balances that circulate through the payment system.[10] A fraction θ of these circulating balances is retained as deposits, so that the recycling condition satisfies

$$dD = dD_0 + \theta\, dL. \tag{27}$$

Note that θ need not equal banks' deposit market share s_d. Balances created by lending circulate through the payment system for as long as the underlying loan remains outstanding, and can flow between accounts without permanently settling. The parameter θ captures the fraction that, at any given point in time, resides as stable deposits from the bank's perspective. This can fall below s_d for two reasons. First, lending-driven payment flows enter the economy as compensation for goods and services rather than as a portfolio allocation. Recipients may spend these balances

[9] This term is derived in the Appendix and captures how the bank considers their own price and cross price elasticity.

[10] Loans are long-maturity claims. Principal is not repaid within the period in which production and payments occur, so repayment and refinancing decisions are abstracted from in the cash-flow accounting. Non-bank lending ΔL_{nb} also finances payments in gross terms, but represents a reallocation of existing cash rather than a source of net new funding.

on their own obligations before any residual enters a savings decision, so that newly created balances circulate at a higher velocity than the existing deposit stock and a smaller fraction is resting as stable funding at any given moment. Second, marginal recipients of lending-driven payment flows—such as firms, institutional investors, or non-bank intermediaries—may have systematically different propensities to hold deposits than the average wealth holder, paralleling the distinction between marginal and average propensities to consume in fiscal policy.

It is worth emphasizing that θ is not a margin the banking system optimizes over. From the perspective of any individual bank, θ is an externality: when a bank extends a loan, the resulting payment flows circulate through the economy and may settle as stable deposits at other banks, or not return to the banking system at all. The originating bank neither captures nor internalizes this margin. Because no individual bank optimizes over θ, it does not enter the first-order conditions of the representative banking sector either. Unlike λ, which emerges endogenously from the bank's optimization, θ is a primitive of the environment that we take to the data in Section 5.

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 regulatory and demand parameters measured directly from balance sheet and pricing data with a GMM procedure that targets three sets of moment conditions. Each exploits a different equilibrium relationship to identify a distinct group of parameters: the banking sector's pricing conditions identify the shadow costs of illiquidity and leverage, the non-bank pricing condition identifies the cost of expanding non-bank credit supply, and the reduced-form multiplier estimates from Section 3 pin down the recycling rate of loan proceeds into stable deposits.

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. The sample spans 1999Q1–2024Q4. 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.3.1.

Two groups of parameters serve as inputs to the estimation, as reported in Table 4. First, the statutory regulatory parameters define the pre- and post-reform regimes, reflecting the substantial tightening of bank regulation following the 2008 financial crisis. In the pre-crisis period, capital requirements are light, reflecting the absence of a binding leverage ratio, and liquidity requirements are accommodating, leaving substantial liquidity slack. In the post-crisis period, capital requirements are tighter to capture the Supplementary Leverage Ratio, and liquidity requirements are more stringent to capture the Liquidity Coverage Ratio.[11] In particular, we set $\kappa_{\text{pre}} = 0.02$, and set our liquidity parameters to $\chi_{\text{pre}} = 0.15$ and $\psi_{\text{pre}} = 0.30$, which corresponds in the data to an average required buffer of roughly 9 percent of total assets against actual liquid holdings of about 22 percent — a liquidity slack of 13 percentage points. In contrast, we set $\kappa_{post} = 0.08$, and set our liquidity parameters to $\chi_{\text{post}} = 0.35$ and $\psi_{\text{post}} = 0.60$, which reduces liquidity slack to only 3 percentage points.[12] We plot the full quarterly series of the quantities K_t and Λ_t in Appendix Figure A.2.

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).[13]

[11] Note that the main capital regime in the pre-period was focused on risk-based capital requirements rather than simple leverage requirements, which became more binding in the post period (Walz, 2026). In this baseline version of the model we isolate the marginal increase in balance sheet costs rather than the effect of risk-based capital requirements which would be qualitatively similar.

[12] We define the capital requirement to be 8% in the post-period. This is greater than the "well capitalized" level specified by the Supplementary Leverage Ratio (Koont and Walz, 2021) due to the fact that the capital measure here is the difference between assets and liabilities, which is a broader measure of capital compared to Tier 1 capital in the SLR.

[13] The exact values for these price elasticities do not qualitatively change the model.

Table 4: Model Inputs

Panel A lists the statutory capital and liquidity parameters for the pre- and post-reform regimes. Panel B lists deposit and loan rate sensitivites.

A. Statutory Parameters

Symbol	Description	Pre-Reform	Post-Reform
κ	Capital requirement	0.020	0.080
χ	Liquidity ratio	0.150	0.350
ψ	Deposit runnability	0.300	0.600

B. Price Elasticities

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

5.2 Moment Conditions and Parameter Estimates

Given these inputs, we estimate the parameter vector $\Theta = \{\, \kappa_W,\ \tau,\ \xi,\ \kappa_{nb}^{pre},\ \kappa_{nb}^{post},\ \theta_{pre},\ \theta_{post}\,\}$ by just-identified GMM, using seven moment conditions derived from the quantitative version of the model in which reserves adjust endogenously as the residual of the balance sheet identity (Appendix A.2.2). In this specification, wholesale funding is an independent choice variable with its own optimality condition, and the bank's pricing conditions reflect the endogenous adjustment of the liquid asset position.

Four sets of moment conditions provide identification, detailed below. Table 5 Panel A reports the parameter estimates from the GMM estimation, and Panel B reports the model-implied credit-deposit multipliers alongside their instrumental variable counterparts. The model produces a pre-period deposit multiplier of 2.60 and a post-period multiplier of 1.30, closely matching the empirical estimates. This alignment provides quantitative discipline for the estimated cost, slack, and recycling parameters.

Shadow costs of illiquidity and leverage. The banking sector's deposit and loan pricing first-order conditions jointly identify τ and ξ. The loan spread $r_t^\ell - r_t^S$ reflects both the bank's market power over borrowers and the shadow cost of liquidity,

$$\frac{1}{T}\sum_{t=1}^{T}\left[(r_t^{\ell}-r_t^{S}) \;-\; \frac{1}{\alpha_\ell(1-s_t^L)} \;-\; \tau\,\frac{1}{\Lambda_t}\right]=0. \tag{28}$$

After subtracting the demand markup, which is pinned down by the input elasticity α_ℓ and the observed loan market share, and given Λ_t, this condition pins down τ directly. The deposit spread then identifies ξ. After subtracting the demand markdown and the liquidity wedge, the residual reflects the shadow cost of capital,

$$\frac{1}{T}\sum_{t=1}^{T}\left[(r_t^{S}-r_t^{d}) \;-\; \frac{1}{\alpha_d(1-s_t^d)} \;+\; \tau\,\frac{1-\chi_t\psi_t}{\Lambda_t} \;-\; \xi\,\frac{\kappa}{K_t}\right]=0. \tag{29}$$

Wholesale funding costs. The wholesale funding first-order condition identifies κ_W, the curvature of the cost of funding outside the deposit base:

$$\frac{1}{T}\sum_{t=1}^{T}\left[\kappa_W\,W_t \;-\; \frac{(1-\chi_t)\,\tau}{\Lambda_t} \;+\; \frac{\kappa\,\xi}{K_t}\right]=0. \tag{30}$$

At the optimum, the marginal cost of an additional dollar of wholesale funding equals its net regulatory benefit: wholesale funding provides liquidity by expanding reserves, but tightens the capital constraint. Given τ and ξ from the pricing conditions, this moment pins down the cost of non-deposit funding, which governs how aggressively banks substitute between deposits and wholesale liabilities when funding conditions change.

Non-bank intermediation costs. The non-bank pricing condition identifies κ_{nb}^{j}, the marginal cost of non-bank credit provision, separately in each regime:

$$\frac{1}{T_j}\sum_{t\in\mathcal{T}_j}\left[(r_t^{nb}-r_t^{S}) \;-\; \frac{1}{\alpha_\ell(1-s_t^{nb})} \;-\; \kappa_{nb}^{j}\,L_{nb,t}\right]=0, \quad j\in\{pre,\,post\}. \tag{31}$$

The logic parallels the bank side: after removing the demand markup, the residual non-bank loan spread must reflect intermediation costs. This parameter matters for the lending response λ because it governs competitive pressure in the loan market. When κ_{nb} is low, non-banks absorb a larger share of any shift in bank credit supply, flattening the residual demand curve facing banks and reducing the equilibrium lending response to deposit inflows.

Deposit recycling. The final two moments use the reduced-form multiplier estimates from Section 3 to discipline the deposit recycling rate θ:

$$\frac{1}{T_j} \sum_{t \in \mathcal{T}_j} \left[\hat{m}_t^j \; - \; \frac{1}{1 - \theta_j \, \lambda_t^j(\Theta)} \right] = 0, \quad j \in \{pre, \, post\}. \tag{32}$$

Given the parameters identified above, the model delivers a lending response λ_t in each quarter. The credit–deposit multiplier then depends on the product $\lambda_t \theta_j$: conditional on λ_t, matching the empirical multiplier pins down θ_j. The empirical multiplier reflects both the bank's optimizing behavior through λ, and the payment system's tendency to recycle loan proceeds into stable deposits through θ. By targeting the multiplier directly, we ensure that the model's two key sufficient statistics are disciplined by the same exogenous variation that underlies our instrumental variables estimates in Section 3.

Table 5: Model Estimates

Panel A reports GMM estimates with standard errors in parentheses. Panel B reports model-implied objects and compares to empirical counterparts from the IV estimation of Section 3.

A. Estimated Parameters

	Estimate	SE
Shadow costs		
τ Illiquidity	0.106	(0.062)
ξ Leverage	0.560	(0.290)
Wholesale funding		
κ_W	3.216	(1.244)
Non-bank costs		
κ_{nb}^{pre}	4.953	(0.829)
κ_{nb}^{post}	2.376	(0.836)
Deposit recycling		
θ_{pre}	0.919	(0.076)
θ_{post}	0.761	(0.215)

B. Model-Implied Objects

	Model	Data
Credit-Deposit Multiplier		
m_{pre}	2.58	2.60
m_{post}	1.33	1.30
Marginal Lending Response		
λ_{pre}	0.661	
λ_{post}	0.319	
Total Effect on Credit		
$\Delta L_{total,pre}$	1.709	
$\Delta L_{total,post}$	0.424	

5.3 The Total Effect of Deposit Flows on Credit

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?

