Reaching for Duration

Thomas M. Mertens   Pascal Paul   Andres Schneider

July 9, 2026*

Abstract

Using historical data on U.S. commercial bank balance sheets, we show that banks' maturity mismatch has more than tripled since the mid-1980s, moving in close lockstep with declining interest rates and term premia. We rationalize these trends in a model of bank portfolio choice in which banks must cover operating costs out of current earnings. When term premia or short-term rates decline, banks extend the duration of their assets to remain profitable. This "reaching for duration" effect is convex in the degree of term premium compression. The resulting maturity mismatch renders banks increasingly vulnerable to self-fulfilling runs by uninsured depositors. Consistent with the model, less profitable banks subsequently raise their asset maturities, particularly in periods of low term premia and large Federal Reserve asset holdings. Quantitative easing, designed to remove duration risk from the private sector, may thus paradoxically concentrate it on bank balance sheets and undermine financial stability.

Keywords: Maturity mismatch, term premium, quantitative easing, financial stability, bank runs, deposit franchise.

JEL Classification: E43, E52, E58, G01, G11, G21.

*Mertens: Federal Reserve Bank of San Francisco, thomas.mertens@sf.frb.org. Paul: Federal Reserve Bank of San Francisco, pascal.paul@sf.frb.org. Schneider: Miami Herbert Business School, andres.schneider@miami.edu. We thank Juliane Begenau, Itay Goldstein, Benjamin Hebert, Arvind Krishnamurthy, and seminar participants at the Federal Reserve Bank of San Francisco for comments and suggestions. We also thank Rohit Garimella and Wesley Wasserburger for excellent research assistance. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of San Francisco, the Board of Governors, or the Federal Reserve System.

1 Introduction

Banks' maturity mismatch—the fact that bank liabilities have a shorter duration than their assets—creates an inherent instability of the financial system. When a bank takes on redeemable deposits and invests in long-term assets, a sudden increase in interest rates lowers the market value of bank assets and can lead to a rapid withdrawal of deposits. Such risks have materialized in the past and resulted in failures of financial institutions, as during the U.S. savings and loan crisis of the 1980s and the more recent regional banking stress in 2023 around the failure of Silicon Valley Bank.

In this paper, we document new empirical evidence on the maturity structure of U.S. commercial bank balance sheets. Using historical Call Report vintages, we show that banks' maturity mismatch has widened significantly over the past decades. We measure this mismatch by the *maturity gap*: the difference between the value-weighted average maturity of a bank's assets and its liabilities. Over our sample, this gap has moved in close lockstep with the general level of interest rates and term premia. In the mid-1980s, following the Volcker disinflation, during which the federal funds rate had well exceeded 10 percent, and with term premia still near their historical highs, the maturity gap reached a low point of around 1.6 years. As both short- and long-term yields declined over subsequent decades and as term premia compressed alongside them, banks steadily extended the duration of their assets. This trend sharply accelerated around the Global Financial Crisis and continued through the prolonged period of near-zero policy rates and successive rounds of Federal Reserve quantitative easing (QE), which aimed to reduce term premia. During the most recent zero-lower-bound episode and the pandemic-era QE of 2020–2021, banks further increased the duration of both their loans and securities, pushing the maturity gap to a peak of 5.7 years in late 2022—an increase of roughly 4 years relative to its trough in the mid-1980s.

These empirical patterns raise a number of salient questions. First, why do banks tend to increase the duration of their assets precisely when interest rates decline and term premia compress—an environment in which the compensation for bearing duration risk is smallest? Second, do banks optimally engage in such behavior, knowing that it increases their exposure to a potential rise in interest rates and the associated risk of deposit outflows and bank runs? Third, why do we not observe a decline in bank maturity mismatch during periods of quantitative easing, as these policies credit banks with zero-maturity reserves and are intended to remove duration risk from the private sector?

To answer these questions, we develop a model of bank portfolio choice with profitability concerns. We consider a bank that is endowed with equity and deposits, which together fund a portfolio of short-term and long-term assets. The short-term asset pays the prevailing policy rate. The long-term asset is a bond that earns a return that decomposes into expected future short rates and a term premium—the compensation for bearing duration risk. The bank possesses deposit market power: it passes through only a fraction of changes in market rates to depositors, generating a valuable deposit franchise. After the bank chooses its portfolio allocation, the short rate is subject to a shock. Shareholders are risk-averse and seek to stabilize the bank's profits by choosing the share invested in the long-term bond.

Within this environment, we show that banks can perfectly insulate their profit margins from interest rate shocks by choosing a "hedging portfolio" allocation. This portfolio matches the duration of assets to the duration of non-repricing funding, that is, the share of the bank's funding whose costs do not adjust with market rates. This result generalizes the central finding of Drechsler, Savov, and Schnabl (2021) to a setting with a more general capital structure. However, banks may choose to deviate from the hedging portfolio and allocate a larger share toward the long-term asset to earn the term premium. The optimal portfolio trades off additional return against variance in profits, and the higher the term premium, the more the bank invests in the long-term bond. Thus, without additional frictions, this model cannot explain the empirical pattern of rising maturity gaps at times when rates and term premia are low.

We therefore take profitability concerns of banks into account. When choosing the optimal portfolio, the bank faces a profitability constraint that requires the bank's earnings to reach at least a certain level, after accounting for its operating expenses. In practice, banks face several operating costs—for branches, personnel, compliance, and technology. We document that such noninterest expenses are substantial in the data, accounting for around 2.5% of total assets for the aggregate U.S. banking sector. In our setting, the constraint is equivalent to a minimum required return on equity, a widespread target in the banking industry (Pennacchi and Santos, 2021). These targets are not merely conventions: because banks build regulatory capital primarily from retained earnings, persistently low income erodes capital and can trigger supervisory intervention, so that the constraint captures the real costs that low current earnings impose on a bank's viability as a going concern.

The profitability constraint gives rise to a *reach-for-duration* effect. For the constraint to remain slack, the bank has to generate sufficient interest income. However, when short-term interest rates or term premia are depressed, the bank's unconstrained portfolio allocation no longer generates enough income to satisfy the constraint. We characterize the critical term premium at which this occurs. Below this threshold, the bank is forced to increase its expected return on equity by investing more in the high-yielding long-term asset and thus extending the duration of its investments.

One prominent reason for a decline in the term premium is quantitative easing. When a central bank conducts QE, it purchases long-term securities from the banking system and credits banks' reserve accounts in return. Because reserves carry zero maturity, a natural first-order view of QE is that it removes duration risk from bank balance sheets. Under this view, banks' maturity mismatch—and hence their exposure to interest rate risk—should decline during periods of large-scale asset purchases, in contrast to the empirical patterns we document.

When banks generate sufficient interest income such that the profitability constraint is not binding, asset purchases by the central bank incentivize banks to lower their long-term bond holdings, as the compensation for duration risk is reduced. However, below the critical term premium, QE has the opposite effect: since banks need to reach for duration to achieve a sufficiently high expected return on their investments, QE *increases* banks' maturity mismatch. Thus, a policy that intends to remove duration risk from private portfolios ends up concentrating such risk within the banking sector. We further show that this effect is convex: the more the term premium is compressed, the larger the incremental increase in duration required to cover costs. The reach-for-duration effect is more likely to arise in low-interest-rate environments, where deposit spreads are diminished, leaving little room to absorb bank operating costs. Our model thus provides a unified explanation for the historical behavior of banks' maturity mismatch: its secular rise as yields and term premia have fallen, and the acceleration of this trend during QE episodes.

We consider two extensions of our baseline model. First, we allow banks to issue loans and explore how this additional margin affects their need to reach for duration. When facing profitability pressure, we show that banks may *reach for credit risk* if they can expand their lending without lowering credit spreads. In contrast, if lending markets are saturated and banks have to lower their credit spreads to attract additional borrowers, they instead adjust their bond portfolio and reach for bond duration. Second, we allow the term premium to respond endogenously to banks' collective demand for long-term assets. This extension is motivated by evidence in Haddad and Sraer (2020) that banks actively influence the pricing of interest rate risk. We show that an amplification mechanism emerges: when banks are under profitability pressure as the term premium declines, they increase their holdings of long-term securities, which in turn compresses the term premium further. This additional compression tightens the profitability constraint, requiring yet more duration exposure.

Next, we introduce bank runs into our baseline setting, building on recent work by Drechsler, Savov, Schnabl, and Wang (2026) and Jiang, Matvos, Piskorski, and Seru (2024). A central premise of these papers is that runs need not rely on the illiquidity of bank assets, as in Diamond and Dybvig (1983). Instead, a run equilibrium can arise even when assets are perfectly liquid—as are the bonds in our model—since the withdrawal of uninsured depositors implies the loss of a bank's valuable deposit franchise, which can render a bank insolvent. Importantly, a larger maturity mismatch implies a greater exposure to a rise in interest rates, which erodes a bank's value and, hence, increases the likelihood of a self-fulfilling run. Thus, when banks reach for duration due to profitability concerns, they also raise their exposure to a withdrawal of uninsured depositors. Banks optimally engage in such risk exposure since they need to generate sufficient income to satisfy the profitability constraint. We characterize the threshold of the term premium below which self-fulfilling runs become possible and show that a sufficiently high level of bank capital rules them out. Quantitative easing can therefore sow the seeds of financial instability: by compressing term premia, it pushes banks into portfolios that are increasingly vulnerable to an interest rate reversal that an eventual policy tightening would bring.

We further generalize our baseline model to a continuous-time dynamic setup, in which banks face persistent interest rate shocks, trade assets dynamically over time, and experience mark-to-market movements in risky asset positions. In this environment, we show that our main predictions from the baseline framework continue to hold and we obtain an additional key prediction: Banks reach for duration even if the profitability constraint does not bind currently, but there is a positive probability of it binding in the future. That is, the presence of concerns about future profitability due to lower rates or term premia already incentivizes banks to reach for duration. This result is driven by risk-averse shareholders' desire to optimally hedge against states in which the constraint is binding, which are episodes of low profitability and highly volatile wealth.

In addition, we use the model as a laboratory to assess the structural forces driving the secular rise in bank maturity mismatch quantitatively. We find that the observed declines in short-term rates and term premia over the past 40 years jointly explain the rise in maturity gaps, almost exactly, and each force contributes around half. Importantly, this shift is exclusively driven by banks' concerns about profitability as the constraint never binds along the transition. We also employ the model to exclude an alternative explanation: a secular rise in bank deposit market power. Such a structural change is unable to match the rise of bank maturity mismatch in the data, and even predicts a decrease when interest rates and term premia are low.

Finally, we test the model's central prediction in the cross-section of banks using bank-level Call Report data. Across a battery of specifications, we obtain three main findings. First, banks with lower profitability—measured by their return on assets—subsequently raise their maturity gap. Second, the response is concentrated on the asset side of the balance sheet, with the maturity of securities responding more strongly than the maturity of loans, and is associated with an increase in banks' net interest margins. And third, the maturity gap response is amplified in periods of low term premia and quantitative easing. Taken together, this bank-level evidence is consistent with the model's main mechanism: when banks face profitability pressure, they extend duration to earn additional interest income.

Related Literature. Our paper contributes to several strands of literature. An extensive body of work studies banks' exposure to interest rate risk, dating back to Flannery (1981, 1983). Recent contributions measure this exposure and its consequences for bank valuations and monetary policy transmission (e.g., English et al., 2018; Gomez et al., 2021; Begenau et al., 2025) or rationalize it as an optimal response to the bank's deposit franchise (Drechsler et al., 2021; Di Tella and Kurlat, 2021). Within this strand, a prominent line of research argues that accommodative monetary policy and persistently low interest rates compress bank profit margins and encourage risk-taking (e.g., Maddaloni and Peydró, 2011; Borio and Zhu, 2012; Dell'Ariccia et al., 2014; Jimenez et al., 2014; Dell'Ariccia et al., 2017; Martinez-Miera and Repullo, 2017; Claessens et al., 2018), with implications for the transmission of monetary policy in general equilibrium (e.g., Ulate, 2021; Abadi et al., 2023; Wang, 2025). Relative to this literature, we underscore the role of the term premium—alongside the level of interest rates—for bank profitability and duration risk-taking, consistent with evidence that banks' profit margins comove with term premia (Paul, 2023). This emphasis connects bank portfolio choice to quantitative easing: since QE is designed to compress the term premium, a policy that intends to remove duration risk from the private sector can, paradoxically, concentrate it in the banking sector.

Second, our paper relates to the literature on *reaching for yield*, which documents that investors tilt toward higher-yielding assets when interest rates decline—among insurers, money market funds, mutual funds, and households (e.g., Becker and Ivashina, 2015; Di Maggio and Kacperczyk, 2017; Choi and Kronlund, 2018; Lian et al., 2019; Daniel et al., 2021). Sharing our focus on duration, Domanski et al. (2017) show that falling long-term rates widen the duration gap of European life insurers, whose rebalancing puts further downward pressure on yields, an amplification mechanism similar to the endogenous term premium response in our framework, and Barth et al. (2024) document that mutual funds use Treasury futures to reach for duration relative to their benchmarks.

Two papers are closest to ours. First, Hanson and Stein (2015) propose a model in which yield-oriented investors, commercial banks in particular, shift toward longer-term bonds when short-term rates fall, compressing term premia and thereby transmitting monetary policy to long-term real rates. In our model, such yield-oriented behavior arises endogenously from a profitability constraint and it responds to the term premium itself rather than only to the short rate—generating the convexity and the implications for QE and financial stability at the center of our analysis. Second, Campbell and Sigalov (2022) show that reaching for yield can arise from rational portfolio choice: an investor who must fund a fixed level of spending out of current income optimally takes more risk as the safe rate declines. Our profitability constraint plays a similar role for banks, standing in for the regulatory and going-concern costs of low earnings, but the margin of adjustment is the duration of bank assets and the relevant compensation is the term premium. Reaching for duration is therefore best understood as a rational response to the costs of low earnings rather than a behavioral anomaly.

Our mechanism is distinct from risk-shifting under limited liability, whereby banks gamble for resurrection as their charter value erodes, since shareholders capture the upside while creditors bear the downside (e.g., Keeley, 1990; Hellmann et al., 2000). In our setting, the bank is risk-averse and takes on duration risk that it would strictly prefer to avoid, driven by an earnings floor rather than by the convexity of the equity payoff. The effect therefore operates even for well-capitalized banks far from default. Moreover, while this literature has largely focused on credit and default risk, we emphasize the duration of bank assets as the margin of adjustment and the term premium as the relevant compensation for risk.

Finally, our paper contributes to a recent literature on banks' vulnerability to runs, motivated by the March 2023 events in particular (Haddad et al., 2023; Jiang et al., 2024; DeMarzo et al., 2024; Drechsler et al., 2026), and to the literature relating monetary policy and financial stability more broadly, including the optimal design of central bank balance sheet policies (Diamond and Rajan, 2012; Stein, 2012; Ajello et al., 2022; Boyarchenko et al., 2022; Alexandrov and Brunnermeier, 2026). We show that QE policies, designed to depress the term premium, incentivize banks to reach for duration, exposing them to the risk of depositor runs. Our results thus highlight a trade-off between unconventional monetary policy and financial stability: while QE can stabilize markets in the short run, it may leave banks increasingly vulnerable to an interest rate reversal once monetary policy eventually tightens.

The remainder of the paper is organized as follows. Section 2 presents the empirical evidence on bank maturity mismatch. Section 3 develops the baseline model and derives the reach-for-duration mechanism. Section 4 introduces deposit franchise runs. Section 5 extends the baseline model to a continuous-time dynamic setup and introduces the concept of ex-ante reaching for duration. Section 6 presents cross-sectional empirical evidence testing our mechanism. Section 7 concludes.

2 Historical Evidence on Bank Maturity Mismatch

This section presents stylized facts about the maturity structure of U.S. commercial bank balance sheets. We show that, since the mid-1980s, the gap between the maturity of assets and liabilities has widened substantially. This development occurred while interest rates and term premia fell. To establish this, we construct long time series of the average maturity of bank assets and liabilities using data from the Federal Financial Institutions Examination Council (FFIEC) Consolidated Reports of Condition and Income ("Call Reports"). While the existing literature has relied on the maturity and repricing schedules available in modern Call Reports—typically from the late 1990s onward (see, e.g., Paul, 2023)—a contribution of this paper is to extend these series back in time using older reporting forms whose maturity data have not previously been utilized: to 1959 for the maturity of bank assets and to 1976 for the maturity of liabilities and, hence, the maturity gap.

2.1 A New Stylized Fact: Banks' Maturity Choices, 1976–2026

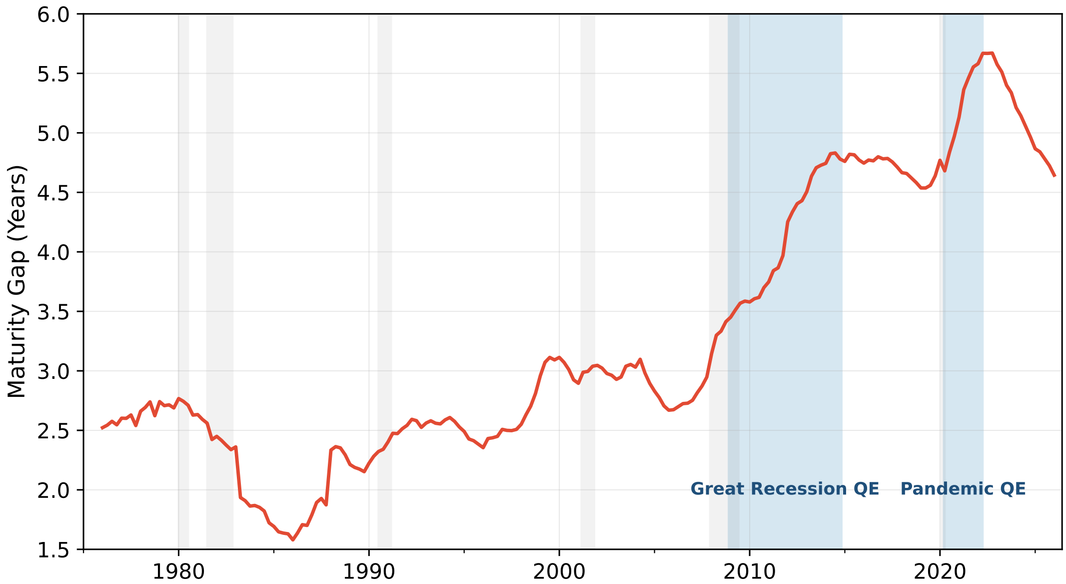
We first present our new stylized fact on the evolution of banks' maturity mismatch. The following subsections describe the underlying data and decompose the maturity gap into its balance sheet components to trace out the drivers of its secular rise.

The maturity gap. The maturity gap—defined as the difference between the average maturity of assets and liabilities—is a summary measure of the duration mismatch on bank balance sheets. For each bank, we compute the dollar-weighted average maturity of its assets and of its liabilities from the maturity and repricing schedules of the Call Reports, described in detail in Section 2.2. The maturity gap is the difference between the two, averaged across all banks. Figure 1 plots this series from 1976:Q1, the first quarter for which both components are available, through 2026:Q1, the end of our sample.

The most striking feature of Figure 1 is the secular increase in the maturity gap over the past five decades. In the early part of the sample, the gap hovered around 2.5–2.75 years. It then declined through the early 1980s, reaching a trough of about 1.6 years in 1986. This trough followed the Volcker disinflation, during which the federal funds rate had well exceeded 10 percent and long-term yields were similarly elevated, and it coincided with term premia close to their historical peak, as shown in Figure 2. Since then, the maturity gap rose steadily as interest rates gradually declined, a trend that accelerated sharply over the last two decades. Since the beginning of the Global Financial Crisis, the maturity gap has increased by almost 3 years, up to a peak of approximately 5.7 years in late 2022. As highlighted in Figure 1, these increases coincided with the large-scale Federal Reserve asset purchases related to the Great Recession and the Pandemic in 2020/21 when short-term policy rates were near zero.