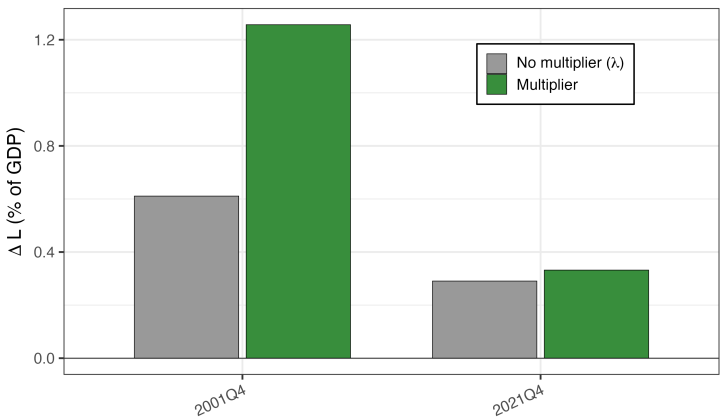
Table 5 reports the model-implied values for λ and θ. Banks' marginal lending response λ has declined from 0.66 to 0.32. In the earlier period, banks channeled roughly two-thirds of each marginal deposit dollar into new lending; in the later period, less than one-third supports credit expansion. The remainder is absorbed by reductions in wholesale funding or increases in liquid asset buffers. Deposits have moved from being primarily a source of credit creation to primarily a tool of balance sheet management.

The deposit recycling rate θ falls from 0.92 to 0.76. In the earlier period, over 90 cents of every dollar of loan-generated payment flows ultimately settled as stable deposits within the banking system, so that nearly all of the funds created by lending returned in a form that relaxed banks' liquidity constraints and could support further credit expansion. In the later period, only about 76 cents returns as stable deposits. This decline is consistent with a reduced specialness of bank deposits as a savings instrument, as households and firms increasingly allocate funds toward non-bank alternatives such as money market funds and other savings vehicles.

As a result, the total effect of a deposit inflow on bank credit, $\Delta L_{total} = \lambda \cdot m_D$, has declined from 1.73 to 0.42. Each dollar of deposits today generates less than a quarter of the credit it once did. Both sufficient statistics contribute to the decline: the weaker lending response reduces the credit generated in each round, while diminished deposit recycling weakens the amplification from one round to the next. Figure 5 illustrates this by comparing the effect from banks' marginal lending response to the total effect following a one percent of GDP deposit inflow in 2001Q4 and 2021Q4. In the earlier period, the multiplier more than doubles the initial credit response; by the later period, amplification is negligible. 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 50% decline, from 0.67 to 0.32. But the total credit effect has fallen by more than 75%, from 1.73 to 0.42. The difference reflects the compounding nature of the multiplier: when both λ and θ weaken simultaneously, the amplification mechanism collapses disproportionately.

Figure 5: First-Round and Total Credit Effects

The figure compares the first-round credit effect and the total credit effect of a one percent of GDP deposit inflow in 2001Q4 and 2021Q4. The first-round effect corresponds to banks' marginal lending response, $\lambda_{\text{total}}\Delta D$, holding deposits fixed and abstracting from recycling. The total effect incorporates full amplification through endogenous deposit re-depositing and multiplier dynamics, $m \times \lambda_{\text{total}}\Delta D$. Differences across periods reflect changes in both marginal pass-through and amplification strength.



6 Decomposing the Decline in Credit Creation

6.1 Drivers of the Decline

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 capital regulation, liquidity regulation, and 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 a tightening of the banks' regulatory environment and the specialness of deposits, as well as a transformation of the balance sheets themselves.

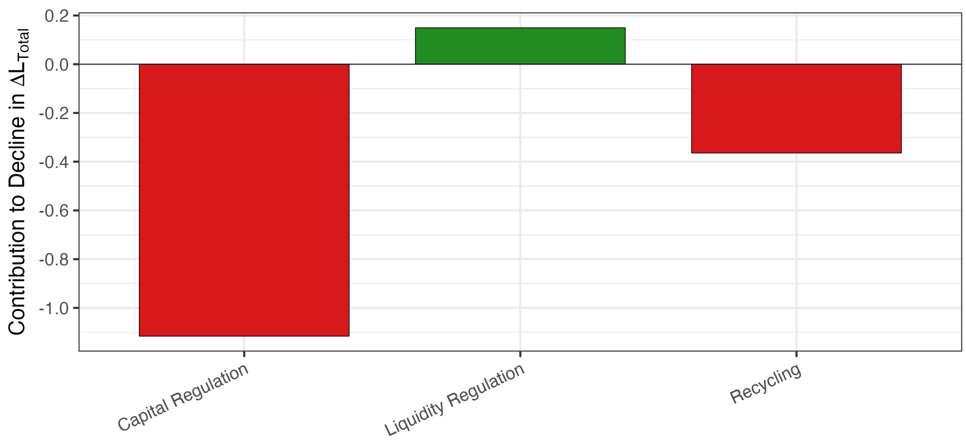
6.1.1 Changing Primitives: Regulation and Deposit Recycling

We begin by asking how much of the decline can be attributed to changes in primitives. Figure 6 reports the contribution of changes in primitives to the decline in the total lending response $\Delta L_{Total} = \lambda \cdot m$ from 2001Q4 to 2021Q4, expressed as a share of the total decline. Starting from the 2001Q4 economy, the bars show the marginal effect of substituting in turn (i) capital regulation κ, (ii) liquidity regulation χ and ψ, and (iii) the deposit recycling parameter θ.

We find that tightening capital regulation more than accounts for the decline in the lending response. On the other hand, changes in liquidity regulation offset part of the contraction. Stricter liquidity requirements reduce effective liquidity slack, which makes stable deposits more valuable at the margin as a source of funding for credit provision, and increases the sensitivity of lending to deposit inflows. Finally, the decline in deposit recycling can explain around one third of the decline in the total credit. This highlights that the diminished tendency for loan proceeds to return as stable deposits is a quantitatively important determinant of the decline in credit creation from deposit inflows.

Figure 6: Drivers of the Decline in Credit Creation

Each bar reports the contribution of a channel to the decline in the total lending response $\Delta L_{Total} = \lambda \cdot m$ from 2001Q4 to 2021Q4, expressed as a share of the total decline. The bars show the marginal effect of substituting in turn (i) capital regulation (κ), (ii) liquidity regulation (χ and ψ), and (iii) the deposit recycling parameter (θ).



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. First, quantitative Easing (QE), which flooded the banking system with deposits and reserves between 2008 and 2014 and again between 2020 and 2022. Next, Quantitative Tightening (QT), which began draining reserves in in 2018 and again in mid-2022. We find that both policies reduce the banking sector's propensity to transform deposit inflows into credit, despite operating in opposite directions on the balance sheet.

Quantitative Easing. When the Federal Reserve purchased Treasury securities and MBS from non-bank counterparties, it simultaneously created reserves on bank balance sheets and deposits in the sellers' bank accounts, expanding the banking system's balance sheet on both sides. By the end of our sample period, QE-created deposits averaged roughly 12% of GDP, a substantial fraction of the banking system's deposit base. These deposits are institutional in nature, landing in accounts of asset managers, pension funds, and dealers' clients, and are functionally closer to wholesale funding than to the stable core deposits that relax banks' liquidity constraints (Blickle et al., 2025). QE therefore expanded balance sheets while simultaneously tightening capital constraints and raising the runnability of the deposit base.

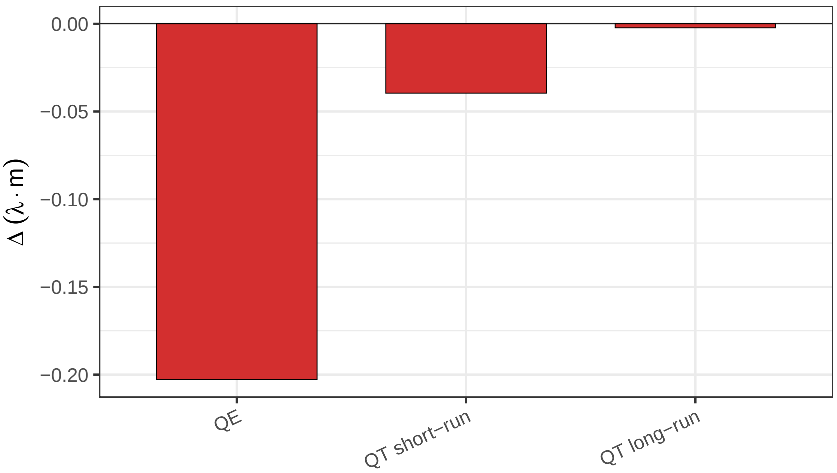
To quantify the effect, we construct a counterfactual banking system in which QE never occurred, stripping out QE-induced deposits and reserves and re-solving for equilibrium lending, non-bank lending, and wholesale funding. Details of the counterfactual construction are provided in Appendix A.4.2. Figure 7 shows that QE reduced the total credit effect $\lambda \cdot m_D$ by 0.20. The result is a larger but more constrained banking sector, with a lower propensity to transform deposit inflows into credit.

Quantitative Tightening. Beginning in mid-2022, the Federal Reserve began allowing its securities holdings to run off. As Treasury and MBS holdings mature without reinvestment, the Treasury reissues the debt to the market. To the extent that banks absorb this issuance, they pay with reserves, which are destroyed when the Treasury uses the proceeds to pay down its account at the Fed. In practice, this has largely manifested as an asset swap on bank balance sheets: reserves fall by X_{QT} and securities holdings rise by X_{QT}, with deposits unchanged (Acharya and Rajan, 2024). QT therefore drains reserves without removing deposits, replacing liquid reserves with less liquid securities on the asset side. This degrades the liquidity buffer without easing the capital constraint.

We report two counterfactuals: a short-run exercise in which the loan book is held fixed and wholesale funding adjusts, and a long-run exercise in which lending, non-bank lending, and wholesale funding all adjust to clear equilibrium conditions. Details are provided in Appendix A.4.2. Figure 7 reports the results. In the short run, QT modestly reduces the total credit effect, as constraints bind harder on both margins. In the long run, however, the effect largely disappears. Given time to adjust, banks shrink their loan books to restore regulatory slack, and a smaller balance sheet with the same equity provides more capital and liquidity headroom per

Figure 7: Effects of QE and QT

Each panel reports the change in the total lending response $\Delta L_{Total} = \lambda \cdot m$ induced by a balance sheet policy under pre-crisis and post-crisis regulation. Panel A evaluates the effect of QE-driven deposit inflows on the post-period banking sector. Panel B evaluates the effect of banks absorbing the decline in Federal Reserve securities holdings from 2022Q2 to 2024Q4.



dollar of lending. QT's effect on the lending response is thus primarily transitional: the banking sector absorbs the reserve drain by scaling down, with little lasting impact on its marginal responsiveness to deposit inflows.