Changes in financing conditions. The comovement between the maturity gap and the general level of interest rates may suggest that banks systematically extend duration as rates decline. The general trend of interest rates also coincides with the behavior of term premia: estimates from Adrian, Crump, and Moench (2013), plotted

Figure 1: The Maturity Gap



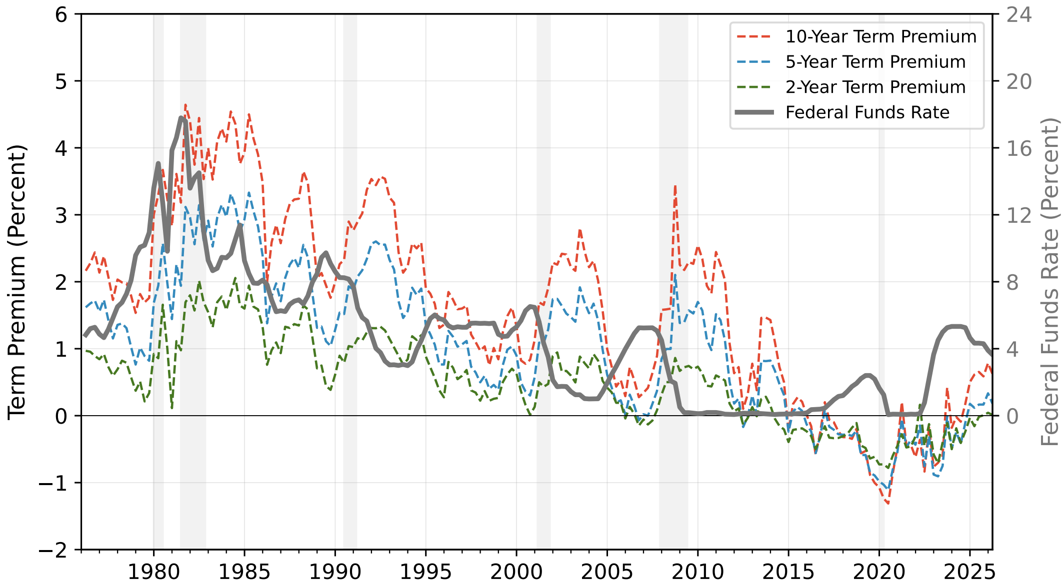
Notes: This figure plots the maturity gap, defined as the difference between the average maturity of total assets and total liabilities (in years), averaged across all banks. The series begins in 1976:Q1, the first quarter for which both asset and liability maturity data are available. Blue shaded areas denote the Federal Reserve's large-scale asset purchase programs: QE1 through QE3 (November 2008 to October 2014) and the pandemic-era purchases (March 2020 to March 2022). Gray shaded areas denote NBER recession dates.

in Figure 2, show that term premia at various maturities peaked in the early 1980s and have since declined to historically low levels.[1] A lower term premium reduces the compensation banks receive for bearing duration risk and, paradoxically, may encourage further reaching for duration as banks seek to maintain net interest margins. Taken together, the evidence points to an environment in which both the low level of interest rates and compressed term premia have contributed to a historically large maturity mismatch in the U.S. banking sector. Motivated by these historical patterns, we develop a model that formalizes such mechanisms in the following sections.

Further evidence. The historical evolution of the maturity gap is robust to several related considerations. First, we convert the maturity measures into duration estimates, which are closely related to the price sensitivity of an interest-bearing asset to yield changes. This adjustment shifts our maturity gap series down (see Appendix Figure 10). However, the downward shift is larger at high interest rates, such that the

[1] Term premia are challenging to estimate and commonly used term structure models often provide different estimates. For example, Bauer and Rudebusch (2020) show that a model with shifting endpoints significantly affects long-run trends in term premia. However, the overall decline based on such a model remains sizable since the mid-1980s (see Figure 7b therein). The estimates by Adrian, Crump, and Moench (2013) show a spike in term premia at the onset of the Great Recession around the first QE programs, which is absent in the model by Kim and Wright (2005) that also incorporates survey data.

Figure 2: Term Premia and Federal Funds Rate



Notes: This figure plots the effective Federal Funds rate and term premia estimates for the 2-, 5-, and 10-year maturities from Adrian, Crump, and Moench (2013). Shaded areas denote NBER recession dates.

overall percentage increase in the duration gap from trough to peak is even larger than in the baseline series, at around 340 percent against 260 percent. Second, we further adjust the duration gap for the prepayment risk inherent in mortgages, for which we observe separate maturity information since 1997:Q2. While this adjustment shifts the series down further when interest rates are low relative to their recent historical path, we still find an overall increase of around 260 percent from trough to peak, comparable to the percentage increase in our baseline maturity gap. Third, a breakdown of the maturity gap into three separate groups of banks ranked by total assets reveals that the secular rise in the maturity gap is prevalent across banks of all sizes, with smaller and mid-size banks showing the strongest increases since the start of the Great Recession (see Appendix Figure 11).

Before discussing the underlying data in more detail, we turn to the natural concern that the savings and loan crisis of the 1980s was an episode of severe maturity mismatch, which seems absent from our data and may thus appear at odds with the patterns we document. Two features of the regulatory landscape reconcile this apparent tension. First, the institutions at the center of that crisis are absent from our sample around this episode: federally chartered savings institutions filed Thrift Financial Reports rather than Call Reports until 2012 and were supervised by separate federal regulators. Second, thrifts and commercial banks operated under different regulatory regimes. Thrifts were by charter specialized housing finance institutions, with tax incentives and asset restrictions that effectively required a high share of assets in residential mortgages, and the long-term fixed-rate mortgage was the standard product, with adjustable-rate mortgages becoming broadly available only in the early 1980s (e.g., Kane, 1989; White, 1991). On the liability side, thrifts had limited access to long-term funding and were predominantly financed by short-term retail deposits. Thrifts therefore had little discretion over the duration of their assets and liabilities. In contrast, commercial bank charters allowed broad asset diversification and funding sources. Put differently, commercial banks could have built sizable portfolios of long-term fixed-rate mortgages during the early 1980s but they chose not to. The low maturity mismatch we document for the Call Report sample therefore reflects portfolio choice under broad discretion, rather than the charter restrictions that bound thrifts.

2.2 Data

Our data are drawn from the Call Reports filed by all commercial banks operating in the United States and regulated by the Federal Reserve System (FRS), the FDIC, and the OCC.[2] These data, digitally available at the Federal Reserve, provide quarterly—and, between 1964 and 1972, semiannual—information on bank balance sheets. The reporting basis varies by form: banks with only domestic offices report domestic-office items, while banks with foreign offices file consolidated reports; the appendix tables list the forms and variable codes used in each regime. The reporting forms have undergone several major revisions since the late 1950s, with substantial changes in the maturity and repricing categories reported. We identify seven distinct reporting regimes and map the relevant variables across each transition. Appendix Table A.1 provides a summary and Appendix Tables A.2–A.10 give complete descriptions of all variables used, including their variable codes, the corresponding maturity or repricing buckets, and the time periods over which each variable is available.

Asset maturity data. On the asset side, we collect data on the maturity or repricing distribution of securities (Treasury notes and bonds, agency obligations, municipal bonds, and other debt securities), loans and leases (broken down by type where available), and other interest-bearing assets. Cash assets, federal funds sold, and securities purchased under agreement to resell are assigned zero maturity. For each maturity bucket, we assign a representative maturity weight equal to the midpoint of

[2] In 2012, savings and loan associations started filing Call Reports. While we leave those institutions in our sample, we test and confirm that their exclusion does not affect our main results. Similarly, our data include a smaller set of FDIC-insured mutual savings banks in the 1980s, though their presence does also not affect our findings. For brevity, we refer to all institutions in our sample as commercial banks throughout.

the bucket, if available (see Tables A.2–A.6 for all weights used).[3]

The data coverage on the asset side begins in 1959:Q2, although the earliest periods provide only limited detail on securities maturities, and no maturity detail was collected between 1965:Q3 and 1974:Q1. From 1974:Q2 onward, the reporting becomes substantially richer, with separate maturity breakdowns for different security types and loan categories for larger banks, which begins in 1976:Q1. To make use of the additional loan information, we restrict the sample in this regime to the large banks and the constructed asset series therefore resume in 1976:Q1. A revised schedule, introduced in 1983:Q2 and carried over to the redesigned reporting forms in 1984, provides a unified maturity/repricing classification covering loans, securities, and other interest-bearing assets. From 1988 onward, fixed-rate and floating-rate positions are reported separately. The post-1997 schedule adds finer maturity bins (up to six buckets per asset class) and separates mortgage pass-throughs, other MBS, closed-end loans, and other loans.

Liability maturity data. On the liability side, we collect maturity and repricing information on time deposits (broken down by size threshold), non-deposit interest-bearing liabilities, subordinated debt, and limited-life preferred stock where available. Demand deposits, transaction deposits, and nontransaction savings deposits are assigned a repricing frequency of zero. Thus, we do not account for the possible stickiness of such deposits but rather associate the maturity with the redeemable nature of these deposits, since our theoretical analysis focuses on the possibility of bank runs.

Liability maturity data begin later than asset data: large banks first reported time deposit maturities in 1976:Q1.The liability reporting also evolves over time. From 1983:Q2 to 1987:Q4, time deposits are split into those above and below $100,000, with a fine repricing grid that also covers nondeposit interest-bearing liabilities, subordinated notes and debentures, and limited-life preferred stock. From 1988 onward, the split between fixed-rate and floating-rate time deposits becomes available. After an interim revision in 1996:Q1, a new set of maturity bins was introduced in 1997:Q2. In 2017, the size threshold for large time deposits shifted from $100,000 to $250,000, reflecting the change in the FDIC insurance limit. Appendix Tables A.7–A.10 document all liability variables across reporting regimes.

[3]For open-ended maturity/repricing buckets (e.g., >5 years), we either assume a specific maturity or match the weighted maturity of subsequent years for the same asset types and maturity spectrum at a point where there is a transition between reporting periods, which reduces possible biases for earlier periods.

2.3 Decomposing the Maturity Gap

Using the data described above, we construct quarterly time series for the average maturity of (i) total assets, (ii) total liabilities, (iii) securities, and (iv) loans. To obtain each series, we first compute the dollar-weighted maturity for each of these balance sheet categories by bank and then take the unweighted average across all banks at each quarter. Value-weighted averages, which aggregate dollar amounts across banks, display a similar secular increase.

Asset and liability maturities. We begin with the two components of the maturity gap: the average maturities of total assets and total liabilities. Total asset maturity combines securities and loans—the two main interest-bearing asset classes, which we examine separately below—with cash, federal funds, and other interest-bearing assets, all aggregated using the same methodology.

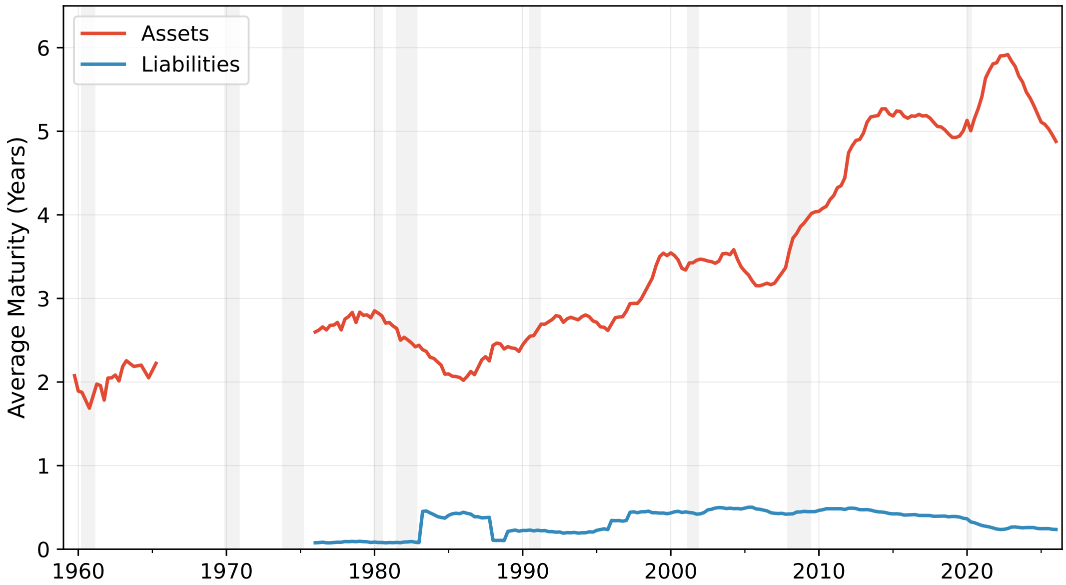
Figure 3 shows that the average maturity of assets begins at around 2 years in the early 1960s and rises gradually over time, reaching about 3.4 years at the onset of the Great Recession. In the years that followed, asset maturity increased substantially, up to a peak of about 5.9 years in late 2022. Liability maturity is an order of magnitude smaller and shows only minor fluctuations over the sample. The series peaks at about 6 months in the mid-2000s and has since trended down to about 3 months in recent years, reflecting the growing dominance of demand and savings deposits in bank funding.

Securities and loan maturities. We next examine the maturity of securities and loans separately. Figure 4 plots the two time series. Securities maturity is available from 1959:Q4 onward; loan maturity begins in 1976:Q1.

Several patterns stand out. In the early 1960s, the average maturity of bank securities portfolios was between 3 and 4 years. Maturity rose through the 1970s, reaching roughly 5.4 years around 1980, before declining to about 4 years by the late 1980s. From the mid-1990s onward, securities maturity increased markedly, with large increases around the Great Recession and the COVID-19 Recession. The series peaked at about 8.8 years in 2021 and stands at roughly 7.8 years at the end of our sample. Loan maturities display a broadly similar upward trend, but start from a lower level. In the late 1970s, the average loan maturity was approximately 2.6 years, reflecting the prevalence of short-term commercial lending. The series remained relatively flat through the early 1990s at around 2.3–2.5 years, then began a sustained increase that brought it to roughly 5 years by 2020. Loan maturities peaked at about 6.1 years in early 2022 and declined thereafter to about 4.8 years at the end of our sample.

Taken together, the decomposition shows that the secular widening of the matu-

Figure 3: Maturity of Assets and Liabilities



Notes: This figure plots the average maturity (in years) of total bank assets and liabilities. Asset maturity is available from 1959:Q4; liability maturity from 1976:Q1. No maturity detail is available between 1965:Q3 and 1975:Q4, and missing quarters during the semiannual reporting of the early 1960s are linearly interpolated. Shaded areas denote NBER recession dates.

rity gap is an asset-side phenomenon. Because liability maturity never exceeds six months, the gap moves almost one-for-one with asset maturity: of the roughly 4-year increase from the 1986 trough to the 2022 peak, nearly all is accounted for by longer asset maturities, with securities extending the most and loan maturities following from the mid-1990s onward. In the following sections, we build a model that links these trends to the concurrent decline in interest rates and term premia.

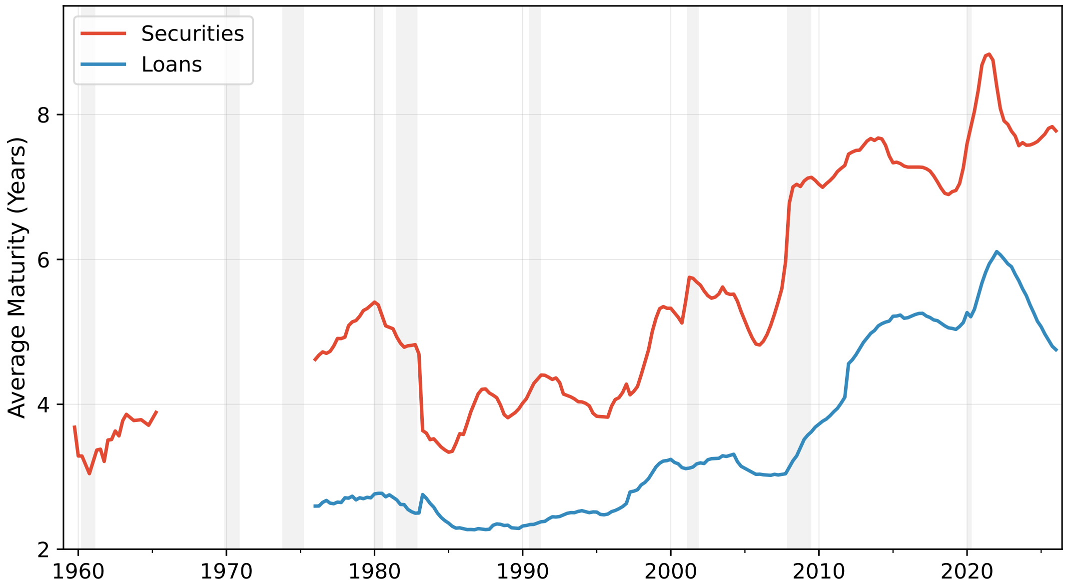
3 Baseline Model

This section develops a model of bank portfolio choice with the following key features: deposit market power, a term premium on long-term assets, stochastic short-term interest rates, bank operating costs, and risk-averse shareholders. We introduce each in turn, derive the banks' optimal portfolios, and show how a decline in term premia—which could be caused, among other forces, by QE—pushes banks into a region in which they have incentives to reach for duration.

3.1 Environment

Consider an economy with three dates ($t = 0, 1, 2$), two periods between them, and a continuum of identical banks. We describe the problem for a representative bank. It is

Figure 4: Maturity of Securities and Loans



Notes: This figure plots the average maturity (in years) of bank securities and loan portfolios. Securities maturity is available from 1959:Q4; loan maturity from 1976:Q1. No maturity detail is available between 1965:Q3 and 1975:Q4, and missing quarters during the semiannual reporting of the early 1960s are linearly interpolated. Shaded areas denote NBER recession dates.

funded by a fixed amount of deposits $(1-e)$ and equity $e \geq 0$ at $t=0$; total assets are normalized to one. As demonstrated in Figure 3, the average maturity gap is mainly driven by changes in the average maturity on the asset side. We therefore model the bank's portfolio choice through the possibility to invest in a short-term, one-period bond and a long-term, two-period bond. The assumption of a single long-term bond is motivated by the average maturities of securities and loan portfolios being highly correlated, as shown in Figure 4. The bank chooses the long-term bond share $\alpha \in [0,1]$ at $t=0$, yielding the balance sheet identity:

$$\underbrace{1-\alpha}_{\text{short bond}} + \underbrace{\alpha}_{\text{long bond}} = \underbrace{(1-e)}_{\text{deposits}} + \underbrace{e}_{\text{equity}}. \tag{1}$$

The short-term bond yields $r_{s,0}$ in the first period and can be rolled over at $t=1$ after an interest rate shock ε realizes. The shock shifts the short rate from $r_{s,0}$ to $r_{s,1} = r_{s,0} + \varepsilon$, and has mean zero ($\mathbb{E}[\varepsilon] = 0$) and variance σ^2. Thus, the shock introduces reinvestment risk for the bank since $r_{s,1}$ is uncertain from the viewpoint of $t=0$.[4] In contrast, the long-term bond pays r_ℓ per period regardless of ε and matures at $t =$

[4]In our baseline model, the bank faces reinvestment risk, but not mark-to-market risk. This changes in the following two sections: first, in Section 4, we introduce bank runs which are determined by the market value of bank assets at $t=1$, and second, in Section 5, we allow banks to trade the long-term asset continuously while facing mark-to-market risk.

2. The rate r_ℓ is equal to the average expected yield on the short-term asset plus an exogenous term premium $\tau > 0$, that is, $r_\ell = \frac{1}{2}(r_{s,0} + \mathbb{E}[r_{s,1}]) + \tau = r_{s,0} + \tau$.[5] The bank takes this premium as given: it is a price-taker in the bond market, where τ reflects the compensation demanded by the marginal investor for bearing duration risk (e.g., Vayanos and Vila, 2021). We treat τ as the main comparative-static parameter of our analysis; the extension in Section 3.7 allows the premium to respond to banks' own demand for duration.

Deposit market power. The bank has pricing power in the market for deposits. It pays a below-market deposit rate $r_{d,0} = \beta r_{s,0}$ in the first period and $r_{d,1} = \beta r_{s,1}$ in the second period, where $\beta \in [0, 1]$ is the *deposit beta*. This parameter measures how much of a change in market rates is passed through to depositors: $\beta = 1$ means full pass-through (no market power), while $\beta = 0$ implies the bank pays no interest on deposits regardless of market rates. We define the non-repricing share of funding

$$\lambda \equiv 1 - \beta(1 - e),$$

which captures the fraction of the bank's total funding that is shielded from interest rate risk. The bank pays floating rates on the remaining share $\beta(1 - e)$. Thus, it faces uncertainty about its interest expenses in the second period on a fraction $(1 - \lambda)$ of its funding.

Operating costs. The bank pays a fixed operating cost $c > 0$ per period. Among other expenses, these costs reflect payments related to running a deposit franchise, which are substantial in practice as documented below. While deposits are generally a cheap source of funding, at low short-term rates, the costs of running a deposit franchise may exceed its immediate benefits, $c > \lambda r_{s,0}$.