Notably, both QE and QT reduce the total credit effect despite operating in opposite directions on the balance sheet. In practice, QE expanded liabilities by adding deposits and reserves, tightening capital constraints while providing some liquidity relief. QT shrank the liquid asset base by replacing reserves with less liquid securities, degrading the liquidity buffer without easing the capital constraint. Together, the two policies dampen the transmission of deposit inflows into credit by tightening the capital and liquidity constraints respectively.

6.2 Implications for Aggregate Credit

The total effect of deposit inflows on bank credit has declined from 1.73 to 0.42. 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 from two angles. We first examine the direct competitive response of non-banks to a deposit shock, asking how much of the lost bank credit is mechanically picked up by non-bank lenders in equilibrium. We then take a broader view that recognizes two complications: deposit inflows themselves are smaller once the multiplier weakens, and non-bank intermediation costs may have fallen independently in ways that partially offset the decline. Within this broader view, we ask whether non-bank costs have in fact declined, and how large a cost improvement would be needed to fully offset the contraction of the bank channel.

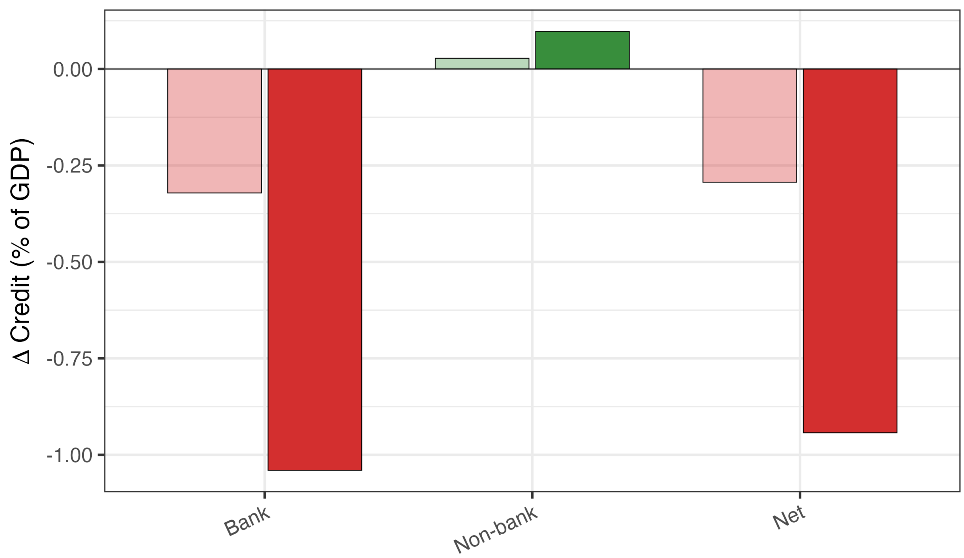
6.2.1 Non-Bank Competitive Response

We begin with the direct margin: how non-bank lenders respond in equilibrium to a deposit inflow that hits the banking system. Figure 8 reports the change in the equilibrium credit response to a one percent of GDP deposit inflow between the early period (2001Q4) and the late period (2021Q4), decomposed into bank credit, non-bank credit, and net credit. Each pair of bars shows the late-minus-early gap for one of these three objects: the lighter bar isolates the first-round effect through banks' marginal lending response λ, while the more opaque bar incorporates the full credit-deposit multiplier and shows the total effect $\lambda \cdot m_D$. The first pair of bars captures the central finding of the paper: bank credit expands substantially less in the late period than in the early period, with the gap widening once the multiplier is taken into account. The second pair shows the corresponding gap in non-bank credit, and the third pair reports the gap in net credit. Comparing the non-bank bars to the bank bars within each specification isolates how much of the weakening of the bank channel is offset by changes in non-bank lending. The figure shows that this offset is small: the non-bank bars are an order of magnitude smaller than the bank bars, and the net credit gap tracks the bank gap almost one-for-one. The competitive response of non-banks to a deposit shock leaves the decline in the bank channel essentially intact in aggregate credit.

The economics behind Figure 8 is as follows. When banks expand lending in response to a deposit inflow, they capture borrowers from non-banks, so non-bank credit contracts. In the post period, banks expand lending by less, so non-banks contract by less—and this smaller contraction shows up as a positive non-bank offset to aggregate credit when comparing the two periods. The magnitude of the displacement in each period is governed by two forces: the substitutability of bank and non-bank credit in borrower demand, and the elasticity of non-bank supply. When borrowers do not view the two as perfect substitutes, or when non-bank supply is not perfectly elastic—for instance, because rising intermediation costs push r^{nb} up as L_{nb} expands and deter marginal borrowers—displacement is incomplete. In our calibration, the dis-

Figure 8: Effect on Aggregate Credit: Late vs. Early Period

The figure reports the change in the equilibrium credit response to a one percent of GDP deposit inflow between the early period (2001Q4) and the late period (2021Q4), decomposed into bank credit, non-bank credit, and net credit. Each pair of bars shows the late-minus-early gap for one of these three objects. Within each pair, the lighter bar isolates the first-round effect operating through banks' marginal lending response λ, while the darker bar incorporates the full credit-deposit multiplier and reports the total effect $\lambda \cdot m_D$. Negative values indicate that the credit response is smaller in the late period than in the early period.



placement rate is roughly 10%, so most of the change in bank credit is absorbed by the outside option of not borrowing rather than by non-bank lenders, and the non-bank offset to aggregate credit is correspondingly small.

This logic also clarifies an important asymmetry in how the competitive channel can "close the gap" between the early and late periods. Our motivating question is whether advancements in the non-bank sector can offset the loss of the bank amplifier in the post period—that is, whether more elastic or more competitive non-banks today can deliver the credit creation that the bank channel no longer provides. The competitive response cannot accomplish this. The reason is that non-banks respond to a deposit inflow by contracting, not expanding: the deposit shock reaches them only indirectly, through the fall in r^ℓ as banks expand, which pulls borrowers away from non-bank credit. A more elastic or more substitutable non-bank sector would only deepen this contraction in the post period, making the gap with the early period wider, not narrower. For non-banks to genuinely compensate for the lost amplification, they must operate through a margin other than the competitive response to a deposit shock, which motivates the analysis of non-bank intermediation costs in the next subsection.

6.2.2 Non-Bank Intermediation Costs

The competitive response of non-banks cannot offset the decline in bank credit creation, but the non-bank sector might still compensate through a different margin: independent growth in non-bank lending capacity. This possibility is worth taking seriously for two reasons. First, the non-bank sector has expanded considerably over recent decades, and it is plausible that improvements in non-bank intermediation technology have made non-bank credit cheaper to supply at the margin. Second, the exercise of comparing the credit response to a given deposit shock across periods may understate the true gap in credit creation, because the weakening of the bank multiplier also means that a smaller volume of deposit inflows now reaches the banking system in the first place. A broader analysis that allows for independent non-bank growth therefore captures part of the picture that the competitive response exercise misses.

Not all forms of non-bank growth, however, can substitute for the lost bank amplification. Because bank credit L includes securities held on bank balance sheets, non-bank credit originated through an originate-to-distribute (OTD) model and ultimately sold to banks is captured within L rather than within non-bank credit (Buchak et al., 2024). Such credit is subject to the same dampened multiplier as any other bank asset: it still relies on the bank balance sheet for funding and faces the same capital and liquidity wedges, so an expansion of OTD origination does not bypass the bank amplification channel but merely repackages credit that flows through it. For non-bank growth to genuinely offset the decline in bank-based credit creation, it must therefore come from credit that is held off bank balance sheets, such as direct lending by private credit funds or assets retained within private equity vehicles.

For our measure of non-bank credit, we find little evidence of faster growth in recent years relative to the early period. To examine this directly, we decompose the structural non-bank cost parameter κ_{nb} into an intermediation cost component κ_{int} and a residual that absorbs common macro forces, as detailed in Appendix A.4.3. This decomposition exploits the relative spread $r^{nb} - r^{\ell}$, which strips out credit risk conditions and other macro factors that affect bank and non-bank pricing symmetrically. We find that κ_{int} has remained roughly stable across the two periods: the decline in κ_{nb} between the early and late periods reflects common macro forces and the tightening of the bank regulatory wedge, not an improvement in non-bank supply capacity.

Our measure of non-bank credit may not fully capture recent growth in the sector, particularly the expansion of off-balance-sheet vehicles such as private equity and private credit. To allow for this possibility, we now ask, starting from the post-period economy, how much would non-bank intermediation costs need to fall to generate enough non-bank credit expansion to offset the bank channel decline? For each counterfactual value of κ_{int}, we solve for the equilibrium rates and compute the resulting change in total credit relative to the baseline. We find that a 4.3% reduction in κ_{int} generates sufficient non-bank credit expansion to offset the first-round bank lending decline in λ. However, offsetting the full decline, including the reduced multiplier m, requires an 11.8% reduction, nearly three times as large. This ratio quantifies the asymmetry between replacing first-round bank lending and replacing the amplification mechanism. Replacing the amplification decline is disproportionately harder because, in the earlier period, each successive round of credit arises endogenously through deposit recycling at no additional cost. Non-bank supply expansion requires independent cost improvements to generate credit that the banking system previously created for free through compounding.

This asymmetry is compounded by the cyclical nature of the deposit multiplier. Each episode of deposit inflows, from monetary easing, fiscal transfers, or portfolio reallocation, triggers successive rounds of credit creation through the recycling mechanism. The weakened multiplier reduces the credit response to each such episode. A reduction in non-bank intermediation costs raises the level of non-bank credit but does not make non-bank credit responsive to subsequent deposit inflows. To the extent that deposit-driven credit amplification recurs over the credit cycle, non-bank intermediation costs would need to decline continually to offset successive episodes of multiplier underperformance.

In sum, the competitive response of non-banks to a deposit shock is small and, by its nature, cannot offset the loss of the bank amplifier. Independent growth in non-bank capacity could in principle close the gap, but the relevant growth must occur off bank balance sheets—credit that flows back onto the bank side through originate-to-distribute channels remains subject to the same dampened multiplier. And even when non-bank growth is genuinely off balance sheet, offsetting the full decline requires a cost improvement roughly three times as large as what would be needed to replace the first-round bank lending response, with that improvement recurring each time deposit-driven amplification would otherwise have been activated. The weakening of the bank channel is therefore difficult to undo through non-bank substitution along any of the margins we consider.

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 a quarter 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 by more than half. Tighter capital regulation has made banks less willing to expand lending against any given funding base, and the diminished specialness of bank deposits means that fewer loan proceeds return as stable funding to support further credit creation. When both margins weaken simultaneously, the compounding nature of the multiplier causes the amplification mechanism to collapse.

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 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, "Banking on Deposits: Maturity Transformation without Interest Rate Risk," *The Journal of Finance*, 2021, *76* (3), 1091–1143.

_ , _ , 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. and Jeremy C. Stein, "The impact of monetary policy on bank balance sheets," *Carnegie-Rochester Conference Series on Public Policy*, 1995, *42*, 151–195.

_ **, 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 and Stefan Walz, "Bank credit provision and leverage constraints: Evidence from the supplementary leverage ratio," Mar 2021.

_ **, 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.

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

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).

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

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.

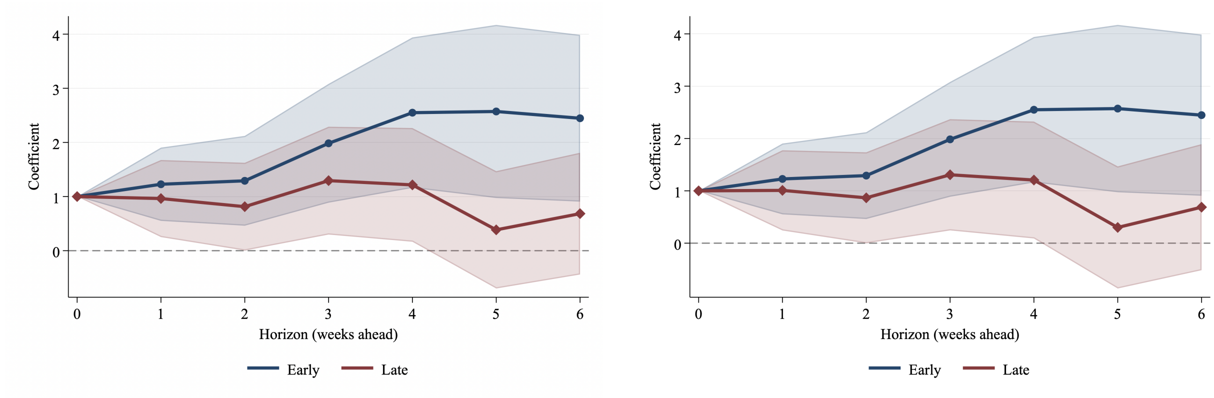
Non-bank Multiplier. For long-term credit provision to the real economy—the object of interest in this paper—we therefore treat $\theta_{nb} \approx 0$, so that the non-bank multiplier $m_{nb} = 1/(1 - \lambda_{nb}\theta_{nb}) \approx 1$. The credit-deposit feedback loop is a feature specific to the 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.1.2 Robustness Tests

For our Treasury cash flow innovation shock for deposits, a concern may be that in the post-crisis period, Treasury cash flows may affect reserves through movements in the Treasury's account at the Federal Reserve. This could generate co-movement between deposits and reserves that is not driven by deposit inflows. To test whether reserve movements are a driver of our results, we re-estimate the late period estimate while controlling for the lagged level of aggregate reserves, R_{t-1}, in Panel (a) of Figure A.1, and the simultaneous change in aggregate reserves, ΔR_t, in Panel (b) of Figure A.1. The results remain very similar, indicating that the estimates primarily reflect deposit inflows rather than reserve co-movement at weekly frequency.

Figure A.1: Deposit Multiplier Estimates: Robustness to Changes in Reserves

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. It compares estimates for the early (2005-2011) versus late (2012-2019) samples. Shaded bands denote 95% confidence intervals.



(a) Level of Reserves (b) Changes in Reserves

A.2 Model Derivations

A.2.1 Baseline Model

A.2.1.1 Bank First Order Conditions

$$\Pi = (r^\ell - r^S)L \;+\; (r^S - r^d)D \;-\; \frac{\kappa_W}{2}W^2 \;+\; \xi \ln\big(K\big) \;+\; \tau \ln\big(\Lambda\big) \tag{1}$$

Re-writing the capital and liquidity slack in terms of choice variables of the bank, using the balance-sheet identity to substitute out wholesale funding,

$$K \;=\; E - \kappa(L + R) \tag{2}$$

$$\Lambda = \chi E + (1-\chi)R - \chi L + \chi(1-\psi)D \tag{3}$$

Loan Rate r^ℓ. The total derivative of profits along bank's optimal responses:

$$\frac{d\Pi}{dr^\ell} = \frac{d\Pi}{dr^\ell} + \underbrace{\frac{d\Pi}{dr^d}}_{=0}\frac{dr^d}{dr^\ell} \tag{4}$$

Via the envelope theorem, the cross terms drop. Thus,

$$\begin{aligned}\frac{d\Pi}{dr^\ell} &= (r^\ell - r^s)\frac{dL}{dr^\ell} + L - \kappa_W W\frac{dW}{dL}\frac{dL}{dr^\ell} + \frac{\xi}{K}\frac{dK}{dL}\frac{dL}{dr^\ell} + \tau\frac{1}{\Lambda}\frac{d\Lambda}{dL}\frac{dL}{dr^\ell} \\ &= (r^\ell - r^s)\frac{dL}{dr^\ell} + L - \kappa_W W\frac{dL}{dr^\ell} - \kappa\frac{\xi}{K}\frac{dL}{dr^\ell} - \chi\frac{\tau}{\Lambda}\frac{dL}{dr^\ell}\end{aligned} \tag{5}$$

where we used

$$\frac{dW}{dL} = 1, \qquad \frac{dK}{dL} = -\kappa, \qquad \frac{d\Lambda}{dL} = -\chi.$$

The first order condition sets this equal to zero:

$$0 = (r^\ell - r^s)\frac{dL}{dr^\ell} + L - \kappa_W W\frac{dL}{dr^\ell} - \kappa\xi\frac{1}{K}\frac{dL}{dr^\ell} - \chi\frac{\tau}{\Lambda}\frac{dL}{dr^\ell} \tag{6}$$

Solving for r^ℓ:

$$r^\ell = r^s + \kappa_W W - \frac{L}{\frac{dL}{dr^\ell}} + \kappa\xi\frac{1}{K} + \chi\frac{\tau}{\Lambda} \tag{7}$$

As before,

$$\frac{dL}{dr^\ell} = -\alpha^\ell(1 - s^\ell)L \tag{8}$$

Substituting back yields

$$r^\ell = r^s + \kappa_W W + \frac{1}{\alpha^\ell(1 - s^\ell)} + \kappa\xi\frac{1}{K} + \chi\frac{\tau}{\Lambda} \tag{9}$$

Deposit Rate r^d The total derivative of profits along bank's optimal responses:

$$\frac{d\Pi}{dr^d} = \frac{d\Pi}{dr^d} + \underbrace{\frac{d\Pi}{dr^\ell}}_{=0}\frac{dr^\ell}{dr^d} \tag{10}$$

Via the envelope theorem, the cross terms drop. Thus,

$$\begin{aligned}\frac{d\Pi}{dr^d} &= (r^s - r^d)\frac{dD}{dr^d} - D - \kappa_W W\frac{dW}{dD}\frac{dD}{dr^d} + \frac{\xi}{K}\frac{dK}{dD}\frac{dD}{dr^d} + \frac{\tau}{\Lambda}\frac{d\Lambda}{dD}\frac{dD}{dr^d} \\ &= (r^s - r^d)\frac{dD}{dr^d} - D + \kappa_W W\frac{dD}{dr^d} + \chi(1-\psi)\frac{\tau}{\Lambda}\frac{dD}{dr^d}\end{aligned} \tag{11}$$

where we used

$$\frac{dW}{dD} = -1, \qquad \frac{dK}{dD} = 0, \qquad \frac{d\Lambda}{dD} = \chi(1-\psi).$$

Setting equal to zero:

$$0 = (r^s - r^d)\frac{dD}{dr^d} - D + \kappa_W W\frac{dD}{dr^d} + \chi(1-\psi)\frac{\tau}{\Lambda}\frac{dD}{dr^d} \tag{12}$$

Using $\frac{dD}{dr^d} = \alpha^d(1-s^d)D$, we obtain

$$r^d = r^s + \kappa_W W - \frac{1}{\alpha^d(1-s^d)} + \chi(1-\psi)\frac{\tau}{\Lambda} \tag{13}$$

A.2.1.2 Derivation of $\lambda = \frac{dL}{dD}$

In this section we derive the pass-through $\lambda = \frac{dL}{dD}$, the change in loans L generated by a marginal change in deposits D, evaluated at the bank's optimum. Three ingredients determine λ: the balance sheet identity, the capital and liquidity identities, and the loan pricing condition (with non-banks).

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

(i) **Interiority:** the optimum lies in the smooth region, with $K > 0$, $\Lambda > 0$, $W > 0$, and 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 ensures a unique, differentiable local equilibrium map $x(z)$, so it is valid to take total differentials of the equilibrium system and solve for the local co-movements dx for any small driver z.[14] Thus, we take total differentials in turn of each of the model's key balance sheet definitions and optimality conditions below to obtain the linearized relations among $(d\Lambda, dK, dL, dD, dW, dL_{nb})$.

$$
\begin{aligned}
L + R &= E + D + W, && \text{(balance sheet)} \quad (14)\\
\Lambda &= R - \chi(\psi D + W), && \text{(liquidity identity)} \quad (15)\\
K &= E - \kappa(L + R), && \text{(capital identity)} \quad (16)\\
r_\ell &= r_S + \frac{1}{\alpha^\ell(1 - s_\ell)} + \kappa_W W + \chi\,\frac{\tau}{\Lambda} + \kappa\,\frac{\xi}{K}, && \text{(bank loan pricing)} \quad (17)\\
r_{nb} &= r_S + \kappa_{nb} L_{nb} + \frac{1}{\alpha^\ell(1 - s_{nb})}, && \text{(non-bank pricing)} \quad (18)
\end{aligned}
$$

In (14)–(16), reserves R are exogenous and fixed, so $dR = 0$ in the comparative statics below.

(a) Balance sheet identity. Differentiating (14) gives,

$$dW = dL - dD. \tag{19}$$

Intuitively, with reserves fixed, an increase in loans must be financed either by additional deposits or by wholesale funding.

(b) Liquidity identity. Differentiating (19) gives,

$$d\Lambda = -\,\chi\big(\psi\, dD + dW\big). \tag{20}$$

Substituting in dW from (19) yields,

$$d\Lambda = -\,\chi\Big(dL + (\psi - 1)\, dD\Big). \tag{21}$$

[14] For instance, a change in the policy r_S or a non-price deposit shifter Z that enters only the deposit demand identity $D = D^d(r_d, r_S; \theta_D, Z)$.

Thus, loan expansion tightens liquidity through its funding needs, while deposits relax liquidity to the extent that they are less runnable than wholesale funding ($\psi < 1$).