3.2 Net Income and Interest Rate Risk Exposure

The bank earns income on its assets, pays interest on deposits, and faces per period operating costs. In what follows, we term the combined profit margin the bank's net income (NI). At $t = 1$, the bank's NI is given by:

$$\text{NI}_1 = \underbrace{(1-\alpha)r_{s,0} + \alpha r_\ell}_{\text{asset income}} - \underbrace{(1-e)\,\beta r_{s,0}}_{\text{deposit cost}} - \underbrace{c}_{\text{operating costs}} = \lambda r_{s,0} + \alpha\tau - c,$$

[5] We take α to be both the unit allocation to the long bond and its $t = 0$ dollar value, since the bond is priced at par with yield-to-maturity r_ℓ.

where $\lambda r_{s,0}$ reflects the base-rate income $r_{s,0}$ on all assets minus the cost of deposits $(1-e)\beta r_{s,0}$. At $t=1$, after the shock ε is realized, the bank reinvests its cash at $r_{s,1}$. Net income in the second period is defined analogously and equals

$$\mathrm{NI}_2 = (1-\alpha)r_{s,1} + \alpha r_\ell - (1-e)\,\beta r_{s,1} - c = \mathrm{NI}_1 + (\lambda-\alpha)\varepsilon,$$

so that expected second-period net income equals first-period net income, $E[\mathrm{NI}_2] = \mathrm{NI}_1$, since $E[\varepsilon]=0$. After payoffs are realized in the second period, the bank pays out its accumulated income to shareholders. This terminal payout cumulates net income over both periods, net of initial equity and principal payments at $t=2$. In the small-rate approximation ($r_{s,0}, \tau, \varepsilon \ll 1$), which ignores this interest on interest, the terminal payout simplifies to:

$$\Pi_2(\varepsilon,\alpha) = \mathrm{NI}_1 + \mathrm{NI}_2 = 2\mathrm{NI}_1 + (\lambda-\alpha)\varepsilon. \tag{2}$$

The terminal payout is linear in the rate shock ε, with an intercept $2\mathrm{NI}_1$ (expected cumulative net income) and slope $\lambda-\alpha$ (the bank's interest rate exposure). When short-term rates rise, $\varepsilon>0$, two forces operate. First, the bank earns more on reinvesting the short-term asset $(1-\alpha)$. Second, the bank has to pay higher interest on its repricing liabilities $(1-\lambda)$. As a result, the bank's combined exposure to interest rate shocks is given by $(\lambda-\alpha)$.

The hedging portfolio. We define the bank's hedging portfolio to be the share α^{hedge} that implies that the terminal payout is deterministic and does not vary with any realization of ε, that is, $\mathrm{Var}(\Pi_2)=0$ and $\Pi_2(\alpha^{\text{hedge}}) = 2\mathrm{NI}_1$. Setting the exposure in (2) to zero gives the hedging portfolio:

$$\alpha^{\text{hedge}} = \lambda = 1-\beta(1-e). \tag{3}$$

Equation (3) demands that the bank should match the duration of its assets to the duration of its non-repricing funding to completely isolate its income from interest rate risk. Alternatively, we can interpret equation (3) in terms of the share invested in the short asset $1-\alpha^{\text{hedge}} = (1-e)\beta$. Intuitively, the bank eliminates variance in its payout by investing an amount $(1-\alpha^{\text{hedge}})$ in the short asset which equals the bank's funding that reprices with short-term rates. Simply put, the gains and losses from interest rate shocks in terms of bank cash-flows are equalized, resulting in an income gap that is equal to zero.[6]

[6] The income gap is defined as the difference between assets and liabilities that reprices over the near-term, normalized by total assets, and has been used empirically for bank holding companies based on FR Y-9C data (see, e.g., Haddad and Sraer, 2020).

Our definition of the hedging portfolio generalizes the central result of Drechsler, Savov, and Schnabl (2021) to a more generic liability structure which takes into account bank equity. When $e = 0$, the hedge reduces to $1 - \beta$, the bank's deposit franchise. With $e > 0$, the hedge increases by $e\beta$, because equity-funded assets create additional non-repricing income that must be hedged with duration.

3.3 The Bank's Problem

Shareholders have mean-variance preferences over the terminal payout Π_2 at $t = 2$:

$$U(\alpha) = \frac{1}{2}\mathbb{E}[\Pi_2] - \frac{\rho}{2}\,\text{Var}[\Pi_2], \tag{4}$$

where $\rho > 0$ represents shareholders' coefficient of risk aversion. The factor $1/2$ converts the two-period payout into a per-period equivalent. There are no frictions between bank management and shareholders.

Substituting the payout equation (2) into the objective function, we obtain the mean-variance problem

$$\max_{\alpha} U(\alpha) = \lambda r_{s,0} + \alpha\tau - c - \frac{\gamma}{2}(\lambda - \alpha)^2, \tag{5}$$

where $\gamma \equiv \rho\sigma^2$ is a joint risk-cost parameter that collects shareholders' risk aversion and the variance of rate shocks. The first term in (5) reflects the common per-period net income since $E[\text{NI}_2] = \text{NI}_1$. The second term reflects the cost of risk, whose minimum value is at the hedge portfolio, $\alpha^{\text{hedge}} = \lambda$.

Unconstrained portfolio choice. The first-order condition to the bank's unconstrained problem (5) gives the interior optimum

$$\alpha^*_{\text{int}} = \lambda + \frac{\tau}{\gamma} = \alpha^{\text{hedge}} + \frac{\tau}{\gamma}. \tag{6}$$

With interest rate risk eliminated at $\alpha^{\text{hedge}} = \lambda$, the bank chooses at least that much duration. If τ is strictly positive, the bank chooses to take on interest rate risk to earn the term premium, scaled by the risk-cost parameter γ.

Profitability constraint. We assume the bank faces a lower bound on its profitability when choosing its portfolio at $t = 0$. The constraint demands that the bank's (expected) net income at dates $t = 1$ and $t = 2$ must reach a certain level κ; that is, the constraint applies to the flow of net income in each period, not to the cumulative payout over the bank's horizon. Since $E[\mathrm{NI}_2] = \mathrm{NI}_1$, this simplifies to a single constraint

$$\mathrm{NI}_1 = \lambda r_{s,0} + \alpha\tau - c \geq \kappa, \tag{7}$$

where κ can be positive, zero, or even negative.[7] We interpret the floor as a reduced form for the costs that low earnings impose on a bank and that are absent from the frictionless mean-variance objective. The most direct of these costs operates through capital regulation and the bank's viability as a going concern. Banks must maintain regulatory capital ratios, and they build capital primarily by retaining earnings, since external equity is costly to issue (Myers and Majluf, 1984). A bank whose earnings are persistently low may therefore be subject to regulatory action.[8] For example, assume the bank's initial capital ratio at $t = 0$ equals its required ratio and that it can neither raise new equity nor shrink its balance sheet as it may be unable to shed deposits (Bolton et al., 2025). A profitability constraint with $\kappa = 0$ then ensures that the bank meets its capital requirement in expectation.

Several other forces point in the same direction. In our setting with fixed equity and total assets, the constraint is equivalent to a minimum required return on equity (ROE) or return on assets (ROA). Credit ratings, analyst expectations, and access to wholesale funding deteriorate when earnings fall, banks are documented to directly target ROE and tie executive compensation to it (Pennacchi and Santos, 2021), and they strongly smooth dividends over time (Floyd et al., 2015). Consistent with these pressures, banks' ROA is remarkably stable at around 1 percent over decades, as shown in Appendix Figure 12. The same emphasis on current earnings is pervasive beyond banking, where a large corporate finance literature documents that managers target earnings benchmarks.[9]

Importantly, the profitability constraint cannot be relaxed through financial transactions at $t = 1$. Not reinvesting the short asset does not generate earnings at $t = 1$. Similarly, borrowing leaves current net income unchanged and lowers future earnings through interest expense, thereby tightening the constraint at $t = 2$. Asset sales do not help either: While realizing gains could raise $t = 1$ net income, they also tighten the

[7]The hard constraint is the limiting case of a steeper cost of missing the earnings floor. Under a smooth penalty the bank begins tilting toward duration before a floor is reached; the dynamic model in Section 5 delivers this property, with banks reaching for duration in anticipation of a constraint that binds only in future states.

[8]U.S. regulators classify banks into capital categories and must impose escalating restrictions—on asset growth, dividends, and expansion—as capital declines, culminating in mandatory resolution of critically undercapitalized institutions. Moreover, stress tests and supervisory ratings like the CAMELS rating system depend in part on whether bank earnings are adequate, sustainable, and sufficient to absorb losses and support growth (the "E" in CAMELS stands for earnings).

[9]See, for example, Degeorge et al. (1999), Graham et al. (2005), Almeida et al. (2016), and Ben-David and Chinco (2023). In particular, the distribution of reported earnings displays a sharp discontinuity at zero, that is, firms try to avoid reporting losses (Burgstahler and Dichev, 1997), consistent with an earnings floor at $\kappa = 0$.

profitability constraint at $t = 2$. Put differently, financial transactions shift resources across periods. This property distinguishes an earnings target from a financing or cash-flow friction, which is about a stock of resources and can be met by moving resources across time. In contrast, an earnings constraint is about a recurring flow, applies period by period, and is therefore structurally immune to intertemporal shifting. Intuitively, the bank must earn κ, not have κ.

Constrained portfolio choice. The profitability constraint implies a minimum level of long-term bond holdings for a given term premium and short rate:

$$\alpha \geq \frac{\kappa + c - \lambda r_{s,0}}{\tau} \equiv \alpha^*_{\text{bind}}. \tag{8}$$

Throughout, we focus on the case $\kappa + c - \lambda r_{s,0} > 0$, in which non-repricing income alone does not cover operating costs and the earnings floor. In the opposite case, the constraint is slack for every portfolio and the model reduces to the unconstrained benchmark. The minimum required duration α^*_{bind} is a decreasing, convex function of the term premium: as τ falls, the bank needs progressively more duration exposure just to cover its costs. The profitability constraint binds whenever the bank's optimal interior portfolio takes on insufficient duration $\alpha^*_{\text{int}} < \alpha^*_{\text{bind}}$. This condition occurs when the term premium is sufficiently low:

$$\lambda + \frac{\tau}{\gamma} \leq \frac{\kappa + c - \lambda r_{s,0}}{\tau} \quad \Longleftrightarrow \quad \tau^2 + \gamma\lambda\tau \leq \gamma(\kappa + c - \lambda r_{s,0}). \tag{9}$$

This holds for τ below the threshold:

$$\tilde{\tau} = \frac{1}{2}\left(-\lambda\gamma + \sqrt{\lambda^2\gamma^2 + 4\gamma(\kappa + c - \lambda r_{s,0})}\right). \tag{10}$$

3.4 Equilibrium

The optimal portfolio at and away from the constraint is

$$\alpha^* = \max\left\{\alpha^*_{\text{int}}, \alpha^*_{\text{bind}}\right\} = \max\left\{\lambda + \frac{\tau}{\gamma}, \ \frac{\kappa + c - \lambda r_{s,0}}{\tau}\right\}. \tag{11}$$

Equation (11) applies whenever the constrained portfolio is feasible, $\alpha^*_{\text{bind}} \leq 1$, which requires $\tau \geq \kappa + c - \lambda r_{s,0}$. At lower term premia, even a portfolio fully invested in the long-term bond cannot generate net income of κ, and the bank is unprofitable; Section 3.6 discusses how this boundary shifts with the bank's leverage capacity. The effects of changes in the term premium and short-term rates on the optimal portfolio depend on whether the profitability constraint binds, as highlighted in the following proposition.

Proposition 1 (Reach for Duration)**.** *Using the threshold $\tilde{\tau}$ for the term premium, we get:*

1. *For $\tau \geq \tilde{\tau}$ (unconstrained regime): the profitability constraint is slack and*

$$\alpha^* = \alpha^*_{\text{int}} = \lambda + \frac{\tau}{\gamma}.$$

A lower term premium tilts the portfolio towards shorter-term assets, while changes in the short-term interest rate have no effect on the optimal portfolio.

2. *For $\kappa + c - \lambda r_{s,0} \leq \tau < \tilde{\tau}$ (constrained regime): the profitability constraint binds and*

$$\alpha^* = \alpha^*_{\text{bind}} = \frac{\kappa + c - \lambda r_{s,0}}{\tau}.$$

The bank reaches for duration: α^* is decreasing in τ and $r_{s,0}$, with

$$\frac{d\alpha^*_{\text{bind}}}{d\tau} = -\frac{\kappa + c - \lambda r_{s,0}}{\tau^2} < 0 \qquad \textit{and} \qquad \frac{d\alpha^*_{\text{bind}}}{dr_{s,0}} = -\frac{\lambda}{\tau} < 0. \tag{12}$$

Lower interest rates or a lower term premium lead banks to invest in longer-term assets. Moreover, the relationship with the term premium is convex: the response to a further compression grows as the premium falls.

Proposition 1 shows that the model is consistent with the stylized facts of Section 2 if the profitability constraint is taken into account, but not otherwise. The historical evidence suggests that banks widen their maturity gaps in response to falling interest rates and term premia by raising the maturity of their asset positions. Proposition 1 shows that a standard model of portfolio choice based on mean-variance preferences is at odds with these stylized facts. The intuition for this result is that a decline in interest rates has no bearing on the portfolio choice since the excess return for holding long-term assets is unaffected. In addition, a falling term premium would lead to the counterfactual prediction that the bank invests more in short-term assets.

Taking the profitability constraint into account overturns these counterfactual predictions. A decline in interest rates reduces deposit spreads and banks' net interest income, which tightens the profitability constraint. As a result, the required duration to satisfy the profitability constraint rises, and the reach-for-duration region expands. Similarly, when the term premium is high, banks voluntarily overshoot their hedge to earn the term premium. A decline in τ reduces this overshoot. However, once τ falls below $\tilde{\tau}$, banks are squeezed: the only way to generate enough profits to cover operating costs is to load up on long-term assets. The flatter the yield curve becomes, the more duration the bank needs, creating a "reach for duration." Note that this effect is nonlinear: When the term premium is small, a further compression in τ would incentivize banks to reach for duration at an even stronger pace.

Quantitative Easing. One potential reason for a fall in the term premium may be quantitative easing. In such a program, the central bank purchases securities that it pays for with reserves. Since reserves carry zero maturity, a natural first-order view of QE is that it removes duration risk from bank balance sheets. Under this view, banks' maturity mismatch—and hence their exposure to interest rate risk—should decline during periods of large-scale asset purchases. In our model, when banks generate sufficient interest income, asset purchases by the central bank would incentivize banks to lower their long-term bond holdings, as the compensation for duration risk is reduced.

However, beyond the critical profitability threshold, QE has the opposite effect: it *increases* bank maturity mismatch. Thus, a policy that intends to remove duration risk from private portfolios ends up concentrating such risk within the banking sector. The reach-for-duration effect is more likely to arise in low-interest-rate environments, where returns to short-term assets and bank deposit franchise values are diminished, leaving little room to absorb bank operating costs. Since the zero lower bound is also the environment in which central banks resort to unconventional tools like QE to compress term premia, the two forces reinforce each other.

3.5 Graphical Illustration and Comparative Statics

Figure 5 illustrates the equilibrium in (τ, α) space for a calibration motivated by U.S. banking data. We set the deposit beta to $\beta = 0.44$, which is the asset-weighted average across U.S. commercial banks, following calculations based on Drechsler, Savov, and Schnabl (2021).[10] The operating cost is set to $c = 0.025$, matching approximately the ratio of aggregate non-interest expenses to total assets reported in the FDIC's Quarterly Banking Profile.[11] The equity ratio is calibrated to $e = 0.06$, in line with regulatory leverage requirements and conservative relative to the equity-to-assets ratio of about 10 percent in the data. We set $r_{s,0} = 0$ to capture the zero lower bound, such that asset income only derives from the term premium. We further assume that $\kappa = 0$, requiring the bank to make non-negative net income. The risk-cost parameter is set to $\gamma = 0.5$ for illustration. This calibration implies a hedge portfolio of 0.586, and a threshold $\tilde{\tau} \approx 0.038$.

The blue line in Figure 5 is the FOC curve $\alpha^*_{\text{int}} = \lambda + \tau/\gamma$, which is linear with intercept λ, the hedge portfolio. The orange line represents the curve where the profitability constraint binds, $\alpha^*_{\text{bind}} = (\kappa + c - \lambda r_{s,0})/\tau$, which is a decreasing convex function. They cross at $\tilde{\tau}$. The equilibrium path (thick line) follows the FOC for $\tau > \tilde{\tau}$ and switches to the portfolio at the binding constraint for $\tau < \tilde{\tau}$. For $\tau < \kappa + c - \lambda r_{s,0}$, even full duration ($\alpha = 1$) cannot deliver κ; at the ZLB calibration this boundary is exactly

[10] See https://pages.stern.nyu.edu/~pschnabl/data/data_deposit_beta.htm.

[11] See https://www.fdic.gov/quarterly-banking-profile.

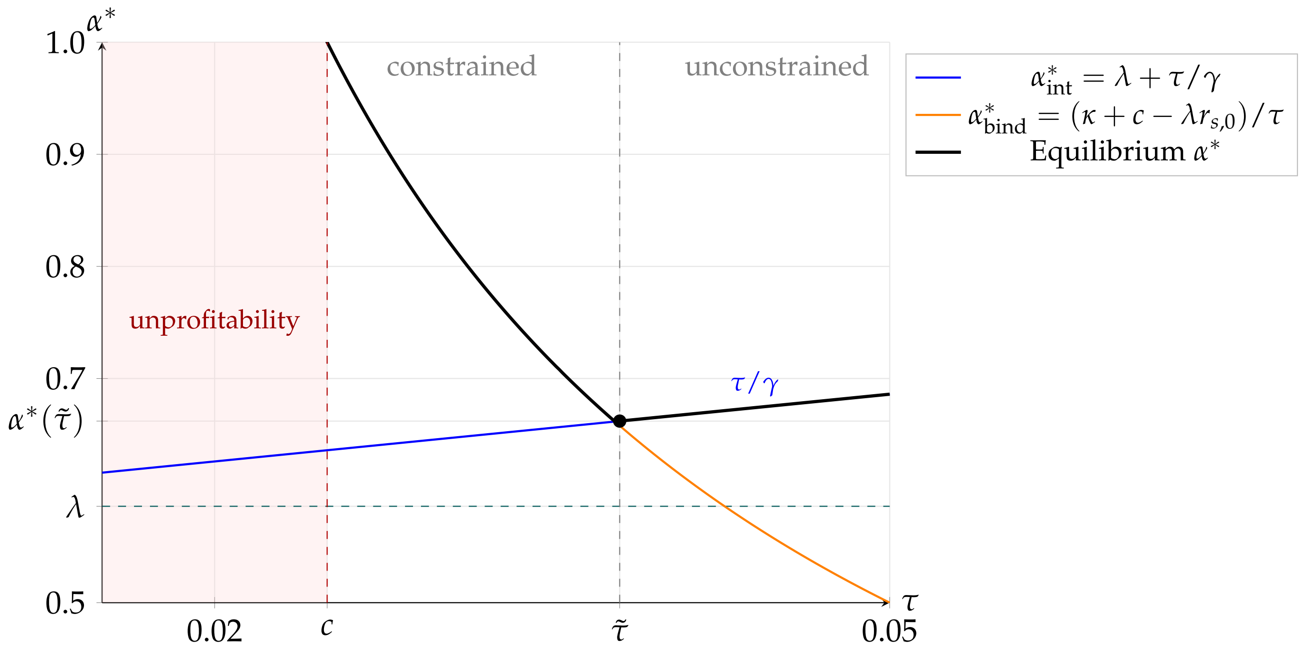


Figure 5: Equilibrium maturity mismatch at the zero lower bound ($r_{s,0} = 0$). Parameters: $\beta = 0.44$, $e = 0.06$, $c = 0.025$, $\kappa = 0$, $\gamma = 0.5$ (so $\lambda = 0.586$, $\tilde{\tau} \approx 0.038$).

$\tau = c$. Beyond the changes in interest rates and term premia, the model generates additional intuitive comparative statics:

Deposit market power (β). A bank with stickier deposits (lower β) has a larger hedge portfolio λ and holds more duration for any term premium τ in the unconstrained regime. It also has a higher non-repricing income $\lambda r_{s,0}$, which relaxes the profitability constraint: $\tilde{\tau}$ is lower, so the reach-for-duration region is smaller. Banks with more deposit market power are therefore both more heavily invested in duration and less vulnerable to the reach-for-duration pressure.

Equity (e). Higher equity increases λ, because $(1 - e)\beta$ falls, raising the hedging portfolio. It also raises the non-repricing income $\lambda r_{s,0}$, relaxing the profitability constraint. Well-capitalized banks hold more duration but are less likely to be squeezed by a decline in the term premium.

3.6 Discussion of Assumptions

The baseline model makes two simplifying assumptions that merit discussion: the portfolio is restricted to $\alpha \in [0, 1]$, and the deposit base is fixed at $1 - e$. Both turn out to be normalizations of the same underlying constraint—the bank's leverage capacity.

Portfolio restriction. If we allow $\alpha \in [0, \bar{\alpha}]$ for some $\bar{\alpha} > 1$, permitting the bank to borrow at the short rate, as in wholesale funding markets, and invest in the long bond, the unconstrained optimum $\alpha^*_{\text{int}} = \lambda + \tau/\gamma$ and the minimum required duration $\alpha^*_{\text{bind}} = (\kappa + c - \lambda r_{s,0})/\tau$ are unchanged, as is the threshold $\tilde{\tau}$, provided the cap does not bind at the interior optimum, $\lambda + \tau/\gamma \leq \bar{\alpha}$—which fails only for term premia above $\gamma(\bar{\alpha} - \lambda)$, far outside the range we consider. The only difference is the unprofitability condition: the bank becomes unprofitable when the required duration α^*_{bind} exceeds $\bar{\alpha}$, i.e., when $\tau < (\kappa + c - \lambda r_{s,0})/\bar{\alpha}$. Higher leverage capacity delays unprofitability in term premium space but does not eliminate the profitability constraint or the reach for duration. In practice, $\bar{\alpha}$ can be interpreted as a leverage constraint.

Fixed deposits. The same leverage capacity also disciplines the bank's deposit demand once the deposit base is made a choice variable. We leave the derivations for this case to Appendix D and briefly summarize the key findings here. Suppose the bank chooses its total assets A, and hence its deposits $A - e$, alongside the portfolio share, subject to $A \leq \bar{\alpha}$, while the per-period operating cost c is unchanged. Deposits are a strictly cheaper source of funds than wholesale borrowing: they cost $\beta r_{s,0}$ rather than $r_{s,0}$, and their non-repricing fraction funds long-term assets, so a marginal deposit dollar invested to leave the bank's interest rate risk exposure unchanged earns the deposit spread evaluated at the long-term rate, $(1 - \beta) r_\ell > 0$. The bank therefore issues deposits up to the limit, whether or not the profitability constraint binds, and the scale decision separates from the portfolio decision.

Given $A = \bar{\alpha}$, the analysis of the baseline goes through unchanged. In dollar terms, the unconstrained long-bond position equals the hedge position $\lambda_{\bar{\alpha}} \equiv e + (1 - \beta)(\bar{\alpha} - e) \geq \lambda$, the bank's non-repricing funding at capacity, plus the myopic component τ/γ, which increases in the term premium. The profitability constraint binds once compressed rates and term premia push the bank's net interest income sufficiently low. The constrained bank's dollar demand for the long-term bond is $(\kappa + c - \lambda_{\bar{\alpha}} r_{s,0})/\tau$, with portfolio shares obtained by dividing by $\bar{\alpha}$. Fixing the deposit base is therefore without loss of generality in this environment.[12]

3.7 Extensions

The baseline model abstracts from two features that may shape banks' responses to a fall in the term premium: the option to invest in risky loans alongside long-term bonds, and the feedback from banks' aggregate demand for duration onto the term

[12] When operating costs instead rise with the size of the deposit base, the balance sheet becomes interior. With a convex cost of deposits, scale is pinned by the deposit franchise, the profitability constraint binds for positive income targets, and the reach for duration effect is still present.

premium itself. Appendix Sections E and F develop both extensions in full; here we summarize the main findings and their implications.

Reaching for credit risk. The first extension allows the bank to invest in risky loans in addition to the short- and long-term assets of the baseline model. Loans earn a credit spread above the bond rate but expose the bank to a credit-loss shock that is orthogonal to interest rate risk. Two frictions discipline the loan portfolio: a downward-sloping credit demand curve, which compresses the marginal spread as lending expands, and credit risk, which the bank dislikes because shareholders are risk-averse. In the unconstrained regime, the bank's portfolio separates: total duration exposure is governed by the term premium as in the baseline, while the loan allocation is pinned by the balance of credit spreads against the two lending frictions.

Once the profitability constraint binds, loans become a tool for managing the shortfall. The bank expands lending as τ falls because spread income partially substitutes for duration income and reduces the required duration overshoot. Whether the loan share rises or falls depends on which friction dominates. When credit risk is the binding friction, the loan share rises as τ falls: the bank substitutes aggressively into loans because each additional loan generates spread income that lowers required duration. When diminishing returns dominate, the loan portfolio is pinned at the spread-maximizing level and the loan share falls: all duration adjustment flows through bonds, as in the baseline. The latter case is consistent with the experience of Silicon Valley Bank, whose limited lending opportunities in the technology sector coincided with a sharp increase in long-duration security holdings ahead of its failure.

Endogenous term premium. The second extension allows the term premium to respond endogenously to banks' aggregate demand for long-term assets, $\tau(\alpha) = \hat{\tau} - \mu\alpha$, where $\hat{\tau}$ is the exogenous component shaped by QE and μ measures the price impact of bank demand. The extension is motivated by evidence in Haddad and Sraer (2020) that banks' aggregate interest rate exposure predicts excess Treasury bond returns. The endogenous term premium component gives rise to a self-reinforcing mechanism. In the constrained regime, banks reach for duration, their collective demand compresses the term premium, the tighter profitability constraint requires yet more duration, and so on. The feedback amplifies the reach-for-duration mechanism: banks hold strictly more duration at every $\hat{\tau}$ than in the exogenous benchmark in the constrained regime, because their own demand compresses the term premium they are trying to earn.

Beyond amplification, an instability threshold emerges at $\hat{\tau}_{\text{crit}} = 2\sqrt{\mu(\kappa + c - \lambda r_{s,0})}$: below it, the feedback admits no fixed point at which banks meet the profitability constraint, so no equilibrium exists. The sensitivity of equilibrium duration to QE diverges as $\hat{\tau}$ approaches the threshold from above.

4 Deposit Franchise Runs

In this section, we extend the baseline model to include deposit franchise runs. The objective is to understand the extent to which reaching for duration can put banks in a fragile position in which self-fulfilling bank runs are more likely to occur. To this end, we build on Drechsler et al. (2026) and Jiang et al. (2024) who show that runs of uninsured depositors can occur even when bank assets are fully liquid.[13]

4.1 Setup and Run Condition

The setting is identical to Section 3, with two modifications. First, the bank's deposit base $1-e$ is split into insured and uninsured deposits. A fraction $u \in (0,1)$ of deposits is uninsured and the remainder $1-u$ is insured. Both types pay the same rate $\beta r_{s,0}$ without loss of generality, so all formulas from Section 3 carry through unchanged. The distinction between insured and uninsured deposits matters only for the possibility of a run. Second, we assume the interest rate shock is uniformly distributed, $\varepsilon \sim U[-\bar{\varepsilon}, \bar{\varepsilon}]$, so that $\mathbb{E}[\varepsilon] = 0$ and the run probability can be computed in closed form. Appendix G contains the derivations for this section.

Run condition. Following Drechsler et al. (2026), when uninsured depositors withdraw, the bank loses the deposit franchise it would have earned on its deposits in period 2. The bank meets uninsured deposit withdrawals by selling assets at their post-shock market value.[14] The lost deposit franchise reduces period 2 net income by

$$\Delta(\varepsilon) = (1-e)(1-\beta)(r_{s,0} + \varepsilon)u. \tag{13}$$

A run is self-fulfilling if, after losing the uninsured franchise, the bank's terminal net worth turns negative: $e + \Pi_2(\varepsilon, \alpha) < \Delta(\varepsilon)$.[15] Using (2) and dropping the small constant

[13] If banks instead respond to profitability pressure by expanding lending, reaching for credit risk as in Appendix E, the classic illiquidity channel of Diamond and Dybvig (1983) becomes relevant in addition to the franchise channel; see the discussion in Appendix E.7.

[14] The asset sale proceeds are assumed to suffice to repay the withdrawing uninsured depositors at face value. As a result, uninsured depositors who withdraw face no expected loss in a run and require no ex-ante risk compensation. In contrast, uninsured depositors who left funds in the bank would hold a claim on an insolvent institution and could not be repaid in full. Joining the run is therefore individually rational. Any remaining losses fall on bank equity and, if equity is exhausted, on insured deposits through the FDIC. An equivalent formulation would have the bank meet withdrawals by borrowing in the short-term funding market at rate $r_{s,0} + \varepsilon$, provided that the new wholesale creditors are paid in full at $t = 2$ and therefore likewise require no risk compensation.

[15] The market value at $t = 1$ is $V_1 = (e + \Pi_2)/(1 + r_{s,0} + \varepsilon)$, since the bank's remaining cash flow to shareholders at $t = 2$ is equity plus the terminal payout. The run condition $e + \Pi_2 < \Delta$ is therefore equivalent to $V_1 < \Delta_V$, where $\Delta_V(\varepsilon) = \Delta(\varepsilon)/(1 + r_{s,0} + \varepsilon)$ is the franchise loss in present-value terms. This condition aligns with Drechsler et al. (2026); see Appendix G for the equivalence. We work with terminal values for algebraic convenience.

$(1-e)(1-\beta)r_{s,0}u$ relative to equity e, the run condition reduces to $\varepsilon > \varepsilon^*(\alpha)$, where:[16]

$$\varepsilon^*(\alpha) = \frac{e + 2(\lambda r_{s,0} + \alpha\tau - c)}{\alpha - \tilde{\lambda}}, \qquad \tilde{\lambda} \equiv \lambda - (1-e)(1-\beta)u. \tag{14}$$

The threshold ε^* has the following interpretation. The numerator is the bank's equity plus its cumulative expected net income. The denominator is the bank's duration in excess of its run-robust funding: using $\lambda = e + (1-e)(1-\beta)$, the adjusted base $\tilde{\lambda} = e + (1-e)(1-\beta)(1-u)$ is the non-repricing funding that survives a run—equity plus the sticky share of insured deposits. Since the equilibrium portfolio always satisfies $\alpha^* \geq \lambda > \tilde{\lambda}$, the denominator is positive. More duration (higher α) or lower net income reduce ε^* and make runs more likely.

We impose two assumptions. First, when a self-fulfilling run is possible ($\varepsilon > \varepsilon^*$), it occurs (worst-case equilibrium selection). Second, the bank has sufficient equity that fundamental insolvency—negative terminal net worth, $e + \Pi_2 < 0$, absent any run—does not occur within the shock's support $[-\bar{\varepsilon}, \bar{\varepsilon}]$.[17]

4.2 Run Onset and the Three Regimes

At the constrained portfolio $\alpha^*_{\text{bind}} = (\kappa + c - \lambda r_{s,0})/\tau$, the profitability constraint binds with $\text{NI}_1 = \kappa$. The run threshold becomes:

$$\varepsilon^*(\tau) = \frac{(e + 2\kappa)\,\tau}{\kappa + c - \lambda r_{s,0} - \tilde{\lambda}\tau}, \tag{15}$$

which is increasing in τ: as the term premium falls, the bank is more fragile. Since $\varepsilon \sim U[-\bar{\varepsilon}, \bar{\varepsilon}]$, runs are possible when $\varepsilon^*(\tau) < \bar{\varepsilon}$, which defines the run onset threshold:

$$\tau_{\text{run}} = \frac{\bar{\varepsilon}\,(\kappa + c - \lambda r_{s,0})}{e + 2\kappa + \tilde{\lambda}\bar{\varepsilon}}. \tag{16}$$

For $\tau < \tau_{\text{run}}$, the constrained bank faces a positive probability of a self-fulfilling run, with run probability $p(\tau) = (\bar{\varepsilon} - \varepsilon^*(\tau))/(2\bar{\varepsilon})$ increasing as τ falls.

We assume that the parameters are such that $\tau_{\text{run}} < \tilde{\tau}$, which holds in our calibration ($\tau_{\text{run}} \approx 0.032 < \tilde{\tau} \approx 0.038$). The equilibrium therefore has three regimes:

1. **Unconstrained, no runs** ($\tau > \tilde{\tau}$): The profitability constraint is slack, $\alpha^* = \lambda + \tau/\gamma$, and $p = 0$.
2. **Constrained, no runs** ($\tau_{\text{run}} < \tau \leq \tilde{\tau}$): The profitability constraint binds, $\alpha^* =$

[16]Retaining the constant term would shift ε^* slightly without changing the qualitative results; at the zero-lower-bound calibration of Figure 6 ($r_{s,0} = 0$), the dropped constant is exactly zero.

[17]Fundamental insolvency requires $\varepsilon > (e + 2\text{NI}_1)/(\alpha - \lambda)$, which exceeds ε^* since $\tilde{\lambda} < \lambda$; see Appendix G. The assumption therefore requires $(e + 2\text{NI}_1)/(\alpha - \lambda) > \bar{\varepsilon}$ for all relevant α, ensuring that the only source of bank failure is a self-fulfilling run.

$(\kappa + c - \lambda r_{s,0})/\tau$, but the bank's buffer keeps $\varepsilon^* > \bar{\varepsilon}$, so $p = 0$.

3. **Constrained, run risk** ($\kappa + c - \lambda r_{s,0} < \tau \leq \tau_{\text{run}}$): The profitability constraint binds and $p(\tau) > 0$, increasing as τ falls.

The run onset threshold τ_{run} is decreasing in equity e (better-capitalized banks tolerate more duration before runs become possible) and increasing in the shock range $\bar{\varepsilon}$ (a more volatile rate environment expands the run-risk region). Indeed, a bank with sufficiently high equity is immune to runs at any term premium: since $\varepsilon^*(\tau)$ is lowest at $\tau = \kappa + c - \lambda r_{s,0}$ (where $\alpha = 1$), runs are impossible for all τ whenever $e \geq e^* \equiv (\bar{\varepsilon}\,\psi - 2\kappa)/(1 + \bar{\varepsilon}\,\psi)$, where $\psi = 1 - (1-u)(1-\beta)$ (see Appendix G).

It may appear inconsistent that a bank concerned with its survival reaches for duration and thereby exposes itself to a run, which threatens its survival as well. Two features reconcile this within our model. First, the two threats differ in nature: the profitability constraint reflects an immediate and certain consequence of low earnings—supervisory action and capital erosion in the current period—whereas a run materializes only in the future and only if a sufficiently large interest rate increase occurs. Second, the constrained bank already holds the smallest duration position consistent with remaining profitable, namely α^*_{bind}, the minimum share that satisfies the constraint. A bank that internalized run risk would wish to hold even less duration, but the constraint prevents it from doing so. Incorporating run aversion into the objective would therefore leave the equilibrium portfolio unchanged: wherever run risk is positive, at term premia below $\tau_{\text{run}} < \tilde{\tau}$, the bank is pinned at α^*_{bind}.

This reconciliation, and in particular the irrelevance of run aversion for the portfolio, rests on the profitability constraint being a hard floor. If the floor were instead a soft penalty that increased with the size of the earnings shortfall, a bank that internalized run risk would face a genuine trade-off: it could accept a larger shortfall, and the associated penalty, in exchange for a lower duration position and a smaller probability of a run. Our analysis isolates the polar case in which earnings are decisive—the cost of missing the target is large enough that the bank meets the floor regardless—and the resulting fragility is the unavoidable byproduct of doing so. More generally, a bank weighing the certain cost of an earnings shortfall against the contingent cost of a run would temper its reach for duration without reversing its direction, since lower term premia continue to push the earnings-constrained bank toward more duration.

4.3 Graphical Illustration

Figure 6 illustrates the three regimes using the Section 3 calibration with uninsured deposit share $u = 0.42$ and shock bound $\bar{\varepsilon} = 0.14$. The uninsured share is consistent with FDIC data: uninsured deposits accounted for approximately 42 percent of

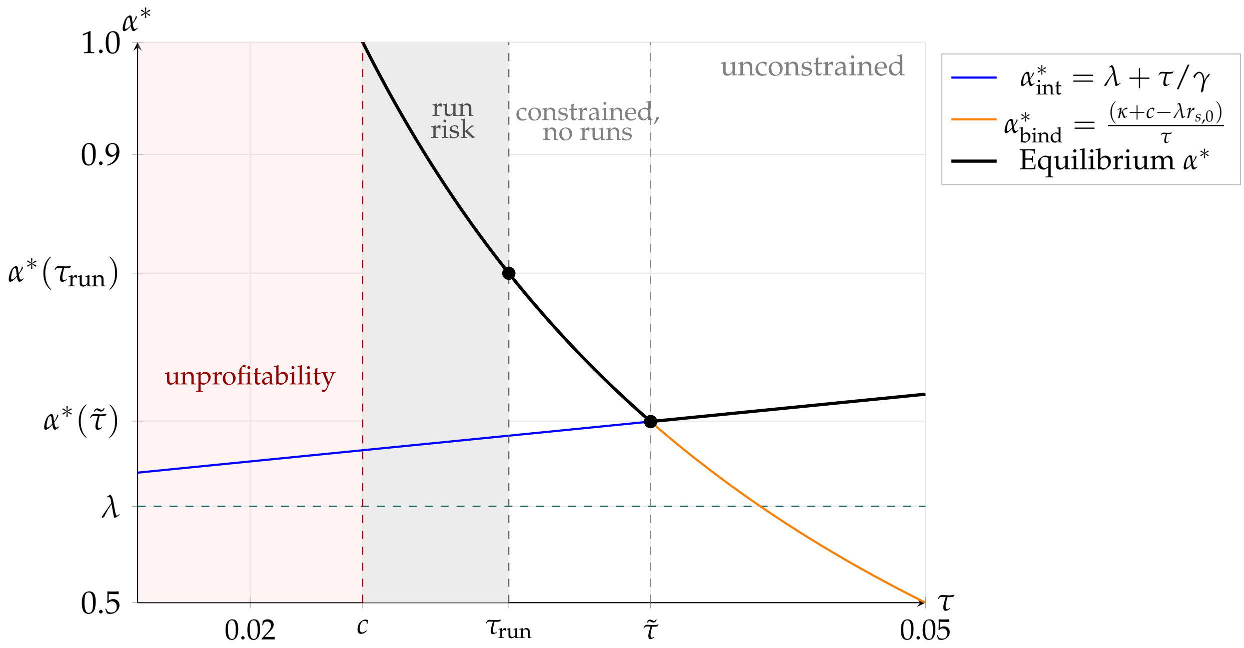


Figure 6: Equilibrium maturity mismatch with deposit franchise run risk. The blue and orange curves are as in Figure 5. To the left of τ_{run} (shaded gray), the constrained bank faces a positive probability of a self-fulfilling run by uninsured depositors. Parameters: $u = 0.42$, $\bar\varepsilon = 0.14$; all others as in Figure 5.

domestic deposits in the U.S. banking sector as of the fourth quarter of 2024.[18] The shock bound is chosen so that fundamental insolvency, i.e. negative terminal net worth $e + \Pi_2 < 0$ absent any run, never occurs: this requires $\bar\varepsilon < (e + 2\kappa)/(1-\lambda)$, which equals $e/(1-\lambda) \approx 0.145$ in the baseline calibration with $\kappa = 0$.[19] The blue and orange curves are identical to Figure 5. The vertical dashed line at $\tau_{\text{run}} \approx 0.032$ separates the constrained region into two zones: to its right, the bank's equity buffer keeps $\varepsilon^* > \bar\varepsilon$ and runs are impossible; to its left (shaded gray), the reach for duration has pushed α high enough that self-fulfilling runs become possible.

5 Dynamic Model

We extend the model presented in Section 3 to a continuous-time, dynamic environment and calibrate it for quantitative analysis. To this end, we make three key changes:

[18]The 42 percent share is computed from estimated uninsured deposits and total domestic deposits reported in Table II-A of the FDIC *Quarterly Banking Profile*, Fourth Quarter 2024; see FDIC (2025).

[19]At the most stressed constrained portfolio $\alpha = 1$ (equivalently $\tau = \kappa + c - \lambda r_{s,0}$), the insolvency threshold is $\tilde\varepsilon = (e + 2\kappa)/(1-\lambda)$, which reduces to $e/(1-\lambda) \approx 0.145$ for $\kappa = 0$. Setting $\bar\varepsilon = 0.14 < 0.145$ ensures that fundamental insolvency lies outside the shock support for all $\alpha \le 1$, so the only source of bank failure is a self-fulfilling run. This assumption also rules out risk-shifting incentives: under limited liability the payoff to shareholders would be $\max(e + \Pi_2, 0)$, which exceeds $e + \Pi_2$ in insolvency states and creates a "put option" whose value is increasing in duration α. By ensuring $e + \Pi_2 > 0$ for all $\varepsilon \in [-\bar\varepsilon, \bar\varepsilon]$, the put is worthless and the bank has no incentive to gamble on duration.

We consider persistent interest rate shocks, allow for time variation in the value of bank's assets, and let the bank trade financial instruments dynamically over time.

Based on this generalized framework, we derive three key takeaways. First, the main conclusions of the static model carry over to a more dynamic setup. Second, the reaching-for-duration effect already alters optimal choices in anticipation of, and thus away from, the occasionally binding profitability constraint. The effect becomes stronger the higher the likelihood of a future binding constraint. This prediction is in contrast to the stylized static model where the reaching-for-duration effect required a binding profitability constraint. Third, we conduct a quantitative evaluation that examines how much the model-implied maturity gap changes in response to declines in the short rate and the term premium of the magnitudes observed over the past 40 years. We find that the model captures the observed increase in the maturity gap almost exactly. Around half of the increase in the maturity gap is attributable to the decline in the term premium while the other half is due to the decline in the level of rates.