(c) Capital identity. Differentiating (16) gives,

$$dK = -\kappa \, dL. \tag{22}$$

Intuitively, with reserves fixed, expanding loans increases total assets one-for-one and tightens the leverage constraint in proportion to κ.

(d) Loan pricing with non-banks. With non-bank credit, bank loan demand depends on both loan rates, $L = L(r_\ell, r_{nb})$, and the non-bank loan rate r_{nb} is pinned down by (18). We therefore differentiate the two pricing conditions (17)–(18) jointly and use logit semi-elasticities to express dL as a function of the wedges $(dW, d\Lambda, dK)$.

Differentiating (17) gives

$$dr_\ell = \frac{1}{\alpha^\ell} \frac{1}{(1-s_\ell)^2} \, ds_\ell + \kappa_W \, dW - \chi \, \frac{\tau}{\Lambda^2} \, d\Lambda - \kappa \, \frac{\xi}{K^2} \, dK. \tag{23}$$

Under multinomial logit,

$$ds_\ell = -\alpha^\ell s_\ell (1 - s_\ell) \, dr_\ell + \alpha^\ell s_\ell s_{nb} \, dr_{nb}, \tag{24}$$

so substituting (24) into (23) yields

$$dr_\ell = \frac{s_\ell s_{nb}}{1 - s_\ell} \, dr_{nb} + (1 - s_\ell) \Big(\kappa_W \, dW - \chi \, \frac{\tau}{\Lambda^2} \, d\Lambda - \kappa \, \frac{\xi}{K^2} \, dK \Big). \tag{25}$$

Differentiating (18) gives

$$dr_{nb} = \kappa_{nb} \, dL_{nb} + \frac{1}{\alpha^\ell} \frac{1}{(1 - s_{nb})^2} \, ds_{nb}. \tag{26}$$

Using logit semi-elasticities,

$$dL_{nb} = -\alpha^\ell(1-s_{nb})L_{nb}\,dr_{nb} + \alpha^\ell s_\ell L_{nb}\,dr_\ell, \tag{27}$$
$$ds_{nb} = -\alpha^\ell s_{nb}(1-s_{nb})\,dr_{nb} + \alpha^\ell s_{nb}s_\ell\,dr_\ell, \tag{28}$$

and substituting (27)–(28) into (26) implies

$$\Big(1+\kappa_{nb}\alpha^\ell(1-s_{nb})L_{nb} + \frac{s_{nb}}{1-s_{nb}}\Big)dr_{nb} = s_\ell\Big(\kappa_{nb}\alpha^\ell L_{nb} + \frac{s_{nb}}{(1-s_{nb})^2}\Big)dr_\ell. \tag{29}$$

Finally, bank loan demand satisfies

$$dL = -\alpha^\ell(1-s_\ell)L\,dr_\ell + \alpha^\ell s_{nb}L\,dr_{nb}. \tag{30}$$

Combining (25), (29), and (30) yields a linear relation of the form

$$dL = -\,\tilde{\kappa}_\ell^{nb}\left(\kappa_W\,dW - \chi\,\frac{\tau}{\Lambda^2}\,d\Lambda - \kappa\,\frac{\xi}{K^2}\,dK\right), \tag{31}$$

where the coefficient $\tilde{\kappa}_\ell^{nb} > 0$ is given explicitly by

$$\tilde{\kappa}_\ell^{nb} = \alpha^\ell L(1-s_\ell)^2 \\ \times \frac{(1-s_\ell)\Big(1+\kappa_{nb}\alpha^\ell(1-s_{nb})L_{nb}+\frac{s_{nb}}{1-s_{nb}}\Big) -\; s_\ell s_{nb}\Big(\kappa_{nb}\alpha^\ell L_{nb}+\frac{s_{nb}}{(1-s_{nb})^2}\Big)}{(1-s_\ell)\Big(1+\kappa_{nb}\alpha^\ell(1-s_{nb})L_{nb}+\frac{s_{nb}}{1-s_{nb}}\Big) -\; s_\ell^2 s_{nb}\Big(\kappa_{nb}\alpha^\ell L_{nb}+\frac{s_{nb}}{(1-s_{nb})^2}\Big)}. \tag{32}$$

When the non-bank rate is held fixed (or $s_{nb} = 0$), the fraction in (32) equals one and (31) would include only the banking sector endogenous response.

(e) Combine to solve for λ. We now substitute the balance-sheet relation (19), the liquidity differential (62), and the capital differential (63) into (31).

First, substituting (19), (62), and (63) into the wedge term gives

$$\begin{aligned}\kappa_W\,dW - \chi\,\frac{\tau}{\Lambda^2}\,d\Lambda - \kappa\,\frac{\xi}{K^2}\,dK &= \kappa_W(dL-dD) - \chi\,\frac{\tau}{\Lambda^2}\Big(-\chi(dL+(\psi-1)dD)\Big) - \kappa\,\frac{\xi}{K^2}\big(-\kappa dL\big)\\ &= \Big(\kappa_W + \frac{\tau\chi^2}{\Lambda^2} + \frac{\xi\kappa^2}{K^2}\Big)dL - \Big(\kappa_W - \frac{\tau\chi^2}{\Lambda^2}(\psi-1)\Big)dD. \end{aligned} \tag{33}$$

Then (31) becomes

$$dL = -\,\tilde{\kappa}_\ell^{nb}\Big(\kappa_W + \frac{\tau\chi^2}{\Lambda^2} + \frac{\xi\kappa^2}{K^2}\Big)dL + \,\tilde{\kappa}_\ell^{nb}\Big(\kappa_W - \frac{\tau\chi^2}{\Lambda^2}(\psi-1)\Big)dD. \tag{34}$$

Re-arranging and solving for $\lambda = \dfrac{dL}{dD}$ yields

$$\begin{aligned}\lambda = \frac{dL}{dD} &= \frac{\tilde{\kappa}_\ell^{nb}\Big(\kappa_W - \frac{\tau\chi^2}{\Lambda^2}(\psi-1)\Big)}{1+\tilde{\kappa}_\ell^{nb}\Big(\kappa_W + \frac{\tau\chi^2}{\Lambda^2} + \frac{\xi\kappa^2}{K^2}\Big)} \\ &= \frac{\Big(\kappa_W + \frac{\tau\chi^2}{\Lambda^2}(1-\psi)\Big)}{\frac{1}{\tilde{\kappa}_\ell^{nb}} + \Big(\kappa_W + \frac{\tau\chi^2}{\Lambda^2} + \frac{\xi\kappa^2}{K^2}\Big)}. \end{aligned} \tag{35}$$

A.2.1.3 Derivation of $\mathcal{W}_D = \frac{dW}{dD}$

We again study small shocks along an interior optimum x^*. With reserves R fixed, the balance sheet identity is

$$L + R = E + D + W. \tag{36}$$

Taking total differentials and using $dR = 0$ and $dE = 0$ yields

$$dL = dD + dW, \tag{37}$$

or equivalently,

$$dW = dL - dD. \tag{38}$$

Dividing both sides of (38) by dD gives

$$\begin{aligned}\mathcal{W}_D \;\equiv\; \frac{dW}{dD} &= \frac{dL}{dD} - 1 \\ &= \lambda - 1. \end{aligned} \tag{39}$$

Thus, when reserves are fixed, wholesale funding adjusts residually to satisfy the balance sheet identity: each additional unit of lending not financed by deposits must be financed one-for-one by wholesale funding.

A.2.2 Quantitative Model

In the main text, we present the model holding central bank reserves R fixed. In this appendix, we generalize the model to allow the banking sector's liquid asset position to adjust endogenously as the residual of its balance sheet optimization. We redefine R to denote the banking sector's total cash and reserve balances, determined residually by the balance sheet identity $L + R = E + D + W$.

We define R broadly to include not only central bank reserves and vault cash but also interbank claims such as fed funds sold and reverse repo. Although interbank claims cancel in aggregate, they reflect an important liquidity management margin: individual banks that lose reserves through lending can restore their liquid asset positions by borrowing in the fed funds market or selling reverse repo. Restricting R to central bank reserves alone would understate the effective liquidity capacity of the banking system, because it would ignore the ability of individual banks to access liquid resources through the interbank market. The broader definition captures this margin in reduced form. In aggregate, interbank activity inflates both R and wholesale funding W by equal amounts. A dollar of gross interbank activity adds one dollar to R but tightens the liquidity constraint by only $\chi < 1$ dollars through W, so the net effect is to raise liquidity slack Λ by $1-\chi$. The broader definition of R therefore captures the additional effective liquidity that interbank markets provide to the banking system.

With R as the residual, wholesale funding W becomes an independent choice variable. The representative banking sector chooses the loan rate r^ℓ, the deposit rate r^d, and wholesale funding W to maximize

$$\Pi = (r^\ell - r^S)L(r^\ell) \; + \; (r^S - r^d)D(r^d) \; - \; \frac{\kappa_W}{2}W^2 \; + \; \xi \ln\big(K\big) \; + \; \tau \ln\big(\Lambda\big), \tag{40}$$

taking r^S and equity E as given. Re-writing the capital and liquidity slack in terms of choice variables of the bank. Using the balance sheet identity

$$L + R = E + D + W \qquad \Rightarrow \qquad R = E + D + W - L, \tag{41}$$

we can substitute out reserves in the liquidity slack:

$$K \;=\; (1-\kappa)E \;-\; \kappa(D+W), \tag{42}$$

$$\begin{aligned}\Lambda &= R - \chi(\psi D + W) \\ &= \big(E + D + W - L\big) - \chi(\psi D + W) \\ &= E - L + (1-\chi\psi)D + (1-\chi)W. \end{aligned} \tag{43}$$

Using (42) and (43), the bank chooses (r^ℓ, r^d, W) taking (r^S, E) as given.