5.1 Setup

Time is continuous and denoted by $t \in [0, T]$. The representative bank solves a dynamic portfolio problem. The bank invests in long-term and short-term assets, and finances these positions with net worth and deposits. We start from the bank's balance sheet in levels instead of shares as in equation (1) of Section 3:[20]

$$L_t + B_t = N_t + D_t, \tag{17}$$

where L_t denotes the value of long-term assets, B_t are short-term assets, and D_t are deposits. Short-term assets pay the interest rate $r_{s,t}$,

$$\mathrm{d}B_t = r_{s,t} B_t \mathrm{d}t,$$

where the short rate $r_{s,t}$ fluctuates exogenously with shocks $W_{r,t}$, following an autoregressive process

$$\mathrm{d}r_{s,t} = \kappa_r \left(\bar{r} - r_{s,t}\right) \mathrm{d}t + \sigma_r \mathrm{d}W_{r,t},$$

with $W_{r,t}$ denoting a standard Brownian motion. The long-term assets' total return is given by

$$\mathrm{d}r_{\ell,t} = (r_{s,t} + \tau)\mathrm{d}t - \sigma_\ell \mathrm{d}W_{r,t},$$

where τ again represents the term premium and σ_ℓ is a parameter capturing the impact of short-rate shocks to the return of long-term assets. Naturally, we have $\sigma_\ell > 0$,

[20]We discuss the equivalence in the Appendix H.

because positive shocks to the short-term rate cause negative returns in long-term assets. The bank pays its deposit base D_t a flow equal to $\beta r_{s,t} D_t \mathrm{d}t$ each instant, with $\beta \in [0,1]$ again representing the deposit beta. The bank also has to pay a cost to maintain the deposit franchise, c per unit of deposits, such that the total cost is cD_t.[21] The evolution of the bank's wealth is therefore

$$\mathrm{d}N_t = \left[(r_{s,t} + \tau)L_t + r_{s,t}B_t - \beta r_{s,t}D_t - cD_t\right]\mathrm{d}t - \sigma_\ell L_t \mathrm{d}W_{r,t}.$$

As in the two-period model, shareholders have preferences over terminal wealth, with risk aversion ρ.[22] At a given t, the optimization problem is

$$V_t = \max_{\{L_s/N_s\}_{s\in[t,T]}} \mathbb{E}_t \left[\exp\left(-\omega\left(T-t\right)\right)\frac{N_T^{1-\rho}}{1-\rho}\right], \tag{18}$$

subject to

$$\begin{aligned} \frac{\mathrm{d}N_t}{N_t} &= \left[r_{s,t} + \frac{L_t}{N_t}\tau + \left(\left(1-\beta\right)r_{s,t} - c\right)\frac{D_t}{N_t}\right]\mathrm{d}t - \frac{L_t}{N_t}\sigma_\ell \mathrm{d}W_{r,t}, &(19)\\ \mathbb{E}_t\left[\frac{\mathrm{d}N_t}{N_t}\right] &\geq \kappa\,\mathrm{d}t, &(20)\\ N_0 &> 0. &(21)\end{aligned}$$

Equation (20) is the dynamic counterpart of the profitability constraint in our baseline model and can bind occasionally. As in Section 3, it requires the bank's expected profitability, scaled by equity, to clear a floor each instant. It thereby only constrains the drift of bank wealth, its expected income, and not the diffusion term. Like its static analog, the constraint cannot be relaxed through financial transactions: The bank can raise its expected profitability only by tilting its portfolio toward the higher-yielding asset. For simplicity, as in Section 3, we assume that banks do not choose the amount of deposits and that the term premium is exogenously given. Deposits evolve according to $D_t = \delta N_t$, so the deposit-to-wealth ratio is constant, and the term premium is constant at τ, which we alter in comparative statics below.[23]

[21]Note that the scaling by deposits is not material in the static model when deposits are exogenous.
[22]We include a time preference parameter in banks' problem, $\omega > 0$, although it does not affect shareholders' marginal decisions.
[23]Extending the model to incorporate persistent term premium shocks does not affect our key results regarding the role of the profitability constraint in banks' optimal decisions.

Solution. We denote the bank's portfolio choice by $\alpha_t = L_t/N_t$.[24] Using the homogeneity of the objective function and the constraints, the global value function takes the form

$$V^*(t, N, r_{s,t}) = \xi^*(t, r_{s,t}) \frac{N^{1-\rho}}{1-\rho} \tag{22}$$

for some unknown function $\xi^*(t, r)$ with terminal condition $\xi^*(T, r) = 1$. The endogenous variable $\xi^*(t, r)$ characterizes the solution to the bank's problem and it captures the bank's future investment opportunities and profits. Let $\alpha^C(r)$ be the duration choice when the profitability constraint (20) binds,

$$\alpha^C(r_{s,t}) = \frac{\kappa + c\delta - [1 + \delta(1-\beta)]\, r_{s,t}}{\tau}.$$

Using the value function (22), the problem (18)–(21) can be characterized recursively by the following Hamilton–Jacobi–Bellman (HJB) equation

$$\begin{aligned} \frac{\omega}{1-\rho} =& \frac{1}{1-\rho}\frac{\xi_t^*}{\xi^*} + \frac{1}{1-\rho}\frac{\xi_r^*}{\xi^*}\kappa_r(\bar{r} - r_{s,t}) + \frac{1}{2(1-\rho)}\frac{\xi_{rr}^*}{\xi^*}\sigma_r^2 \\ &+ \max_{\alpha \geq \alpha^C(r_{s,t})} \left\{ \alpha\tau + [1 + \delta(1-\beta)]\, r_{s,t} - c\delta - \frac{\rho}{2}\alpha^2\sigma_\ell^2 - \frac{\xi_r^*}{\xi^*}\alpha\sigma_\ell\sigma_r \right\}, \end{aligned} \tag{23}$$

where ξ_t^* represents the time-derivative of ξ^*, ξ_r^* its derivative with respect to $r_{s,t}$, and ξ_{rr}^* denotes the respective second derivative. The first-order condition is given by

$$\tau - \rho\alpha^*(t, r_{s,t})\sigma_\ell^2 - \frac{\xi_r^*(t, r_{s,t})}{\xi^*(t, r_{s,t})}\sigma_\ell\sigma_r \leq 0,$$

with equality when the profitability constraint is slack. Therefore, when unconstrained, the bank chooses a portfolio

$$\alpha^U(t, r_{s,t}) = \frac{\tau}{\rho\sigma_\ell^2} - \frac{\xi_r^*(t, r_{s,t})}{\xi^*(t, r_{s,t})}\frac{\sigma_r}{\rho\sigma_\ell}. \tag{24}$$

Note that (24) is the standard portfolio expression in Merton (1971). As in the two-period model discussed in Section 3, banks' holdings of the long-term asset are motivated by a *myopic* component—the first term in (24)—and a *hedging* component—the second term in (24). The difference, though, is that in the dynamic model the hedging motive is driven by the covariance between changes in the banker's investment opportunity set and the banker's realized returns on wealth (the last term of the HJB); while in the two-period model it was due to the variance of the payoffs. In both cases,

[24] In the two-period model, the choice variable is the fraction of long-term assets as a share of total assets. In the dynamic model, this quantity is simple the scaled $\alpha_t/(1+\delta)$, which comes from the accounting identity (17) and the constant deposit-to-wealth ratio. We provide a detailed mapping between the two-period model and the dynamic model in the Appendix.

hedging motives are strongly driven by banks' ability to issue deposits, as we discuss below. Combining the portfolio choices, the bank's global portfolio choice is

$$\alpha^*(t, r_{s,t}) = \max\left\{\alpha^U(t, r_{s,t}), \alpha^C(r_{s,t})\right\},$$

with an endogenous boundary between the slack and binding regions. As opposed to the two-period model, the optimal unconstrained portfolio choice is influenced by the possibility of becoming constrained in the future. The logic is that the continuation value, ξ^*, determines the unconstrained solution and thus the boundary where the constraint binds, whose location feeds back into the continuation value.

The key implication from a potentially binding constraint in the future is that the bank may reach for duration before the profitability constraint binds. As the probability of hitting the constraint rises, the global continuation value becomes more sensitive to the short rate. Such changes in the sensitivity of the bank's continuation value affect the hedging component and induces the bank to increase duration exposure even in states in which the constraint is not currently binding.

This result is driven by the hedging component, more precisely by the semi-elasticity of the value function with respect to changes in the short rate, $\xi_r^*(t, r_{s,t})/\xi^*(t, r_{s,t})$, which captures how sensitive the bank's investment opportunity set is to the short rate $r_{s,t}$. Intuitively, the bank dislikes states in which profitability is low and its wealth is too volatile due to the possibility of a binding profitability constraint, causing $\xi_r^*(t, r_{s,t})$ to be relatively more sensitive to changes in $r_{s,t}$ than when the constraint is slack. In Proposition 2, we derive a closed-form solution for the semi-elasticity in the polar cases in which the constraint is either always binding or always slack. We find that banks' deposit market power, as well as other key parameters discussed in the proposition, play a key role in shaping the strength of the hedging demand.

Proposition 2 (Polar cases in the dynamic model)**.** *Suppose the profitability constraint is either always slack or always binding. If the profitability constraint is always slack (so that the bank is always unconstrained), the continuation-value function is*

$$\xi^{AU}(t, r_{s,t}) = \exp\left(\Omega^{AU}(t) + \Lambda^{AU}(t) r_{s,t}\right), \qquad \frac{\xi_r^{AU}(t, r_{s,t})}{\xi^{AU}(t, r_{s,t})} = \Lambda^{AU}(t),$$

where

$$\Lambda^{AU}(t) = \frac{1-\rho}{\kappa_r}\left[1 + \delta(1-\beta)\right]\left(1 - \exp\left(-\kappa_r(T-t)\right)\right). \tag{25}$$

If the profitability constraint is always binding,

$$\alpha^C(r_{s,t}) = \frac{\kappa + c\delta - \left[1 + \delta(1-\beta)\right] r_{s,t}}{\tau}$$

and

$$\xi^C(t, r_{s,t}) = \exp\left(\Omega^C(t) + \Lambda^C(t) r_{s,t} + \Psi^C(t) r_{s,t}^2\right),$$

so that

$$\frac{\xi_r^C(t, r_{s,t})}{\xi^C(t, r_{s,t})} = \Lambda^C(t) + 2\Psi^C(t) r_{s,t}. \tag{26}$$

The expressions for $\Psi^C(t)$ and $\Lambda^C(t)$ are shown in the appendix.

Equation (25) shows the semi-elasticity of the value function with respect to the short rate for the case in which the constraint is always slack. The deposit spread, $1 - \beta$, and the amount of deposits, represented by δ, are key parameters in shaping the semi-elasticity, and hence the bank's dynamic hedging motives. As we illustrate in the appendix, the semi-elasticity in the constrained case, (26), is lower, and thus more negative, than in the unconstrained case. Intuitively, when constrained, expected returns are low while the bank's wealth is more volatile. As we discuss below in the global solution, the bank's semi-elasticity, and hence its dynamic hedging motives, is stronger than in the always unconstrained case, even if the bank is currently unconstrained. This is due to the presence of the occasionally binding profitability constraint: Even though the bank may not be constrained, the probability of a binding profitability constraint in the future, which occurs when profits are low, strengthens the bank's dynamic hedging motive ex-ante.

Numerical solution to the global problem. We solve the global HJB numerically and describe the procedure in Appendix H. Proposition 2 provides the intuition for the solution in the two polar cases, that is, when the constraint is always binding or always slack. This exercise is useful because the global solution can be thought as a mix of the two polar cases. Our parameterization, which is chosen to match the two-period model as closely as possible, is shown in Table 1. For the interest-rate dynamics, we set the parameters of the short-rate process, κ_r, $\bar{r}$, and σ_r, to approximately match the persistence, mean, and standard deviation of the short rate in the sample discussed in Section 2. We set the term premium to 2%, also to approximately match the sample average in the evidence discussed in Section 2. We then conduct comparative statics with respect to these parameters. Finally, we set κ to 3% as a plausible lower bound for return on equity, which is the empirical counterpart of $\mathbb{E}_t[dN_t/N_t]$. The qualitative results do not depend on the specific value of κ. The remaining parameters come from Section 3 and are approximately consistent with the banking evidence.

Figure 7 shows the global solution with an occasionally binding profitability constraint. Panel A plots the continuation-value index $\xi^*(t, \bar{r})$ at the unconditional mean of the short rate. Since $V_N^*(t, N, r_{s,t}) = \xi^*(t, r_{s,t}) N^{-\rho}$, ξ^* can be interpreted as the marginal continuation value of wealth, relative to the normalization $\xi^*(T, r_{s,t}) = 1$ at

Table 1: Parameters for the dynamic model

Symbol	Value	Description
ω	0.04	Discount rate
ρ	2.0	Risk aversion
β	0.44	Deposit pass-through
δ	15.67	Deposit-to-net-worth ratio
c	0.025	Cost per unit of deposits
κ_r	0.05	Mean reversion of the short rate
$\bar{r}$	0.04	Long-run short rate
σ_r	0.01	Short-rate volatility
σ_ℓ	0.10	Exposure of long asset to rate shocks
τ	0.02	Term premium
κ	0.03	Profitability bound

the terminal date. Early on, this marginal value is high because the bank internalizes the possibility that the profitability constraint may bind in future states. In those states, the bank loses portfolio flexibility and wealth dynamics become more volatile. As a result, an additional unit of wealth is especially valuable as a buffer. As time passes, there are fewer instances for the constraint to bind and this precautionary component declines. Near the terminal date, the continuation-value index converges back to the terminal condition $\xi^*(T, r) = 1$.

Panel B plots $\xi^*(0, r)$ across the state space $r_{s,t}$. The figure shows that the marginal value of wealth is especially high when the short rate is low. In low-rate states, the profitability constraint is more likely to bind and the constrained duration position $\alpha^C(r_{s,t})$ is larger. As a result, the bank must take a more volatile duration position to satisfy the constraint, making an additional unit of wealth particularly valuable. As the short rate rises, the deposit franchise becomes more profitable, the profitability constraint is less relevant, and the marginal continuation value of wealth declines sharply.

Panel C plots the global duration policy at $r = \bar{r}$. The blue line is the actual global policy $\alpha^*(t, \bar{r})$. The purple dashed line is the unconstrained candidate computed using the global continuation value, α_t^U, while the yellow dotted line is the profitability lower bound, α^C. The red dashed line is the always-unconstrained benchmark, α^{AU}.

The key result is that the global unconstrained candidate lies above the always-unconstrained benchmark. This is the case because the global continuation value internalizes the possibility that the profitability constraint may bind in future states. Since $\sigma_\ell > 0$, a more negative semi-elasticity ξ_r^*/ξ^* raises the hedging component of duration demand. Therefore, even when the bank is not currently constrained, it chooses more duration than in the always-unconstrained benchmark as a hedge against future constrained states. Over time, there is less remaining time for the constraint to bind, the hedging motive weakens, and α_t^U declines. Once α_t^U falls below the profitability lower bound, the actual policy follows α^C.

Panel D plots the semi-elasticity with respect to the short rate, ξ_r^*/ξ^*, which governs the hedging component of duration demand. This object describes the mechanism related to Panel C well. Since $\sigma_\ell > 0$, a more negative semi-elasticity raises the hedging demand for long-term assets. In the global problem, ξ_r^*/ξ^* is more negative than in the always-unconstrained benchmark because the bank internalizes that low-rate states are states in which the profitability constraint is more likely to bind. In those states, wealth has a higher marginal continuation value, so the continuation value becomes more sensitive to the short rate. This stronger sensitivity increases the hedging motive and leads the bank to choose more duration even before the constraint binds. As time passes, there is less remaining time for the constraint to bind, so this additional precautionary component fades and the global semi-elasticity moves closer to the always-unconstrained benchmark.

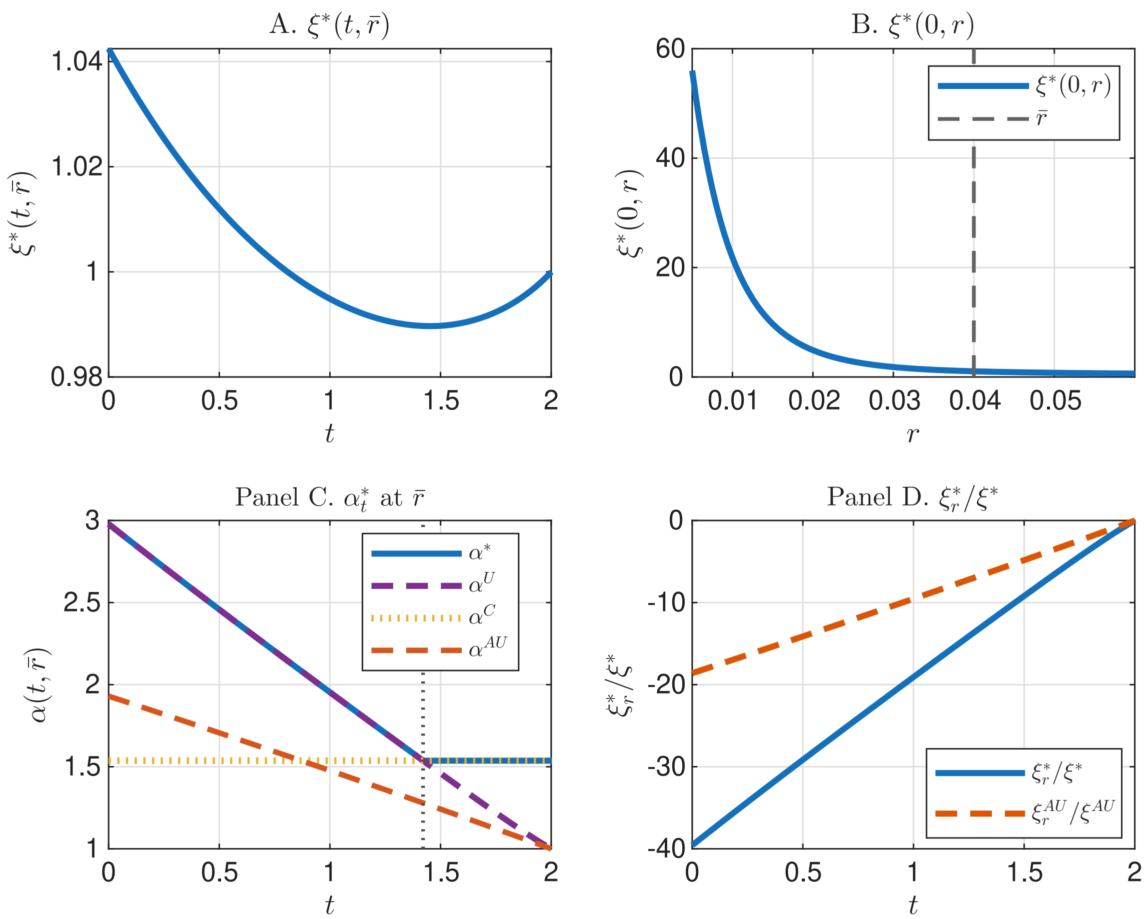


Figure 7: Global solution with an occasionally binding profitability constraint. Panel A shows the continuation-value index $\xi^*(t,\bar{r})$ across time. Panel B shows $\xi^*(0,r_{s,t})$ across the state space. Panel C shows the global duration policy α_t^*. Panel D shows the semi-elasticity ξ_r^*/ξ^*, which determines the hedging component of duration demand.

We next solve the global problem for different values of the term premium τ. Figure 8 shows the resulting policy at a fixed point in time and at $r_{s,t} = \bar{r}$. As the term premium declines, the profitability lower bound becomes tighter. However, the bank does not wait until the constraint binds. Instead, the global policy reacts in advance because the continuation value already reflects the risk of future binding constraints.

This result highlights the main distinction between the static and dynamic models. In the static model, reaching for duration occurs when the profitability constraint binds contemporaneously. In the dynamic model, reaching for duration can also arise in the slack region because the bank hedges against the possibility of becoming constrained in the future.