A.2.2.1 Bank First Order Conditions

Loan Rate r^ℓ Differentiating profits with respect to r^ℓ yields

$$0 = \frac{\partial \Pi}{\partial r^\ell} = (r^\ell - r^S)\frac{dL}{dr^\ell} + L + \tau \frac{1}{\Lambda}\frac{\partial \Lambda}{\partial L}\frac{dL}{dr^\ell}. \tag{44}$$

From (43), $\frac{\partial \Lambda}{\partial L} = -1$, so

$$0 = (r^\ell - r^S)\frac{dL}{dr^\ell} + L - \frac{\tau}{\Lambda}\frac{dL}{dr^\ell}, \tag{45}$$

which implies

$$r_\ell = r_S - \frac{L}{\frac{dL}{dr^\ell}} + \frac{\tau}{\Lambda}. \tag{46}$$

Under logit loan demand, $\frac{dL}{dr^\ell} = -\alpha_\ell(1-s_\ell)L$, so the loan pricing equation is

$$r_\ell = r_S + \frac{1}{\alpha_\ell(1-s_\ell)} + \frac{\tau}{\Lambda}. \tag{47}$$

Deposit Rate r^d Differentiating profits with respect to r^d yields

$$0 = \frac{\partial \Pi}{\partial r^d} = (r^S - r^d)\frac{dD}{dr^d} - D + \xi \frac{1}{K}\frac{\partial K}{\partial D}\frac{dD}{dr^d} + \tau \frac{1}{\Lambda}\frac{\partial \Lambda}{\partial D}\frac{dD}{dr^d}. \tag{48}$$

From (42), $\frac{\partial K}{\partial D} = -\kappa$, and from (43), $\frac{\partial \Lambda}{\partial D} = 1 - \chi\psi$, so

$$0 = (r^S - r^d)\frac{dD}{dr^d} - D - \kappa\frac{\xi}{K}\frac{dD}{dr^d} + (1 - \chi\psi)\frac{\tau}{\Lambda}\frac{dD}{dr^d}, \tag{49}$$

which implies

$$r_d = r_S - \frac{D}{\frac{dD}{dr^d}} - \kappa\frac{\xi}{K} + (1 - \chi\psi)\frac{\tau}{\Lambda}. \tag{50}$$

Under logit deposit demand, $\frac{dD}{dr^d} = \alpha_d(1 - s_d)D$, so the deposit pricing equation is

$$r_d = r_S - \frac{1}{\alpha_d(1 - s_d)} - \kappa\frac{\xi}{K} + (1 - \chi\psi)\frac{\tau}{\Lambda}. \tag{51}$$

Wholesale Funding W Differentiating profits with respect to W yields

$$0 = \frac{\partial \Pi}{\partial W} = -\kappa_W W + \xi\frac{1}{K}\frac{\partial K}{\partial W} + \tau\frac{1}{\Lambda}\frac{\partial \Lambda}{\partial W}. \tag{52}$$

From (42), $\frac{\partial K}{\partial W} = -\kappa$, and from (43), $\frac{\partial \Lambda}{\partial W} = 1 - \chi$, so

$$0 = -\kappa_W W - \kappa\frac{\xi}{K} + (1 - \chi)\frac{\tau}{\Lambda}. \tag{53}$$

Equivalently, the wholesale funding condition is

$$\kappa_W W = \frac{(1 - \chi)\tau}{\Lambda} - \frac{\kappa\xi}{K}. \tag{54}$$

A.2.2.2 Derivation of $\lambda = \dfrac{dL}{dD}$

In this section we derive the pass-through

$$\lambda = \frac{dL}{dD},$$

the change in bank lending L generated by a marginal change in deposits D, evaluated at the equilibrium of the model. Three ingredients determine λ: the balance sheet identity, the capital and liquidity identities, and the joint bank–nonbank pricing system. We study small shocks around an interior equilibrium

$$x^* = (r_d^*, r_\ell^*, r_{nb}^*, R^*, W^*, D^*, L^*, L_{nb}^*),$$

in which capital slack $K > 0$, liquidity slack $\Lambda > 0$, wholesale funding $W > 0$, and all market shares lie strictly between zero and one. Because the equilibrium is interior and the system of pricing conditions, demand equations, and balance sheet identities is locally well behaved, small shocks to exogenous drivers induce well-defined local co-movements among the endogenous variables. We therefore take total differentials of the equilibrium system and solve for the implied linear relations among

$$(d\Lambda, dK, dR, dL, dD, dW, dr_d, dr_\ell, dr_{nb}, dL_{nb}).$$

The key equilibrium conditions are:

$$\begin{aligned}
L + R &= E + D + W, && \text{(balance sheet identity)} && (55)\\
\Lambda &= R - \chi(\psi D + W), && \text{(liquidity identity)} && (56)\\
K &= (1-\kappa)E - \kappa(D + W), && \text{(capital identity)} && (57)\\
\kappa_W W &= \frac{\tau(1-\chi)}{\Lambda} - \frac{\xi\kappa}{K}, && \text{(wholesale FOC)} && (58)\\
r_\ell &= r_S + \frac{1}{\alpha^\ell(1-s_\ell)} + \frac{\tau}{\Lambda}, && \text{(bank loan pricing)} && (59)\\
r_d &= r_S - \frac{1}{\alpha_d(1-s_d)} + \chi(1-\psi)\frac{\tau}{\Lambda} - \kappa\frac{\xi}{K}, && \text{(bank deposit pricing)} && (60)\\
r_{nb} &= r_S + \kappa_{nb} L_{nb} + \frac{1}{\alpha^\ell(1-s_{nb})}, && \text{(non-bank pricing)} && (61)
\end{aligned}$$

In this general version, equity E is treated as fixed, so $dE = 0$ in the comparative statics below, while reserves R adjusts endogenously through (55). Differentiating Λ produces

$$\begin{aligned}
d\Lambda &= (dD + dW - dL) - \chi(\psi dD + dW)\\
&= -\,dL + (1-\chi\psi)\,dD + (1-\chi)\,dW. && (62)
\end{aligned}$$

Loan expansion tightens liquidity through its effect on reserve holdings, while deposits relax liquidity to the extent that they are less runnable than wholesale funding ($\psi < 1$). Differentiating the capital identity (57) gives

$$dK = -\kappa(dD + dW). \tag{63}$$

Intuitively, expanding liabilities reduces capital slack in proportion to κ. With non-bank credit, bank loan demand depends on both loan rates, $L = L(r_\ell, r_{nb})$, and the non-bank loan rate r_{nb} is pinned down by (61). We therefore differentiate the two pricing conditions (59)–(61) jointly and use logit semi-elasticities to express dL as a function of $(dW, d\Lambda, dK)$. Differentiating (59) gives

$$dr_\ell = \frac{1}{\alpha^\ell}\frac{1}{(1-s_\ell)^2}\, ds_\ell - \frac{\tau}{\Lambda^2}\, d\Lambda. \tag{64}$$

Differentiating our logit demand functions we have

$$ds_\ell = -\alpha^\ell s_\ell(1-s_\ell)\, dr_\ell + \alpha^\ell s_\ell s_{nb}\, dr_{nb}. \tag{65}$$

Next we substitute (65) into (64) to yield

$$dr_\ell = \frac{s_\ell s_{nb}}{1-s_\ell}\, dr_{nb} - (1-s_\ell)\frac{\tau}{\Lambda^2}\, d\Lambda. \tag{66}$$

Differentiating (61) gives

$$dr_{nb} = \kappa_{nb}\, dL_{nb} + \frac{1}{\alpha^\ell}\frac{1}{(1-s_{nb})^2}\, ds_{nb}. \tag{67}$$

Using logit semi-elasticities,

$$dL_{nb} = -\alpha^\ell(1-s_{nb})L_{nb}\, dr_{nb} + \alpha^\ell s_\ell L_{nb}\, dr_\ell, \tag{68}$$

$$ds_{nb} = -\alpha^\ell s_{nb}(1-s_{nb})\, dr_{nb} + \alpha^\ell s_{nb}s_\ell\, dr_\ell, \tag{69}$$

and substituting (68)–(69) into (67) implies

$$\Big(1 + \kappa_{nb}\alpha^\ell(1-s_{nb})L_{nb} + \frac{s_{nb}}{1-s_{nb}}\Big)dr_{nb} = s_\ell\Big(\kappa_{nb}\alpha^\ell L_{nb} + \frac{s_{nb}}{(1-s_{nb})^2}\Big)dr_\ell. \tag{70}$$

Finally, bank loan demand satisfies

$$dL = -\alpha^\ell(1-s_\ell)L\, dr_\ell + \alpha^\ell s_{nb}L\, dr_{nb}. \tag{71}$$

Combining (66), (70), and (71) yields

$$dL = -\,\tilde{\kappa}_\ell^{nb}\Big(-\frac{\tau}{\Lambda^2}\,d\Lambda\Big), \tag{72}$$

where the non-bank-adjusted loan semi-elasticity $\tilde{\kappa}_\ell^{nb} > 0$ is given explicitly by

$$\begin{aligned}\tilde{\kappa}_\ell^{nb} &= \alpha^\ell L(1-s_\ell)^2 \\ &\quad\times \frac{(1-s_\ell)\Big(1+\kappa_{nb}\alpha^\ell(1-s_{nb})L_{nb}+\frac{s_{nb}}{1-s_{nb}}\Big) - \ s_\ell s_{nb}\Big(\kappa_{nb}\alpha^\ell L_{nb}+\frac{s_{nb}}{(1-s_{nb})^2}\Big)}{(1-s_\ell)\Big(1+\kappa_{nb}\alpha^\ell(1-s_{nb})L_{nb}+\frac{s_{nb}}{1-s_{nb}}\Big) - \ s_\ell^2 s_{nb}\Big(\kappa_{nb}\alpha^\ell L_{nb}+\frac{s_{nb}}{(1-s_{nb})^2}\Big)}.\end{aligned} \tag{73}$$

When $s_{nb} = 0$ (or the non-bank rate is fixed), the fraction in (73) equals one and (72) reduces to the banking-sector-only response. Note that by using this term the bank considers their own price and cross price elasticity. We now substitute the liquidity differential (62) and the wholesale-optimality condition into the linearized loan-demand relation (72) to solve for the pass-through $\lambda \equiv \dfrac{dL}{dD}$.