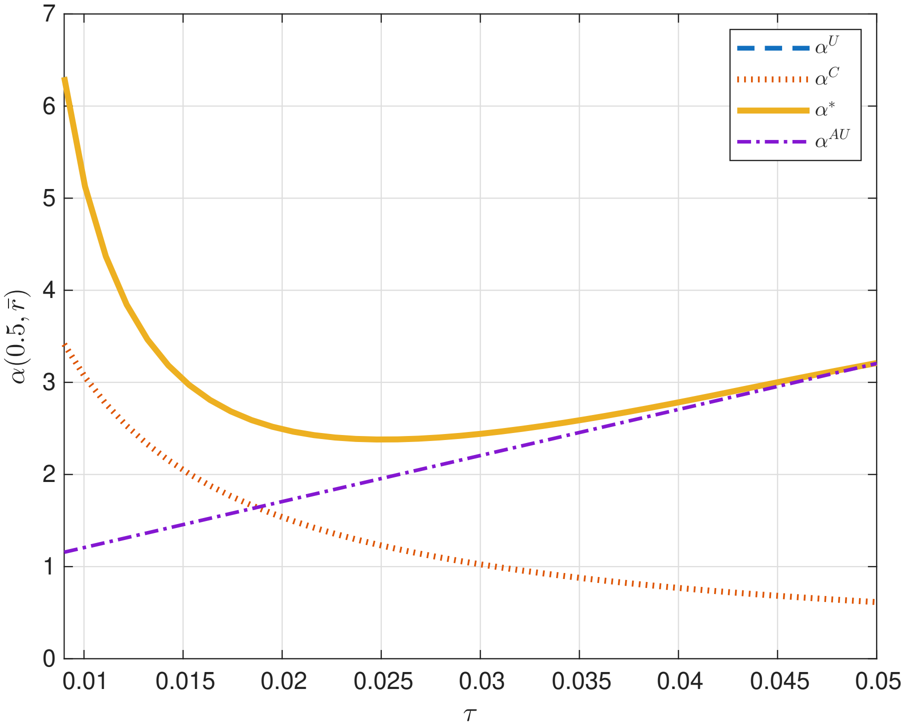


Figure 8: Global duration policy for different values of the term premium. The figure shows that the bank increases duration exposure as the term premium declines, even in regions in which the profitability constraint is not currently binding. The anticipation of future binding constraints affects the global continuation value and strengthens the hedging demand for duration. Note that α^U lies behind the yellow line, α^*, because the bank is unconstrained in the global solution.

Quantitative evaluation. We conclude this section by evaluating the quantitative implications of the dynamic model. To this end, we compare the model-implied maturity gap under two interest-rate and term-premium environments: One in which both are high, and one in which both are low. In each case, we solve the global dynamic problem and evaluate the bank's optimal duration policy at $t = 0$.

We discipline the levels of $r_{s,t}$ and τ using the evidence documented in Figure 2 for the Federal funds rate and the 10-year term premium. For the high-rate environment, we set $\tau = 4\%$ and evaluate the solution at $r_{s,t} = 8\%$, matching the average levels observed in 1985. For the low-rate environment, we set $\tau = 0.6\%$ and evaluate the solution at $r_{s,t} = 4.2\%$, corresponding to the average levels observed in 2025. We then compare the resulting model-implied maturity gaps with their empirical counterparts in 1985 and 2025.

To map the model into a maturity gap, note first that the dynamic model defines $\alpha_t = L_t / N_t$, while total assets are $L_t + B_t = N_t + D_t = (1 + \delta) N_t$. Thus, the share of total assets allocated to the long-term asset is

$$\frac{L_t}{L_t + B_t} = \frac{\alpha_t}{1 + \delta}.$$

Second, the duration of the long-term asset is implied by its return sensitivity to short-rate shocks. Since a short-rate shock of size $\sigma_r dW_{r,t}$ produces a long-asset return innovation of $-\sigma_\ell dW_{r,t}$, the implied modified duration of the long-term asset is

$$D_L = \frac{\sigma_\ell}{\sigma_r}.$$

To express this duration in years of maturity, we interpret the representative long-term asset as a zero-coupon bond, or equivalently use the standard zero-coupon approximation under which duration and maturity coincide. Finally, because liabilities have instantaneous maturity in the model, their maturity is zero. The model-implied maturity gap is therefore

$$MG_t = \frac{\alpha_t}{1 + \delta} \frac{\sigma_\ell}{\sigma_r}.$$

Figure 9 shows the results. The first and second bars, in blue and gray, show the maturity gap in the model and in the data, respectively, when the short rate and the term premium are high, at the levels observed in 1985. The third and fourth bars show the maturity gap, in the model and in the data, respectively, when using relatively lower levels of the term premium and the short rate as observed in 2025. The model matches the observed level of the maturity gap almost exactly.

The colored bars decompose the maturity gap into contributions of the decline in the level of the short rate versus the decline in the level of the term premium. To this end, we conduct an Aumann–Shapley attribution exercise.[25] The red and yellow bars show how much of the increase in the model-implied maturity gap, which moves from 1.76 years to 4.83 years once the short rate and the term premium decline, can be attributed to the short rate and the term premium. Approximately 55% of the increase

[25] The attribution lets $r_{s,t}$ and τ decline jointly and proportionally from their 1985 to their 2025 values and assigns each variable the integral of its marginal contribution along the path; the two shares sum exactly to the total change and, unlike sequential decompositions, do not depend on the order of the two declines.

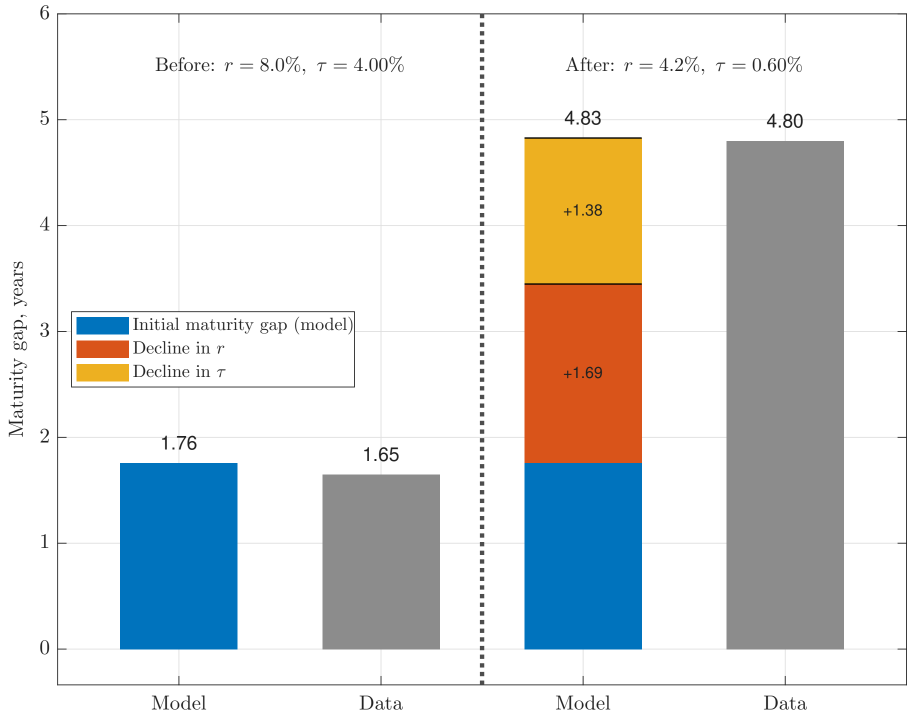


Figure 9: Model-implied and empirical maturity gaps. The data in the “before" (“after") bucket corresponds to the maturity gap, the level of the federal funds rate, and the 10-year term premium in 1985 (2025).

is due to the decline in the level of rates, while the remaining 45% can be attributed to the decline in the term premium. Thus, both the decline in the term premium and the decline in the level of rates are crucial for rationalizing the increase in the maturity gap.

In Appendix H, we show that the increase in the maturity gap is driven by banks' concerns about profitability: In the low-rate and low-term-premium environment, the bank is not constrained but concerned about the probability of being so at some point in the future. In contrast, absent the profitability constraint, the bank would have decreased its maturity gap in response to the decline in rates and term premia, highlighting the importance of incorporating concerns about profitability into banks' duration decisions.

Finally, we use the model to explore an alternative explanation for the secular increase in bank maturity mismatch: a concurrent increase in bank deposit market power, possibly driven by a switch from term deposits to savings deposits (Supera, 2021). Assuming a deposit beta of around half of its original calibration, we find that such a structural change barely moves the maturity gap in isolation, and even leads to a sharp decrease of it in the low-rate and low-term-premium environment as banks' concerns about profitability are alleviated.

6 Bank-Level Evidence

Section 2 documents a more-than-threefold widening of the maturity gap of U.S. commercial banks since the mid-1980s, in close correspondence with the secular decline in interest rates and term premia. The aggregate pattern is consistent with the reach-for-duration mechanism developed in Section 3, but could also be driven by other developments in recent decades, such as a structural rise of agency mortgage securitization, evolving bank regulation, or a shift in the asset preferences of long-duration investors. To distinguish our mechanism from such alternatives, we turn to a sharper cross-sectional prediction: banks whose profitability is squeezed should subsequently extend the duration of their assets more than banks whose profitability is not. We test this prediction in the cross-section using Call Report data for FDIC-insured banks.

Across a battery of specifications, we find that banks with lower net income relative to assets significantly raise their maturity gap in the following quarters. The headline coefficient is stable across specifications that include various fixed effects and bank controls, and robust to replacing net income over assets with two alternative profitability measures that the model also speaks to. In the appendix, we further document that the maturity extension is concentrated on the asset side of the balance sheet and that the response is amplified when term premia are compressed—whether measured by term structure estimates or by the Federal Reserve's securities holdings as an indicator of the stance of quantitative easing. Taken together, the evidence supports the model's central mechanism: when banks face profitability pressure, they extend duration to earn additional interest income.

6.1 Empirical Strategy

The model in Section 3 predicts that banks reach for duration when their profitability constraint binds. While several primitives can account for the profitability pressure—a low term premium, a low short rate that compresses deposit spreads, or high operating costs—they all imply that bank net income is too low relative to a bank's target. Net income is, in this sense, a sufficient statistic for the various forces that push a bank toward duration extension. We exploit this implication for our empirical specification and use a bank's profitability, scaling net income either by a bank's assets or its core income, as the regressor of interest. Since it is challenging to find a plausible instrument for profitability across a large panel of banks, we instead assess the stability and statistical significance of our estimates across various empirical specifications and discuss the direction of possible biases. For bank i at time t, our baseline specification is:

$$\text{Maturity Gap}_{i,t} = \phi_0 + \beta\, Z_{i,t-1} + \theta\, \text{Maturity Gap}_{i,t-1} + u_{i,t}, \tag{27}$$

where $Z_{i,t-1}$ is a measure of the bank's profitability and we control for the lagged dependent variable to isolate changes in the maturity gap. Our baseline measure of Z is $\text{NetIncome}_{i,t-1}/\text{Assets}_{i,t-1}$. We also consider two alternatives: $\text{NetIncome}_{i,t-1}/\text{NII}_{i,t-1}$, which scales net income by the bank's core income (net interest income, denoted NII), and $\text{NoninterestExpenses}_{i,t-1}/\text{NII}_{i,t-1}$, which captures the burden of operating costs relative to core income. NetIncome/Assets is expressed in percent, so a one-unit change corresponds to one percentage point of quarterly return on assets; the two NII-based measures are unit-free ratios. Our model predicts that $\beta < 0$ for the first two measures (lower profitability leads to subsequent maturity extension) and $\beta > 0$ for the third (higher cost burden leads to subsequent maturity extension). We consider the lagged level of Z, as opposed to its recent change, because the model's mechanism speaks to the level of profitability relative to a target rather than to recent dynamics: a bank can experience declining net income while still operating well above its profitability constraint. Since we do not have information on banks' profitability targets, which may well be implicit and can differ in the cross-section of banks, we are unable to consider deviations from targets.

Starting from this baseline, we progressively saturate regression (27). We add time fixed effects ϕ_t to absorb aggregate variation in the maturity gap and bank profitability. We next include bank fixed effects ϕ_i to absorb time-invariant heterogeneity in banks' maturity structures and profitability. Finally, we include a standard vector of controls: log assets, liabilities over assets, deposits over assets, securities over assets, and a once-further-lagged value of Z to account for possible correlations with various bank characteristics and a bank's recent profitability trajectory.

While these additional controls and fixed effects diminish possible omitted variable biases, our estimates may still be subject to a reverse-causality concern. However, such a bias likely works in the opposite direction of the model's mechanism: a bank that previously extended its maturity gap would subsequently earn higher interest income and therefore have higher net income, generating a *positive* coefficient on lagged profitability. In this sense, estimated coefficients with $\beta < 0$ can be interpreted as conservative: if reverse causality is operative, it biases these estimates toward positive values.

The dependent variable is the maturity gap measured in years. A one-quarter change in it, however, reflects the renewal of only a small fraction of the bank's balance sheet, since costs and constraints prevent banks from selling legacy assets and quickly revamping their positions. To make β interpretable as the net maturity of newly chosen investments, we multiply the maturity gap by $(2M - 1)$, where M is the average maturity of bank assets across all banks over the sample (in quarters). The transformation rests on three assumptions: the adjustment occurs on the asset side of the balance sheet, with liability maturity unchanged—consistent with the decomposition in Appendix I, where the maturity gap response coincides with the asset-maturity response and the response of liability maturity is zero and statistically insignificant; asset maturities are distributed approximately uniformly around the average M; and a fraction $1/(2M-1)$ of the asset portfolio matures and is renewed each quarter. Under this normalization, a one-unit change in the dependent variable corresponds to a one-year increase in the maturity of newly chosen assets relative to the level that keeps the maturity gap constant. In Appendix I.2, we consider the response of the maturity gap over time, adjusting dynamically for the growing fraction of assets that can be renewed.

Data and sample. We use quarterly Call Report data from 1997:Q2 through 2026:Q1. The starting date coincides with the introduction of the post-1997 maturity schedule, which provides the finer maturity bins to construct the maturity gap (see Section 2). Given these data, we apply the following sample restrictions and cleaning steps: (i) we drop non-depository trust companies, cooperative banks, and foreign banking organizations from the sample, (ii) we exclude observations with quarter-on-quarter asset growth above 100 percent (typically reflecting mergers), (iii) we drop observations that violate basic balance sheet logic such as negative assets or liabilities, deposits exceeding liabilities, or individual asset classes exceeding total assets, and (iv) we exclude weighted maturity observations if the underlying asset or liability values deviate from the corresponding balance sheet values by more than 10 percent in absolute terms. In addition, the profitability measures and the income-based outcome variables are trimmed at their 1st and 99th percentiles to limit the influence of outliers. Table A.12 lists the Call Report codes used to construct all non-maturity variables; the maturity and repricing variables underlying the maturity gap are documented in Tables A.6 and A.10.

6.2 Results

Table 2 reports the main estimates based on two-way clustered standard errors by bank and time. For our baseline profitability measure $Z = \text{NetIncome}/\text{Assets}$, columns (i)–(v) progressively add fixed effects and controls. The coefficient on Z is negative, consistent with our model prediction, and statistically different from zero at the 1 percent level in every column. Across specifications, the magnitude of the coefficient is stable: the simple regression in column (i) delivers $\beta = -0.54$, and the most saturated specification with two-way fixed effects and the full control vector in column (v) gives $\beta = -0.40$. Quantitatively, these coefficients imply that a one-percentage-point decline in a bank's return on assets, which is close to a one-standard deviation change within our sample, is associated with a subsequent extension of the net maturity of

Table 2: Maturity Gap and Bank Profitability.

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
NI/Assets	-0.54***	-0.45***	-0.42***	-0.43***	-0.40***		
	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)		
NI/NII						-0.64***	
						(0.18)	
Nonint.Exp./NII							1.23***
							(0.28)
Time FE		✓	✓	✓	✓	✓	✓
Firm FE			✓		✓	✓	✓
Controls				✓	✓	✓	✓
Partial R^2	0.01	0.04	0.08	0.04	0.08	0.08	0.08
Observations	530,026	530,026	529,735	516,792	516,540	517,305	517,869
Number of Banks	11,201	11,201	10,910	10,989	10,737	10,739	10,730

Notes: Estimation results for regression (27). NI/Assets refers to net income over assets, NI/NII to net income over net interest income, and Nonint.Exp./NII to noninterest expenses over net interest income. NI/Assets is expressed in percent; the NII-based measures are unit-free ratios. Controls are log assets, liabilities over assets, deposits over assets, securities over assets, and a once-further-lagged value of the respective profitability measure. The partial R^2 conditions on the variation explained by the lagged dependent variable. Standard errors in parentheses are two-way clustered by bank and time. Sample: 1997:Q2–2026:Q1. Notation: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

newly chosen assets of around five to six months.

Columns (vi) and (vii) report the estimates for the two alternative profitability measures. Column (vi) uses $Z = \text{NetIncome}/\text{NII}$ and delivers a coefficient of -0.64, again statistically significant at the 1 percent level. Column (vii) uses $Z = \text{NoninterestExp.}/\text{NII}$ and gives a coefficient of $+1.23$. The sign flip is consistent with the model's mechanism: a higher burden of operating costs relative to core income is a form of profitability pressure, so the predicted response is of the opposite sign.

In Appendix I we extend the analysis along four directions. First, we trace out the dynamic response of the maturity gap with horizon-by-horizon local projections; the coefficients on profitability are negative on impact and remain so at longer horizons, which allow for a growing fraction of assets to be renewed—a decline in profitability thus predicts a persistent widening of the maturity gap. Second, we decompose the maturity response into asset and liability sides of the balance sheet, and into securities and loans within the asset side; the entire response is concentrated on the asset side, with securities maturity responding more strongly than loan maturity. Third, we compute interest income responses, showing that banks with low profitability subsequently generate additional interest income on securities and loans, and raise their net interest margin. Fourth, we interact the profitability measure with two proxies for term premia—an estimated term premium series from Adrian, Crump, and Moench

(2013) and the Federal Reserve's securities holdings relative to GDP, an indicator of the stance of QE. Consistent with the model's prediction, the reach-for-duration effect is amplified when estimated term premia are low or the Federal Reserve's holdings of securities are high.

Taken together, the bank-level evidence supports the model's mechanism along its three central predictions: low profitability predicts subsequent maturity extension; the extension operates through the asset side and generates additional interest income; and the response is amplified when term premia are compressed—whether measured directly or through the stance of quantitative easing.

7 Conclusion

This paper documents a pronounced increase in banks' maturity mismatch over the past four decades and proposes an explanation for it. Using historical data on U.S. commercial bank balance sheets, we show that the maturity gap between assets and liabilities has widened by roughly 4 years since the mid-1980s, tracking the secular decline in interest rates and term premia. We develop a model in which banks face profitability constraints from noninterest expenses that must be covered by current earnings. When term premia fall below a critical threshold—whether through market forces or quantitative easing—banks extend the duration of their assets to generate sufficient income to maintain profitability, a mechanism we term "reaching for duration." This portfolio adjustment increases banks' vulnerability to interest rate reversals and can lead to self-fulfilling runs by uninsured depositors.

Two further results support this mechanism. In a dynamic version of the model, banks reach for duration even before the profitability constraint binds, as shareholders hedge against future states in which it does—consistent with the steady widening of the maturity gap in the years leading up to the Great Recession. And in the cross-section of banks, less profitable institutions subsequently increase their maturity gap, with the response concentrated in securities and amplified in periods of low term premia and large Federal Reserve asset holdings.

Our analysis provides a mechanism through which quantitative easing can increase financial instability. This opens up natural questions about optimal policy design. How should large-scale asset purchases be structured to achieve their intended macroeconomic stimulus objectives while limiting the buildup of duration risk in the banking sector? How could macroprudential tools complement monetary policy in addressing this trade-off? We believe that finding answers to these important questions is a fruitful avenue for future work.

References

ABADI, J., M. BRUNNERMEIER, AND Y. KOBY (2023): "The Reversal Interest Rate," *American Economic Review*, 113, 2084–2120.

ADRIAN, T., R. K. CRUMP, AND E. MOENCH (2013): "Pricing the Term Structure with Linear Regressions," *Journal of Financial Economics*, 110, 110–138.

AJELLO, A., N. BOYARCHENKO, F. GOURIO, AND A. TAMBALOTTI (2022): "Financial Stability Considerations for Monetary Policy: Theoretical Mechanisms," Staff Report 1002, Federal Reserve Bank of New York.

ALEXANDROV, A. AND M. K. BRUNNERMEIER (2026): "Optimal (Un)Conventional Monetary Policy," Working Paper, Princeton University.

ALMEIDA, H., V. FOS, AND M. KRONLUND (2016): "The Real Effects of Share Repurchases," *Journal of Financial Economics*, 119, 168–185.

BARTH, D., R. J. KAHN, P. MONIN, AND O. SOKOLINSKIY (2024): "Reaching for Duration and Leverage in the Treasury Market," Finance and Economics Discussion Series 2024-039, Board of Governors of the Federal Reserve System.

BAUER, M. D. AND G. D. RUDEBUSCH (2020): "Interest Rates under Falling Stars," *American Economic Review*, 110, 1316–1354.

BECKER, B. AND V. IVASHINA (2015): "Reaching for Yield in the Bond Market," *Journal of Finance*, 70, 1863–1902.

BEGENAU, J., M. PIAZZESI, AND M. SCHNEIDER (2025): "Banks' Risk Exposures," NBER Working Paper 21334, National Bureau of Economic Research, revised August 2025.

BEN-DAVID, I. AND A. CHINCO (2023): "Modeling Managers as EPS Maximizers," NBER Working Paper 31125, National Bureau of Economic Research.

BOLTON, P., Y. LI, N. WANG, AND J. YANG (2025): "Dynamic Banking and the Value of Deposits," *Journal of Finance*, 80, 2063–2105.

BORIO, C. AND H. ZHU (2012): "Capital Regulation, Risk-Taking and Monetary Policy: A Missing Link in the Transmission Mechanism?" *Journal of Financial Stability*, 8, 236–251.

BOYARCHENKO, N., G. FAVARA, AND M. SCHULARICK (2022): "Financial Stability Considerations for Monetary Policy: Empirical Evidence and Challenges," Staff Report 1003, Federal Reserve Bank of New York.

BRUNNERMEIER, M. K. AND L. H. PEDERSEN (2009): "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, 22, 2201–2238.

BURGSTAHLER, D. AND I. DICHEV (1997): "Earnings Management to Avoid Earnings Decreases and Losses," *Journal of Accounting and Economics*, 24, 99–126.

CAMPBELL, J. Y. AND R. SIGALOV (2022): "Portfolio Choice with Sustainable Spending: A Model of Reaching for Yield," *Journal of Financial Economics*, 143, 188–206.

CHOI, J. AND M. KRONLUND (2018): "Reaching for Yield in Corporate Bond Mutual Funds," *Review of Financial Studies*, 31, 1930–1965.

CLAESSENS, S., N. COLEMAN, AND M. DONNELLY (2018): "'Low-For-Long' Interest Rates and Banks' Interest Margins and Profitability: Cross-Country Evidence," *Journal of Financial Intermediation*, 35, 1–16.

DANIEL, K., L. GARLAPPI, AND K. XIAO (2021): "Monetary Policy and Reaching for Income," *Journal of Finance*, 76, 1145–1193.

DEGEORGE, F., J. PATEL, AND R. ZECKHAUSER (1999): "Earnings Management to Exceed Thresholds," *Journal of Business*, 72, 1–33.

DELL'ARICCIA, G., L. LAEVEN, AND R. MARQUEZ (2014): "Real Interest Rates, Leverage, and Bank Risk-Taking," *Journal of Economic Theory*, 149, 65–99.

DELL'ARICCIA, G., L. LAEVEN, AND G. A. SUAREZ (2017): "Bank Leverage and Monetary Policy's Risk-Taking Channel: Evidence from the United States," *Journal of Finance*, 72, 613–654.

DEMARZO, P. M., A. KRISHNAMURTHY, AND S. NAGEL (2024): "Interest Rate Risk in Banking," NBER Working Paper 33308, National Bureau of Economic Research.

DI MAGGIO, M. AND M. KACPERCZYK (2017): "The Unintended Consequences of the Zero Lower Bound Policy," *Journal of Financial Economics*, 123, 59–80.

DI TELLA, S. AND P. KURLAT (2021): "Why Are Banks Exposed to Monetary Policy?" *American Economic Journal: Macroeconomics*, 13, 295–340.

DIAMOND, D. W. AND P. H. DYBVIG (1983): "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy*, 91, 401–419.

DIAMOND, D. W. AND R. G. RAJAN (2012): "Illiquid Banks, Financial Stability, and Interest Rate Policy," *Journal of Political Economy*, 120, 552–591.

DOMANSKI, D., H. S. SHIN, AND V. SUSHKO (2017): "The Hunt for Duration: Not Waving but Drowning?" *IMF Economic Review*, 65, 113–153.

DRECHSLER, I., A. SAVOV, AND P. SCHNABL (2017): "The Deposits Channel of Monetary Policy," *Quarterly Journal of Economics*, 132, 1819–1876.

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

DRECHSLER, I., A. SAVOV, P. SCHNABL, AND O. WANG (2026): "Deposit Franchise Runs," *Journal of Finance*, 81, 1573–1617.

ENGLISH, W. B., S. J. VAN DEN HEUVEL, AND E. ZAKRAJŠEK (2018): "Interest Rate Risk and Bank Equity Valuations," *Journal of Monetary Economics*, 98, 80–97.

FEDERAL DEPOSIT INSURANCE CORPORATION (2025): "Quarterly Banking Profile, Fourth Quarter 2024," FDIC Quarterly, 19(1), https://www.fdic.gov/quarterly-banking-profile/quarterly-banking-profile-q4-2024.

FLANNERY, M. J. (1981): "Market Interest Rates and Commercial Bank Profitability: An Empirical Investigation," *Journal of Finance*, 36, 1085–1101.

——— (1983): "Interest Rates and Bank Profitability: Additional Evidence: Note," *Journal of Money, Credit and Banking*, 15, 355–362.

FLOYD, E., N. LI, AND D. J. SKINNER (2015): "Payout policy through the financial crisis: The growth of repurchases and the resilience of dividends," *Journal of Financial Economics*, 118, 299–316.

GOMEZ, M., A. LANDIER, D. SRAER, AND D. THESMAR (2021): "Banks' Exposure to Interest Rate Risk and the Transmission of Monetary Policy," *Journal of Monetary Economics*, 117, 543–570.

GRAHAM, J. R., C. R. HARVEY, AND S. RAJGOPAL (2005): "The Economic Implications of Corporate Financial Reporting," *Journal of Accounting and Economics*, 40, 3–73.

GÜRKAYNAK, R. S., B. SACK, AND J. H. WRIGHT (2007): "The U.S. Treasury Yield Curve: 1961 to the Present," *Journal of Monetary Economics*, 54, 2291–2304.

HADDAD, V., B. HARTMAN-GLASER, AND T. MUIR (2023): "Bank Fragility When Depositors Are the Asset," Working Paper, UCLA.

HADDAD, V. AND D. SRAER (2020): "The Banking View of Bond Risk Premia," *Journal of Finance*, 75, 2465–2502.

HANSON, S. G. AND J. C. STEIN (2015): "Monetary Policy and Long-Term Real Rates," *Journal of Financial Economics*, 115, 429–448.

HELLMANN, T. F., K. C. MURDOCK, AND J. E. STIGLITZ (2000): "Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough?" *American Economic Review*, 90, 147–165.

JIANG, E. X., G. MATVOS, T. PISKORSKI, AND A. SERU (2024): "Monetary Tightening and U.S. Bank Fragility in 2023: Mark-to-Market Losses and Uninsured Depositor Runs?" *Journal of Financial Economics*, 159, 103899.

JIMENEZ, G., S. ONGENA, J.-L. PEYDRO, AND J. SAURINA (2014): "Hazardous Times for Monetary Policy: What Do Twenty-Three Million Bank Loans Say About the Effects of Monetary Policy on Credit Risk-Taking?" *Econometrica*, 82, 463–505.

KANE, E. J. (1989): *The S&L Insurance Mess: How Did It Happen?*, Washington, DC: Urban Institute Press.

KEELEY, M. C. (1990): "Deposit Insurance, Risk, and Market Power in Banking," *American Economic Review*, 80, 1183–1200.

KIM, D. H. AND J. H. WRIGHT (2005): "An Arbitrage-Free Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yields and Distant-Horizon Forward Rates," Finance and Economics Discussion Series 2005-33, Board of Governors of the Federal Reserve System.

LIAN, C., Y. MA, AND C. WANG (2019): "Low Interest Rates and Risk-Taking: Evidence from Individual Investment Decisions," *Review of Financial Studies*, 32, 2107–2148.

MADDALONI, A. AND J.-L. PEYDRÓ (2011): "Bank Risk-Taking, Securitization, Supervision, and Low Interest Rates: Evidence from the Euro-Area and the U.S. Lending Standards," *Review of Financial Studies*, 24, 2121–2165.

MARTINEZ-MIERA, D. AND R. REPULLO (2017): "Search for Yield," *Econometrica*, 85, 351–378.

MERTON, R. C. (1971): "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," *Journal of Economic Theory*, 3, 373–413.

MYERS, S. C. AND N. S. MAJLUF (1984): "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have," *Journal of Financial Economics*, 13, 187–221.

PAUL, P. (2023): "Banks, Maturity Transformation, and Monetary Policy," *Journal of Financial Intermediation*, 53, 101011.

PENNACCHI, G. G. AND J. A. C. SANTOS (2021): "Why Do Banks Target ROE?" *Journal of Financial Stability*, 54, 100856.

STEIN, J. C. (2012): "Monetary Policy as Financial Stability Regulation," *Quarterly Journal of Economics*, 127, 57–95.

SUPERA, D. (2021): "Running Out of Time (Deposits): Falling Interest Rates and the Decline of Business Lending, Investment and Firm Creation," Working Paper, Wharton School, University of Pennsylvania.

ULATE, M. (2021): "Going Negative at the Zero Lower Bound: The Effects of Negative Nominal Interest Rates," *American Economic Review*, 111, 1–40.

VAYANOS, D. AND J.-L. VILA (2021): "A Preferred-Habitat Model of the Term Structure of Interest Rates," *Econometrica*, 89, 77–112.

WANG, O. (2025): "Banks, Low Interest Rates, and Monetary Policy Transmission," *Journal of Finance*, 80, 1379–1416.

WHITE, L. J. (1991): *The S&L Debacle: Public Policy Lessons for Bank and Thrift Regulation*, New York: Oxford University Press.

Appendix for

"Reaching for Duration"

by Thomas Mertens, Pascal Paul, and Andres Schneider

A Maturity Data Documentation

We construct the data using several FFIEC reporting forms: the FFIEC 010 (domestic banks, 1959–1983; from December 1978 restricted to those with less than $100M in total assets), the FFIEC 012 (domestic offices only, 1978–1983), the FFIEC 014 (banks with foreign offices, 1969–1983), the FFIEC 015 (large banks supplement, 1976–1983), the FFIEC 031 (banks with foreign offices, 1984–present), the FFIEC 032–034 (domestic-only banks by size, 1984–2000), and the FFIEC 041 (all domestic-only banks, 2001–present).

Table A.1: Historical Call Report Reporting Forms

Form	Filing Period	Frequency	Coverage
FFIEC 010	1959:Q2–1983:Q4	SA/Q	Domestic Report of Condition (<$100M)
FFIEC 012	1978:Q4–1983:Q4	Q	Domestic offices only
FFIEC 014	1969:Q2–1983:Q4	Q	Domestic & foreign subsidiaries
FFIEC 015	1976:Q1–1983:Q1	Q	Large banks supplement
FFIEC 031	1984:Q1–present	Q	Domestic & foreign offices
FFIEC 032	1984:Q1–2000:Q4	Q	Domestic only, assets $\geq$$300M
FFIEC 033	1984:Q1–2000:Q4	Q	Domestic only, $100M–$300M
FFIEC 034	1984:Q1–2000:Q4	Q	Domestic only, <$100M
FFIEC 041	2001:Q1–present	Q	Domestic only (replaces 032–034)

Notes: This table lists the FFIEC reporting forms used to construct the maturity data. “SA” denotes semiannual reporting (FFIEC 010 from 1964 to 1972); “Q” denotes quarterly. The data are digitally available at the Federal Reserve. Historical reporting forms can be accessed via the FFIEC website and the Micro Data Reference Manual at the Federal Reserve.

Reporting Form Details

The following provides detailed descriptions of the FFIEC reporting forms used to construct the maturity data.

- **FFIEC 010** — Domestic Report of Condition. Filed from June 10, 1959 to December 31, 1983. Prior to December 31, 1960, it was reported by FRS member banks and a handful of nonmember banks; afterwards, by all insured commercial banks, savings banks, nondeposit trust companies, and industrial banks with domestic offices only. From December 31, 1978, coverage was restricted to banks with domestic offices only and with less than $100 million in total assets. From 1964 to 1972, data are reported semiannually; otherwise, quarterly.

- **FFIEC 012** — Consolidated Report of Condition for a Bank and its Domestic Subsidiaries (Domestic Offices Only). Filed from December 31, 1978 to December 31, 1983, by all insured commercial banks, savings banks, nondeposit trust companies, and industrial banks with domestic offices only.

- **FFIEC 014** — Consolidated Report of Condition for a Bank and its Domestic and Foreign Subsidiaries. Filed from June 30, 1969 to December 31, 1983, by all insured commercial banks with domestic offices and foreign branches, foreign subsidiaries, Edge Act and agreement subsidiaries, and/or branches in Puerto Rico or U.S. territories and possessions.

- **FFIEC 015** — Large Banks Supplement. Filed from March 31, 1976 to March 31, 1983. Provided additional detail on maturity breakdowns for large banks, including time deposit maturities on the liability side starting in 1976:Q1.

- **FFIEC 031** — Consolidated Reports of Condition and Income for a Bank with Domestic and Foreign Offices. Filed from March 31, 1984 to the present by all national banks, state member banks, and insured state nonmember banks with branches or consolidated subsidiaries in a foreign country or with international banking facilities; savings associations joined the Call Report panel in March 2012, when the Thrift Financial Report was discontinued (see Section 2).

- **FFIEC 032–034** — Consolidated Reports of Condition and Income for banks with domestic offices only, stratified by asset size. The FFIEC 032 covered banks with total assets of $300 million or more; the FFIEC 033 covered those with total assets of $100 million or more but less than $300 million; the FFIEC 034 covered those with total assets less than $100 million. All three were filed from March 31, 1984 to December 31, 2000.

- **FFIEC 041** — Consolidated Reports of Condition and Income for a Bank with Domestic Offices Only. Since March 31, 2001, the FFIEC 041 replaced the FFIEC 032, 033, and 034 forms. It is filed by all national banks, state member banks, and insured state nonmember banks with domestic offices only, and—since March 2012—savings associations.

Table A.2: Asset Maturity Variables: 1959:Q2–1965:Q2

Variable	Code	Maturity Bucket	Weight (Mo.)	Period
Treasury Notes				
	RCON 0294	<12 months	6	1959:Q4–1965:Q2
	RCON 0304	>12 months	24	1959:Q4–1965:Q2
Treasury Bonds (1959:Q2–1960:Q4)				
	RCON 0309	<60 months	30	1959:Q2–1960:Q4
	RCON 0315	60–120 months	90	1959:Q2–1960:Q4
	RCON 0324	120–240 months	180	1959:Q2–1960:Q4
	RCON 0328	>240 months	300	1959:Q2–1960:Q4
Treasury Bonds (1961:Q1–1965:Q2)				
	RCON 0298	<12 months	6	1961:Q1–1965:Q2
	RCON 0308	12–60 months	36	1961:Q1–1965:Q2
	RCON 0315	60–120 months	90	1961:Q1–1965:Q2
	RCON 0320	>120 months	150	1961:Q1–1965:Q2
Cash				
	RCON 0010	0	0	1959:Q4–1965:Q2

Notes: This table lists the Call Report variables used to measure asset maturity from 1959:Q2 to 1965:Q2. Data are from the FFIEC 010 reporting forms. "Weight" denotes the representative maturity (in months) assigned to each bucket.

Table A.3: Asset Maturity Variables: 1974:Q2–1983:Q1

Variable	Code	Maturity Bucket	Weight (Mo.)	Large Only
U.S. Treasuries				
	RCON 0405	<12 months	6	
	RCON 0410	12–60 months	36	
	RCON 0415	60–120 months	90	
	RCON 0420	>120 months	150	
U.S. Agency Obligations				
	RCON 0610	<12 months	6	
	RCON 0615	12–60 months	36	
	RCON 0620	60–120 months	90	
	RCON 0625	>120 months	150	
Municipal Bonds				
	RCON 0910	<12 months	6	
	RCON 0915	12–60 months	36	
	RCON 0920	60–120 months	90	
	RCON 0925	>120 months	150	
Other Bonds, Notes, and Debentures				
	RCON 0960	<12 months	6	
	RCON 0965	12–60 months	36	
	RCON 0970	60–120 months	90	
	RCON 0975	>120 months	150	
Construction & Development Loans				
	RCOS 1416	<12 months	6	✓
	RCOS 1417	12–60 months	36	✓
	RCOS 1418	>60 months	120	✓
Other RE Loans				
	RCOS 1486	<12 months	6	✓
	RCOS 1487	12–60 months	36	✓
	RCOS 1488	>60 months	120	✓
C&I Loans				
	RCOS 1601	<12 months	6	✓
	RCOS 1602	12–60 months	36	✓
	RCOS 1603	>60 months	120	✓
Other Loans				
	RCOS 2086	<12 months	6	✓
	RCOS 2087	12–60 months	36	✓
	RCOS 2088	>60 months	120	✓
Cash, Fed Funds, Securities to Resell				
	RCON 0010	0	0	
	RCON/RCFD 1350	0	0	

Notes: This table lists the Call Report variables used to measure asset maturity from 1974:Q2 to 1983:Q1. Data are from FFIEC 010, 012, 014, and 015 reporting forms. Variables marked "Large Only" are from the FFIEC 015 supplement and are available only for large banks. Loan variables use the RCOS prefix. The FFIEC 015 supplement begins in 1976:Q1; because the sample for this regime is restricted to FFIEC 015 filers, the series constructed from these variables begin in 1976:Q1 (see also Table A.7).

Table A.4: Asset Maturity/Repricing Variables: 1983:Q2–1987:Q4

Variable	Code	Maturity/Repr. Bucket	Weight (Mo.)
Loans and Leases			
	RCON/RCFD 1425	0 (noninterest)	0
	RCON/RCFD 1426	<3 months	1.5
	RCON/RCFD 1427	3–6 months	4.5
	RCON/RCFD 1428	6–12 months	9
	RCON/RCFD 1429	12–60 months	36
	RCON/RCFD 1431	>60 months	165
Debt Securities			
	RCON/RCFD 1546	0 (noninterest)	0
	RCON/RCFD 1547	<3 months	1.5
	RCON/RCFD 1548	3–6 months	4.5
	RCON/RCFD 1549	6–12 months	9
	RCON/RCFD 1551	12–60 months	36
	RCON/RCFD 1552	>60 months	130
Other Interest-Bearing Assets			
	RCON/RCFD 1554	0 (noninterest)	0
	RCON/RCFD 1556	<3 months	1.5
	RCON/RCFD 1557	3–6 months	4.5
	RCON/RCFD 1558	6–12 months	9
	RCON/RCFD 1559	12–60 months	36
	RCON/RCFD 1561	>60 months	130
Cash, Fed Funds, Securities to Resell			
	RCON/RCFD 0081 + 0071 or 0010	0	0
	RCON/RCFD 1350	0	0

Notes: This table lists the Call Report variables used to measure asset maturity from 1983:Q2 to 1987:Q4. Data for 1983:Q2–1983:Q4 are from the FFIEC 010, 012, and 014 forms, which introduced the revised maturity/repricing schedule, and from the FFIEC 031–034 forms thereafter. Buckets represent maturity for fixed-rate positions and earliest repricing date for floating-rate positions.

Table A.5: Asset Maturity/Repricing Variables: 1988:Q1–1997:Q1

Variable	Code	Maturity/Repr. Bucket	Weight (Mo.)
Fixed-Rate Debt Securities (Remaining Maturity)			
	RCON/RCFD 0343	<3 months	1.5
	RCON/RCFD 0344	3–12 months	7.5
	RCON/RCFD 0345	12–60 months	36
	RCON/RCFD 0346	>60 months	130
Floating-Rate Debt Securities (Repricing Frequency)			
	RCON/RCFD 4544	<3 months	1.5
	RCON/RCFD 4545	3–12 months	7.5
	RCON/RCFD 4551	12–60 months	36
	RCON/RCFD 4552	>60 months	130
Fixed-Rate Loans (Remaining Maturity)			
	RCON/RCFD 0348	<3 months	1.5
	RCON/RCFD 0349	3–12 months	7.5
	RCON/RCFD 0356	12–60 months	36
	RCON/RCFD 0357	>60 months	165
Floating-Rate Loans (Repricing Frequency)			
	RCON/RCFD 4554	<3 months	1.5
	RCON/RCFD 4555	3–12 months	7.5
	RCON/RCFD 4561	12–60 months	36
	RCON/RCFD 4564	>60 months	165
Cash, Fed Funds, Securities to Resell			
	RCON/RCFD 0081 + RCON/RCFD 0071	0	0
	RCON/RCFD 1350	0	0

Notes: This table lists the Call Report variables used to measure asset maturity from 1988:Q1 to 1997:Q1. Data are from FFIEC 031–034 reporting forms. Fixed-rate positions are classified by remaining maturity; floating-rate positions by repricing frequency.