From (72) and (72)–(64), we have

$$dL = -\,\tilde{\kappa}_\ell^{nb}\Big(-\frac{\tau}{\Lambda^2}\,d\Lambda\Big) = \tilde{\kappa}_\ell^{nb}\,\frac{\tau}{\Lambda^2}\,d\Lambda, \tag{74}$$

so that

$$d\Lambda = \frac{\Lambda^2}{\tau}\,\frac{1}{\tilde{\kappa}_\ell^{nb}}\,dL. \tag{75}$$

Next, differentiating the wholesale FOC (58) implies

$$\kappa_W\,dW = -\frac{\tau(1-\chi)}{\Lambda^2}\,d\Lambda + \frac{\xi\kappa}{K^2}\,dK. \tag{76}$$

Using (63) to substitute $dK = -\kappa(dD + dW)$ into (76) yields

$$\Big(\kappa_W + \frac{\xi\kappa^2}{K^2}\Big)dW = -\frac{\tau(1-\chi)}{\Lambda^2}\,d\Lambda - \frac{\xi\kappa^2}{K^2}\,dD. \tag{77}$$

Substituting (75) and rearranging produces

$$dW = -\frac{(1-\chi)}{\kappa_W + \frac{\xi\kappa^2}{K^2}}\,\frac{1}{\tilde{\kappa}_\ell^{nb}}\,dL - \frac{1}{\kappa_W + \frac{\xi\kappa^2}{K^2}}\,\frac{\xi\kappa^2}{K^2}\,dD. \tag{78}$$

Finally, substitute (78) into the liquidity differential (62), and use (75) to eliminate $d\Lambda$:

$$\begin{aligned}\frac{\Lambda^2}{\tau}\frac{1}{\tilde{\kappa}_\ell^{nb}}dL &= -\,dL + (1-\chi\psi)\,dD + (1-\chi)\,dW \\ &= -\,dL + (1-\chi\psi)\,dD + (1-\chi)\Big(-\frac{(1-\chi)}{\kappa_W + \frac{\xi\kappa^2}{K^2}}\frac{1}{\tilde{\kappa}_\ell^{nb}}\,dL - \frac{1}{\kappa_W + \frac{\xi\kappa^2}{K^2}}\frac{\xi\kappa^2}{K^2}\,dD\Big).\end{aligned} \tag{79}$$

Collecting dL and dD terms and solving for $\lambda = \dfrac{dL}{dD}$ yields

$$\lambda = \frac{dL}{dD} = \frac{\frac{\tau}{\Lambda^2}\,\tilde{\kappa}_\ell^{nb}\left[(1-\chi\psi)\ -\ (1-\chi)\,\frac{\frac{\xi\kappa^2}{K^2}}{\kappa_W+\frac{\xi\kappa^2}{K^2}}\right]}{1+\frac{\tau}{\Lambda^2}\tilde{\kappa}_\ell^{nb}+\frac{\tau(1-\chi)^2}{\Lambda^2\left(\kappa_W+\frac{\xi\kappa^2}{K^2}\right)}}. \tag{80}$$

A.2.2.3 Derivation of $\mathcal{W}_D = \dfrac{dW}{dD}$

We can start with the differential form of $d\Lambda$

$$d\Lambda = (1-\chi\psi)\,dD + (1-\chi)\,dW - dL. \tag{81}$$

Next differentiate the wholesale FOC,

$$\kappa_W W = \frac{(1-\chi)\tau}{\Lambda} - \frac{\kappa\xi}{K}, \tag{82}$$

to get

$$\kappa_W\,dW = -(1-\chi)\tau\,\frac{1}{\Lambda^2}\,d\Lambda + \kappa\xi\,\frac{1}{K^2}\,dK. \tag{83}$$

Substituting into (83) yields

$$\left(\kappa_W + \frac{\xi\kappa^2}{K^2}\right)dW = -(1-\chi)\tau\,\frac{1}{\Lambda^2}\,d\Lambda - \frac{\xi\kappa^2}{K^2}\,dD. \tag{84}$$

Now impose the definition $\lambda \equiv dL/dD$ and rewrite (81) as

$$d\Lambda = \Big[(1-\chi\psi) - \lambda\Big]\, dD + (1-\chi)\, dW. \tag{85}$$

Substituting (85) into (84) gives

$$\begin{aligned} \left(\kappa_W + \frac{\xi\kappa^2}{K^2}\right) dW &= -(1-\chi)\tau\,\frac{1}{\Lambda^2}\Big(\big[(1-\chi\psi)-\lambda\big]\, dD + (1-\chi)\, dW\Big) - \frac{\xi\kappa^2}{K^2}\, dD \\ \Rightarrow \quad \left(\kappa_W + \frac{\xi\kappa^2}{K^2} + \frac{\tau(1-\chi)^2}{\Lambda^2}\right) dW &= -\left(\frac{\tau(1-\chi)}{\Lambda^2}\big[(1-\chi\psi)-\lambda\big] + \frac{\xi\kappa^2}{K^2}\right) dD. \end{aligned} \tag{86}$$

Dividing by dD yields the compact expression

$$\mathcal{W}_D \;\equiv\; \frac{dW}{dD} = -\frac{\frac{\tau(1-\chi)}{\Lambda^2}\big[(1-\chi\psi)-\lambda\big] + \frac{\xi\kappa^2}{K^2}}{\kappa_W + \frac{\xi\kappa^2}{K^2} + \frac{\tau(1-\chi)^2}{\Lambda^2}}. \tag{87}$$

A.2.3 How Deposit Inflows Propagate into Credit

To build intuition for how deposit inflows generate credit, consider an exogenous stable deposit inflow dD_0.[15] In equilibrium, total deposits dD and total loans dL are jointly determined by two conditions,

$$dL = \lambda dD \tag{88}$$

$$dD = dD_0 + \theta dL. \tag{89}$$

The first, Equation (88), is the bank's optimal lending response which states that the bank extends λ dollars of new loans for each dollar of deposits. The second, Equation (89), is the deposit recycling condition which states that total deposits equal the exogenous inflow plus the endogenous deposits created by lending: when the bank extends dL in loans, borrowers spend those funds, generating payment flows of which a fraction θ is retained within the banking system as stable deposits. Solving this system yields directly the deposit multiplier and the corresponding

[15] Such an inflow arises naturally from a shift in depositor preferences ξ_d that reallocates household savings from non-bank vehicles into bank deposits, or from a shift in borrower demand ξ_ℓ that stimulates lending and generates payment flows that settle as stable deposits. In either case, the inflow represents a change in the composition of bank liabilities rather than an injection of outside reserves.

increase in bank loans,

$$dD = \frac{dD_0}{1-\lambda\theta}, \quad \text{and} \quad dL = \lambda\frac{dD_0}{1-\lambda\theta}. \tag{90}$$

To see how this equilibrium quantity of deposits and loans arises, consider the round-by-round balance sheet dynamics. Since no new reserves accompany the deposit inflows, the initial effect is a pure liability recomposition: the stable deposit dD_0 replaces the wholesale funding one-for-one, with no change in balance sheet size. Because stable deposits are less runnable than wholesale, this recomposition relaxes the bank's liquidity constraint, creating room for balance sheet expansion.

The bank responds by originating λdD_0 in new loans. When the bank lends, it simultaneously creates a new liability — the borrower's account is credited. This new liability is initially flighty, because the borrower may intend to spend the funds immediately. The balance sheet has expanded by λdD_0 on both sides, with new loans on the asset side, and new wholesale funding on the liability side. The original stable deposit dD_0 remains in place, backing existing reserves. It was not used to fund the loan, rather it relaxed the liquidity constraint that made the bank willing to expand.[16]

As the borrowers spend the loan proceeds and the resulting payment flows circulate through the economy, a fraction θ of the initially flighty liabilities settle into stable core deposits. This adds $\theta\lambda dD_0$ in stable funding to the banking system without changing balance sheet size — it is a liability recomposition from flighty to stable funding. The remaining $(1-\theta)\lambda dD_0$ stay as flighty liabilities that do not support further expansion. The new stable deposits relax the liquidity constraint further, enabling the bank to originate another $\lambda^2\theta dD_0$ in loans, again creating flighty liabilities that partially stabilize through θ. The process continues with diminishing rounds, converging to the fixed point derived above.

[16] An equivalent accounting decomposition is to view the deposit inflow dD_0 as directly funding the loan: a fraction λ finances new lending while the remaining $1-\lambda$ retires wholesale. Both decompositions produce identical balance sheet outcomes: assets increase by λdD_0, deposits increase by dD_0, and wholesale falls by $(1-\lambda)dD_0$. However, the split decomposition obscures the role of θ, since the stable funding appears to leave when the borrower spends, with only $\theta\lambda dD_0$ returning. The self-funding interpretation avoids this confusion: the stable deposit remains in place, lending creates flighty liabilities, and θ unambiguously adds stable funding each round.

A.2.4 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.3 Details on Model Calibration

A.3.1 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. Equity is constructed residually as the difference between assets and liabilities.

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. 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 Baa 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 and statutory capital requirements. Capital slack and liquidity slack are computed as functions of equity, loans, core deposits, and wholesale funding, and therefore vary endogenously over time with the evolution of bank balance sheets.

A.3.2 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 loans 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

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

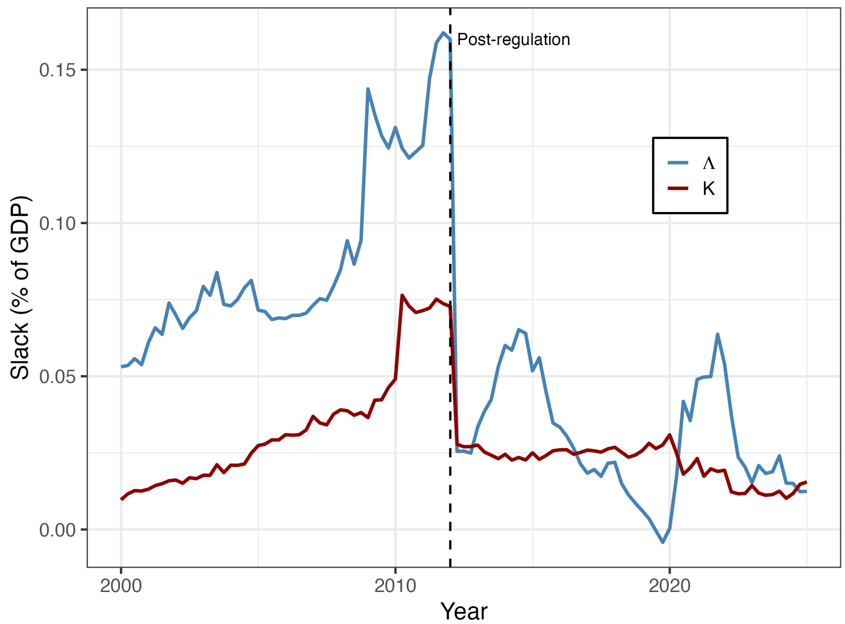
where $\eta_b > 1$ pins down the average share of non-borrowers. Borrower market shares are $s_t^L \equiv L_t/W_t^b$ and $s_t^{nb} \equiv L_{nb,t}/W_t^b$, with the no-borrowing share given by $s_t^0 \equiv 1 - s_t^L - s_t^{nb}$.

A.3.3 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 statutory capital requirement. Liquidity slack is defined as the difference between liquid assets and runnable liabilities scaled by regulatory liquidity parameters. Both slack variables are functions of balance sheet quantities and therefore vary over time with the evolution of deposits, loans, wholesale funding, and equity. Figure A.2 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.

Figure A.2: Model Implied Structural Parameters

The figure plots the implied slacks (K_t, Λ_t). The vertical line marks the start of the post-reform regime. Lower values indicate tighter constraints.



A.4 Additional Counterfactual Results

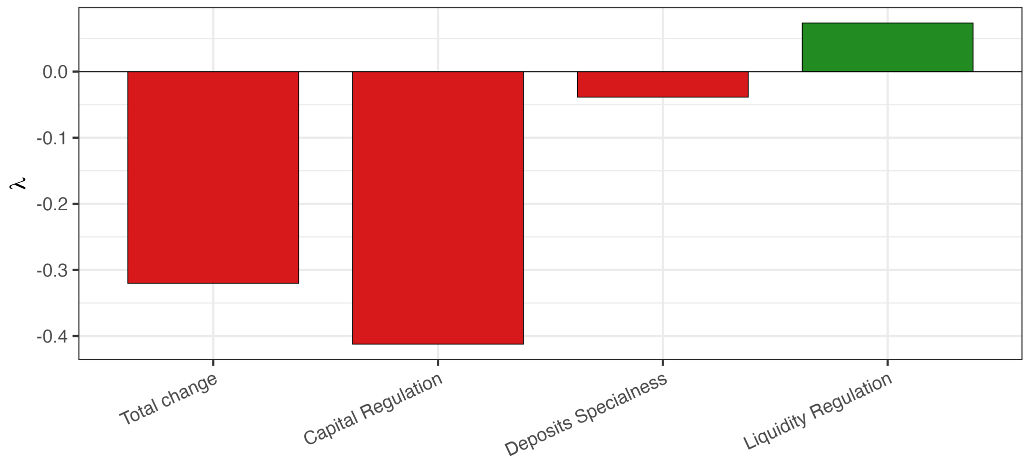
A.4.1 Decomposition

Panel (a) reports banks' marginal lending response λ. Banks' marginal lending response declines from roughly 0.65 in 2001Q4 to about 0.32 in 2021Q4, a reduction of approximately 0.33. Tightening capital regulation more than accounts for this decline: moving to the post-regulation capital wedge reduces λ by roughly 0.40 on its own. Changes in deposit runnability further lower λ by about 0.05, reflecting a weaker funding advantage of deposits in the post period. In contrast, the post-regulation liquidity regime offsets part of the contraction, increasing λ by roughly 0.08–0.10 relative to the capital-only counterfactual. Stricter liquidity requirements reduce effective liquidity slack, which makes stable deposits more valuable at the margin as a source of funding for credit provision and increases the sensitivity of lending to deposit inflows.

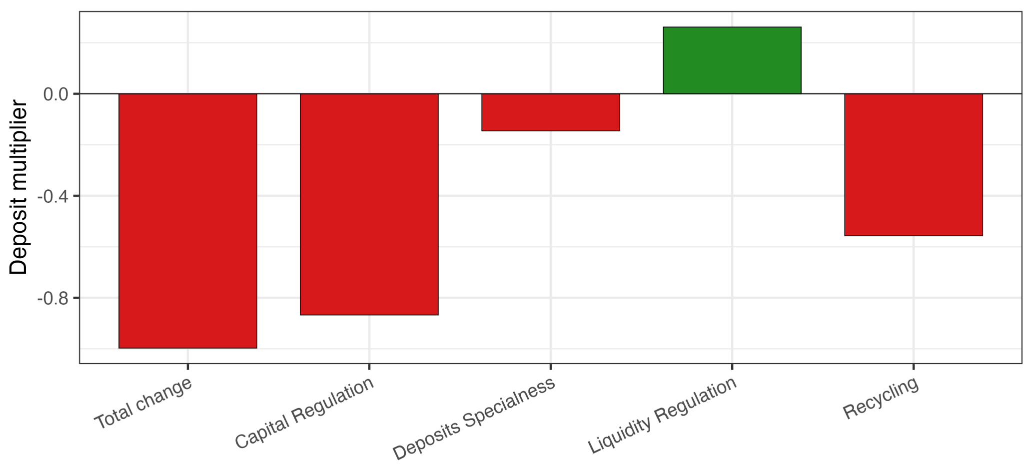
Panel (b) reports the deposit multiplier m. The multiplier falls from approximately 2.6 to 1.3, a decline of about 1.3. Tighter capital regulation can explain a significant fraction of the observed reduction, lowering m by roughly 0.8 in isolation. Reduced deposit runnability contributes an additional decline of about 0.2. The post-regulation liquidity regime again partially offsets these forces, raising the multiplier by roughly 0.2 relative to the capital and runnability adjustments. Crucially, the decline in the recycling parameter θ lowers the multiplier by approximately 0.6, accounting for nearly half of the total decline. This highlights that changes in banks' balance sheet optimization alone cannot explain the weakening of credit creation: the diminished tendency for loan proceeds to return as stable deposits is a quantitatively important source of the decline. As fewer loan-generated payment flows settle as deposits, the endogenous feedback loop that once amplified deposit inflows into large credit expansions has progressively unwound.

Figure A.3: Decomposing Decline in Credit Creation

Each panel reports a sequential decomposition of the change from the 2001Q4 baseline to the 2021Q4 baseline. The first bar shows the total change. The subsequent bars report the marginal contribution of sequentially substituting (i) capital regulation, (ii) deposit runnability (deposits specialness), and (iii) liquidity regulation. The multiplier in panel (b) also incorporates the influence of (iv) the recycling parameter, with each effect computed relative to the previous bar.



(a) Banks' Marginal Lending Response λ



(b) Credit-Deposit Multiplier m

A.4.2 Counterfactual Construction: QE and QT

A.4.2.1 Quantitative Easing

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

$$D_{QE,t} = \phi \cdot \max(R_t - \bar{R}_{pre}, 0), \tag{91}$$

where $\bar{R}_{pre}$ is the average reserve-to-GDP ratio over 2005–2007, before QE began, and ϕ is the share of reserve-created deposits that land in core deposits. Our baseline sets $\phi = 1$. We construct a counterfactual balance sheet by removing QE-created deposits and the associated reserves:[17]

$$\tilde{D}_t = D_t - D_{QE,t}, \qquad \tilde{R}_t = R_t - D_{QE,t}. \tag{92}$$

Because QE deposits are institutional and flighty, we assign them a runnability of $\psi_{QE} = 1$, equivalent to wholesale funding. Removing them lowers the overall runnability of the deposit base:

$$\tilde{\psi}_t = \frac{\psi_t D_t - \psi_{QE} \cdot D_{QE,t}}{\tilde{D}_t}. \tag{93}$$

We evaluate the counterfactual at the post-period average balance sheet. Given the smaller deposit base $\tilde{D}$ and adjusted runnability $\tilde{\psi}$, we re-solve for equilibrium lending, non-bank lending, and wholesale funding, holding all structural parameters at their estimated values.[18] We then evaluate the lending response $\tilde{\lambda}$ at the counterfactual balance sheet and compute the counterfactual multiplier $\tilde{m} = 1/(1 - \bar{\theta}\tilde{\lambda})$, where $\bar{\theta} = (\theta_{pre} + \theta_{post})/2$. This midpoint reflects that QE likely depressed θ by compressing deposit rate spreads and diluting the stable deposit base, so the no-QE counterfactual should use a higher recycling rate than θ_{post}.

[17] We remove QE deposits from the depositor market entirely rather than reallocating them to money market funds or other savings vehicles. This reflects the fact that QE deposits were not an equilibrium outcome of the depositor choice problem—they were a mechanical byproduct of the Fed's asset purchases. The sellers of securities to the Fed were institutional investors whose funds, absent QE, would have remained in securities outside the deposit-MMF-cash choice set.

[18] The borrower's outside option is calibrated to reproduce observed post-period market shares and held fixed across counterfactuals, so that differences reflect only the change in bank funding conditions.

A.4.2.2 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. In our baseline we assume that banks absorb the entire runoff. The parameter ω governs the liquidity treatment of the newly acquired securities: a fraction ω is treated as liquid, comparable to reserves, so the liquidity buffer tightens by only $(1-\omega) \times X_{QT}$ rather than the full amount. We set $\omega = 0.5$, reflecting the mixed liquidity characteristics of the securities banks absorb. The capital constraint tightens fully since κ applies to all non-reserve assets regardless. In the short-run counterfactual, wholesale funding adjusts to satisfy the first-order condition while the loan book is held fixed. In the long-run counterfactual, bank lending L, non-bank lending L_{nb}, and wholesale funding W all adjust to clear the equilibrium conditions.

A.4.3 Backing Out Non-Bank Intermediation Costs κ_{int}

In the model, κ_{nb} is a parsimonious parameter that captures not only intermediation costs but also macro trends. To isolate genuine non-bank intermediation costs, we exploit the relative spread $r_{nb} - r_\ell$. Common forces affecting both bank and non-bank pricing—credit risk conditions, general financial conditions—cancel in this difference. From the bank and non-bank first-order conditions:

$$r_{nb} - r_\ell = (r_{10yr} - r_S) + \left(\frac{1}{\alpha_\ell(1-s_{nb})} - \frac{1}{\alpha_\ell(1-s_\ell)}\right) + \kappa_{int} \cdot L_{nb} - \frac{\tau}{\Lambda}, \tag{94}$$

where τ/Λ is the bank regulatory wedge and $(r_{10yr} - r_S)$ is the term spread. Including τ/Λ_t strips out the bank regulatory tightening that compressed $r_{nb} - r_\ell$ from the bank side. Any common credit risk force that contaminates τ/Λ_t also moves r_{nb} proportionally—since borrowers come from the same pool—so these effects cancel.

We back out κ_{int} directly,

$$\kappa_{int} = \frac{(r_{nb} - r_\ell) - (r_{10yr} - r_S) - \left(\frac{1}{\alpha_\ell(1-s_{nb})} - \frac{1}{\alpha_\ell(1-s_\ell)}\right) + \frac{\tau}{\Lambda_t}}{L_{nb}}, \tag{95}$$

evaluated at period-specific averages. Every term on the right is either observed data or already estimated from the bank side.

We find that $\kappa_{int,pre} \approx \kappa_{int,post}$, indicating that non-bank-specific intermediation costs were approximately stable across periods. The entire apparent decline in κ_{nb} was driven by common macro forces and bank regulatory tightening, not by genuine improvement in non-bank supply capacity. This is consistent with the observation that non-bank credit relative to GDP did not expand between periods.

Intuitively, κ_{nb} is identified from how much non-bank spreads rise as non-bank lending volume expands: a steep relationship implies high marginal intermediation costs, while a flat relationship implies low costs. Bank regulatory tightening pushes borrowers from banks to non-banks, raising L_{nb}. If non-bank spreads do not rise proportionally—for instance, because the marginal borrower displaced from banks is similar in quality to existing non-bank borrowers—the estimated κ_{nb} mechanically falls, even though nothing has changed about non-bank intermediation technology itself. The κ_{int} decomposition isolates genuine improvements in non-bank capacity from this kind of regulatory spillover.