Table A.6: Asset Maturity/Repricing Variables: 1997:Q2–Present

Variable	Code	Maturity/Repr. Bucket	Weight (Mo.)	Period
Debt Securities—Government Agencies				
	RCON/RCFD A549	<3 months	1.5	1997:Q2-present
	RCON/RCFD A550	3–12 months	7.5	1997:Q2-present
	RCON/RCFD A551	12–36 months	24	1997:Q2-present
	RCON/RCFD A552	36–60 months	48	1997:Q2-present
	RCON/RCFD A553	60–180 months	120	1997:Q2-present
	RCON/RCFD A554	>180 months	240	1997:Q2-present
Debt Securities—Mortgage Pass-Throughs				
	RCON/RCFD A555	<3 months	1.5	1997:Q2-present
	RCON/RCFD A556	3–12 months	7.5	1997:Q2-present
	RCON/RCFD A557	12–36 months	24	1997:Q2-present
	RCON/RCFD A558	36–60 months	48	1997:Q2-present
	RCON/RCFD A559	60–180 months	120	1997:Q2-present
	RCON/RCFD A560	>180 months	240	1997:Q2-present
Debt Securities—Other MBS				
	RCON/RCFD A561	<36 months	18	1997:Q2-present
	RCON/RCFD A562	>36 months	60	1997:Q2-present
Loans and Leases—Closed-End				
	RCON A564	<3 months	1.5	1997:Q2-present
	RCON A565	3–12 months	7.5	1997:Q2-present
	RCON A566	12–36 months	24	1997:Q2-present
	RCON A567	36–60 months	48	1997:Q2-present
	RCON A568	60–180 months	120	1997:Q2-present
	RCON A569	>180 months	240	1997:Q2-present
Loans and Leases—Excluding Closed-End				
	RCON/RCFD A570	<3 months	1.5	1997:Q2-present
	RCON/RCFD A571	3–12 months	7.5	1997:Q2-present
	RCON/RCFD A572	12–36 months	24	1997:Q2-present
	RCON/RCFD A573	36–60 months	48	1997:Q2-present
	RCON/RCFD A574	60–180 months	120	1997:Q2-present
	RCON/RCFD A575	>180 months	240	1997:Q2-present
Cash, Fed Funds, Securities to Resell				
	RCON/RCFD 0081 + RCON/RCFD 0071	0	0	1997:Q2-present
	RCON/RCFD 1350	0	0	1997:Q2-2001:Q4
	RCON B987 + RCON/RCFD B989	0	0	2002:Q1-present

Notes: This table lists the Call Report variables used to measure asset maturity from 1997:Q2 to the present. Data are from FFIEC 031–034 and FFIEC 041 reporting forms. Fixed-rate positions are classified by remaining maturity; floating-rate positions by repricing frequency.

Table A.7: Liability Maturity Variables: 1976:Q1–1983:Q1

Variable	Code	Maturity/Repr. Bucket	Weight (Mo.)	Large Only
Time CDs ≥$100k				
	RCOS 6637	<3 months	1.5	✓
	RCOS 6639	3–6 months	4.5	✓
	RCOS 6641	6–12 months	9	✓
	RCOS 6643	>12 months	36	✓
Zero-Maturity Liabilities				
	RCON 2220	Demand deposits	0	
	RCON 6650	Savings deposits	0	

Notes: This table lists the Call Report variables used to measure liability maturity from 1976:Q1 to 1983:Q1. Time deposit data are from the FFIEC 015 large-bank supplement (RCOS prefix).

Table A.8: Liability Maturity/Repricing Variables: 1983:Q2–1987:Q4

Variable	Code	Maturity/Repr. Bucket	Weight (Mo.)
Time Deposits ≥$100k			
	RCON 2459	0 (noninterest)	0
	RCON 2460	<3 months	1.5
	RCON 2461	3–6 months	4.5
	RCON 2462	6–12 months	9
	RCON 2464	12–60 months	36
	RCON 2465	>60 months	90
All Other Time Deposits			
	RCON 2466	0 (noninterest)	0
	RCON 2467	<3 months	1.5
	RCON 2470	3–6 months	4.5
	RCON 2471	6–12 months	9
	RCON 2472	12–60 months	36
	RCON 2473	>60 months	90
Nondeposit Interest-Bearing Liabilities			
	RCON/RCFD 2492	0 (noninterest)	0
	RCON/RCFD 2493	<3 months	1.5
	RCON/RCFD 2495	3–6 months	4.5
	RCON/RCFD 2497	6–12 months	9
	RCON/RCFD 2499	12–60 months	36
	RCON/RCFD 2501	>60 months	90
Subordinated Notes & Debentures (1984:Q1–1985:Q1)			
	RCON/RCFD 3201	<12 months	6
	RCON/RCFD 3203	12–24 months	18
	RCON/RCFD 3204	24–36 months	30
	RCON/RCFD 3206	36–48 months	42
	RCON/RCFD 3208	48–60 months	54
	RCON/RCFD 3209	>60 months	90
Limited-Life Preferred Stock (1984:Q1–1985:Q1)			
	RCON/RCFD 3276	<12 months	6
	RCON/RCFD 3277	12–24 months	18
	RCON/RCFD 3278	24–36 months	30
	RCON/RCFD 3279	36–48 months	42
	RCON/RCFD 3280	48–60 months	54
	RCON/RCFD 3281	>60 months	90
Zero-Maturity Liabilities			
	RCON 2215 / 2220	Demand & transaction deposits	0
	RCON 2389 / 6650	Nontransaction savings deposits	0

Notes: This table lists the Call Report variables used to measure liability maturity from 1983:Q2 to 1987:Q4. Data for 1983:Q2–1983:Q4 are from the FFIEC 010, 012, and 014 forms, which introduced the revised maturity/repricing schedule, and from the FFIEC 031–034 forms thereafter.

Table A.9: Liability Maturity/Repricing Variables: 1988:Q1–1997:Q1

Variable	Code	Maturity/Repr. Bucket	Weight (Mo.)
Time Deposits <\$100k & Open-Account TDs ≥\$100k (1988:Q1–1995:Q4)			
	RCON 0359	<3 months	1.5
	RCON 3644	3–12 months	7.5
Fixed-Rate Time Deposits ≥\$100k (1988:Q1–1995:Q4)			
	RCON 2761	<3 months	1.5
	RCON 2762	3–12 months	7.5
	RCON 2763	12–60 months	36
	RCON 2765	>60 months	90
Floating-Rate Time Deposits ≥\$100k (1988:Q1–1995:Q4)			
	RCON 4568	<3 months	1.5
	RCON 4569	3–12 months	7.5
	RCON 4571	12–60 months	36
	RCON 4572	>60 months	90
Fixed-Rate Time Deposits <\$100k (1996:Q1–1997:Q1)			
	RCON A225	<3 months	1.5
	RCON A226	3–12 months	7.5
	RCON A227	>12 months	16
Floating-Rate Time Deposits <\$100k (1996:Q1–1997:Q1)			
	RCON A228	<3 months	1.5
	RCON A229	3–12 months	7.5
	RCON A230	>12 months	16
Fixed-Rate Time Deposits ≥\$100k (1996:Q1–1997:Q1)			
	RCON A232	<3 months	1.5
	RCON A233	3–12 months	7.5
	RCON A234	12–60 months	36
	RCON A235	>60 months	90
Floating-Rate Time Deposits ≥\$100k (1996:Q1–1997:Q1)			
	RCON A236	<3 months	1.5
	RCON A237	3–12 months	7.5
	RCON A238	12–60 months	36
	RCON A239	>60 months	90
Zero-Maturity Liabilities			
	RCON 2215	Demand & transaction deposits	0
	RCON 2389	Nontransaction savings deposits	0

Notes: This table lists the Call Report variables used to measure liability maturity from 1988:Q1 to 1997:Q1, covering two sub-regimes: 1988:Q1–1995:Q4 and 1996:Q1–1997:Q1. Data are from FFIEC 031–034 reporting forms. RCON A231 and RCON A240 report floating-rate time deposits with remaining maturity <12 months (1996:Q1–1997:Q1).

Table A.10: Liability Maturity/Repricing Variables: 1997:Q2–Present

Variable	Code	Maturity/Repr. Bucket	Weight (Mo.)	Period
Time Deposits <$100k (1997:Q2–2016:Q4)				
	RCON A579	<3 months	1.5	1997:Q2–2016:Q4
	RCON A580	3–12 months	7.5	1997:Q2–2016:Q4
	RCON A581	12–36 months	24	1997:Q2–2016:Q4
	RCON A582	>36 months	60	1997:Q2–2016:Q4
Time Deposits ≥$100k (1997:Q2–2016:Q4)				
	RCON A584	<3 months	1.5	1997:Q2–2016:Q4
	RCON A585	3–12 months	7.5	1997:Q2–2016:Q4
	RCON A586	12–36 months	24	1997:Q2–2016:Q4
	RCON A587	>36 months	60	1997:Q2–2016:Q4
Time Deposits <$250k (since 2017:Q1)				
	RCON HK07	<3 months	1.5	2017:Q1–present
	RCON HK08	3–12 months	7.5	2017:Q1–present
	RCON HK09	12–36 months	24	2017:Q1–present
	RCON HK10	>36 months	60	2017:Q1–present
Time Deposits ≥$250k (since 2017:Q1)				
	RCON HK12	<3 months	1.5	2017:Q1–present
	RCON HK13	3–12 months	7.5	2017:Q1–present
	RCON HK14	12–36 months	24	2017:Q1–present
	RCON HK15	>36 months	60	2017:Q1–present
Zero-Maturity Liabilities				
	RCON 2215	Demand & transaction deposits	0	
	RCON 2389	Nontransaction savings deposits	0	

Notes: This table lists the Call Report variables used to measure liability maturity from 1997:Q2 to the present. Data are from FFIEC 031–034 and FFIEC 041 reporting forms. In 2017:Q1, the size threshold for large time deposits changed from $100,000 to $250,000.

B Additional Evidence on Historical Maturity Data

The baseline maturity gap in Section 2 weights each Call Report category by its bucket midpoint—the representative remaining maturity for fixed-rate items, or the earliest repricing date for floating-rate items. This measure is directly comparable across reporting regimes, but it does not capture two features of interest-rate-sensitive instruments that are important for translating maturity into price sensitivity: the coupon-driven cash-flow profile of fixed-rate assets, and the prepayment behavior of mortgage-backed securities and residential mortgages. In this appendix, we present two adjustments that move the measure closer to a duration measure and assess the robustness of the patterns documented in Section 2.

Duration adjustment. The first adjustment replaces each bucket midpoint M (in months) with the Macaulay duration of a par bond at the bucket-matched yield. For a par bond with maturity M months and yield y percent, the Macaulay duration in months is

$$D(M, y) \;=\; 12 \cdot \frac{1 + y/100}{y/100} \cdot \left[1 - (1 + y/100)^{-M/12}\right].$$

We use the 3-month Treasury bill rate (FRED series DTB3) for sub-1-year buckets and the Gürkaynak, Sack, and Wright (2007) zero-coupon yield—as an approximation to the corresponding par yield—for buckets of one year or longer, linearly interpolating between adjacent integer maturities for non-integer buckets. When the relevant yield is missing or below one basis point, the weight falls back to the bucket midpoint M. The resulting weight $D(M, y_t)$ is time-varying because y_t varies over time. We then aggregate the maturity gap exactly as in the main text, but with these duration weights in place of the bucket midpoints. In reporting regimes where fixed-rate and floating-rate positions are reported separately (1988:Q1–1997:Q1), we apply the conversion to fixed-rate positions only, since the buckets of floating-rate positions reflect repricing dates rather than remaining maturities. In regimes where the two are combined within the same bucket, we apply the conversion uniformly; the resulting bias is small because floating-rate positions are concentrated in the shortest buckets, for which $D(M, y) \approx M$.

Two features of this adjustment are worth highlighting. First, $D(M, y) \leq M$ for $y > 0$, with the gap widening at higher yields and longer maturities. Put differently, the duration weight equals the bucket midpoint when yields are zero and falls below it when yields rise. Second, the downward shift in the maturity gap is therefore larger at high interest rates as the adjustment compresses the early-sample (high-rate) observations more than the recent (low-rate) ones.

Prepayment adjustment. The second adjustment further accounts for the negative convexity of mortgage-backed securities and residential mortgages. As current mortgage rates fall below their recent historical level, borrowers prepay existing loans at faster rates, shortening the effective duration of mortgage cash flows; as rates rise, prepayments slow and effective duration extends. We capture this dynamic with a time-varying scaling factor applied only to the mortgage-related codes—namely mortgage pass-through MBS (RCON A555–A560), other mortgage-backed securities (A561–A562), and closed-end 1-4 family residential mortgages (A564–A569), when these distinctions become available in 1997:Q2.

For each quarter t, we compute a refinancing-moneyness measure $m_t = \bar{r}_t^{30} - r_t^{30}$, where r_t^{30} is the contemporaneous 30-year fixed mortgage rate from Freddie Mac (FRED series MORTGAGE30US) and $\bar{r}_t^{30}$ is its trailing five-year average. Positive moneyness corresponds to current rates below recent history, when refinancing incentives are strong; negative moneyness corresponds to the reverse. The prepayment scaling factors are linear in moneyness around base levels,

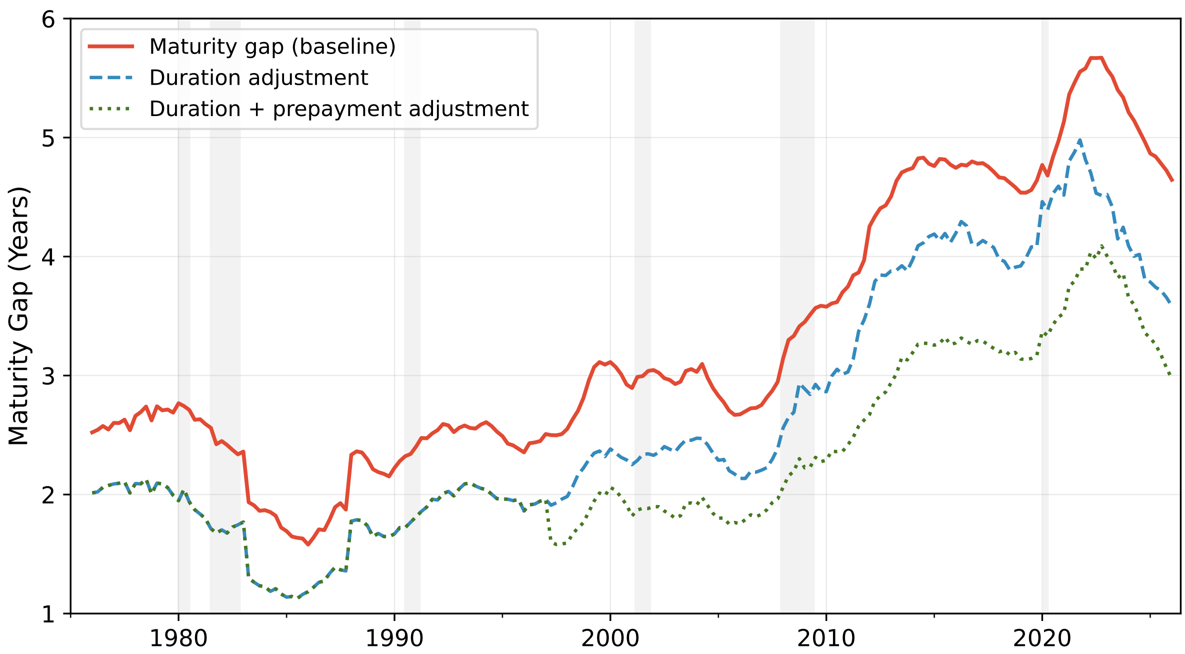
$$\kappa_t^{\text{mbs}} = 0.40 - 0.05 \cdot m_t, \qquad \kappa_t^{\text{mortgage}} = 0.50 - 0.05 \cdot m_t,$$

clamped to $[0.20, 0.60]$ and $[0.30, 0.70]$, respectively. The base levels are calibrated to reproduce the through-the-cycle ratio of effective to stated duration for agency MBS, roughly five years against twelve stated, and for bank-held 1-4 family residential mortgages, roughly six years against twelve stated. The sensitivity of 0.05 per 100 basis points of moneyness reflects the observed responsiveness of MBS effective duration to refinancing incentives over the 2008–2024 period, and the clamping bounds span the two-to-seven-year range of effective duration observed for 30-year MBS and mortgages over the same period.

For each mortgage-related code, the weight that enters the dollar-weighted average is $\kappa_t^{\text{mbs}} \cdot M$ or $\kappa_t^{\text{mortgage}} \cdot M$, where M is the stated bucket midpoint; the Macaulay conversion is *not* applied to these codes, because their effective duration is dominated by prepayment rather than by the coupon effect that the Macaulay formula captures, and the scaling factors are calibrated directly as the ratio of effective to stated duration. All other codes carry their Macaulay duration weights $D(M, y_t)$ as in the duration-adjusted series and are unaffected by the prepayment scaling.

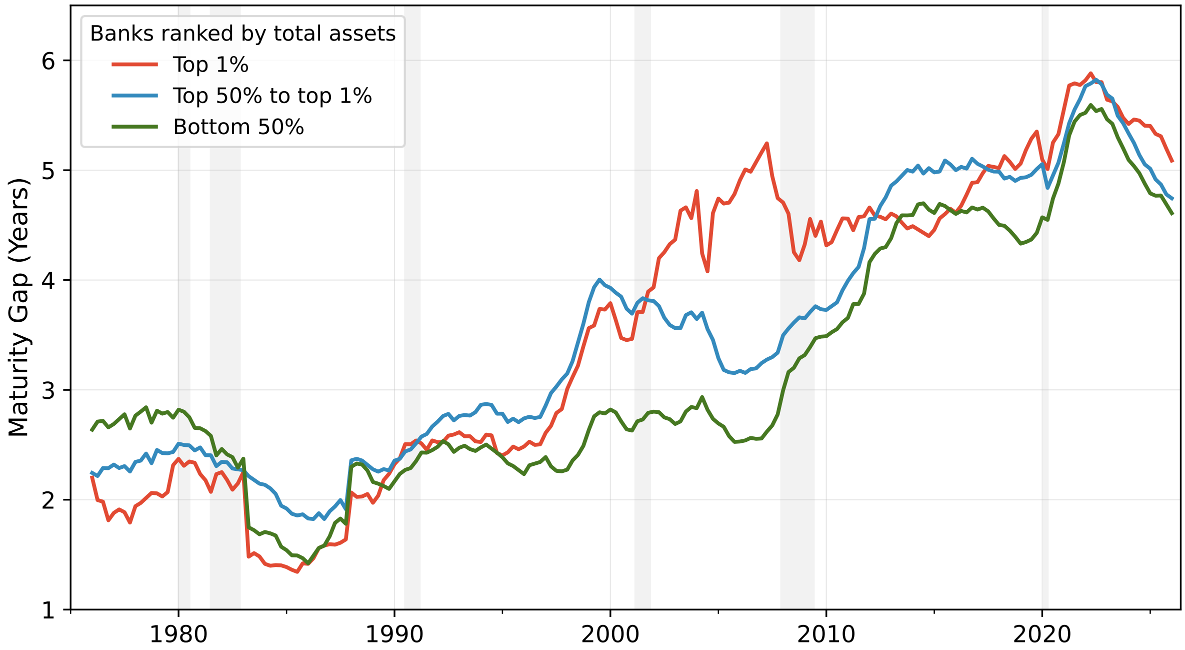
Results. Figure 10 plots the baseline maturity gap alongside the two adjusted series. Two features stand out. First, both adjusted series lie below the baseline throughout the sample, with the duration adjustment shifting the level down by roughly one year in recent quarters and the prepayment adjustment shifting it down by an additional half to one year. Second, the time path of all three series is qualitatively similar: each rises from a trough in 1985–1986 to a peak around 2021–2022 and declines in the recent tightening cycle. The trough-to-peak increases are roughly 4.1 years for the baseline (from 1.6 to 5.7 years), 3.9 years for the duration-adjusted series (from 1.1 to 5.0), and 3.0 years for the duration-plus-prepayment series (from 1.1 to 4.1). The corresponding percentage increases are around 260, 340, and 260 percent, respectively. The secular widening of the maturity gap documented in Section 2 is therefore robust to both adjustments, and the duration adjustment in fact amplifies the percentage increase because it compresses the high-rate observations of the early sample more than the low-rate ones of the recent period.

Figure 10: Maturity Gap: Baseline, Duration-, and Prepayment-Adjusted



Notes: This figure plots the baseline maturity gap and two adjusted versions. The duration-adjusted series replaces each bucket midpoint with the Macaulay duration of a par bond at the bucket-matched yield, using the 3-month Treasury bill rate for sub-1-year buckets and the Gürkaynak, Sack, and Wright (2007) zero-coupon yield for buckets of one year or longer. For the duration- and prepayment-adjusted series, the Macaulay weight of mortgage pass-through securities, other mortgage-backed securities, and closed-end 1-4 family residential mortgages is replaced by the stated bucket midpoint scaled by a time-varying factor κ_t^{mbs} or $\kappa_t^{\text{mortgage}}$ that depends on the deviation of the current 30-year fixed mortgage rate from its trailing five-year average. Shaded areas denote NBER recession dates.

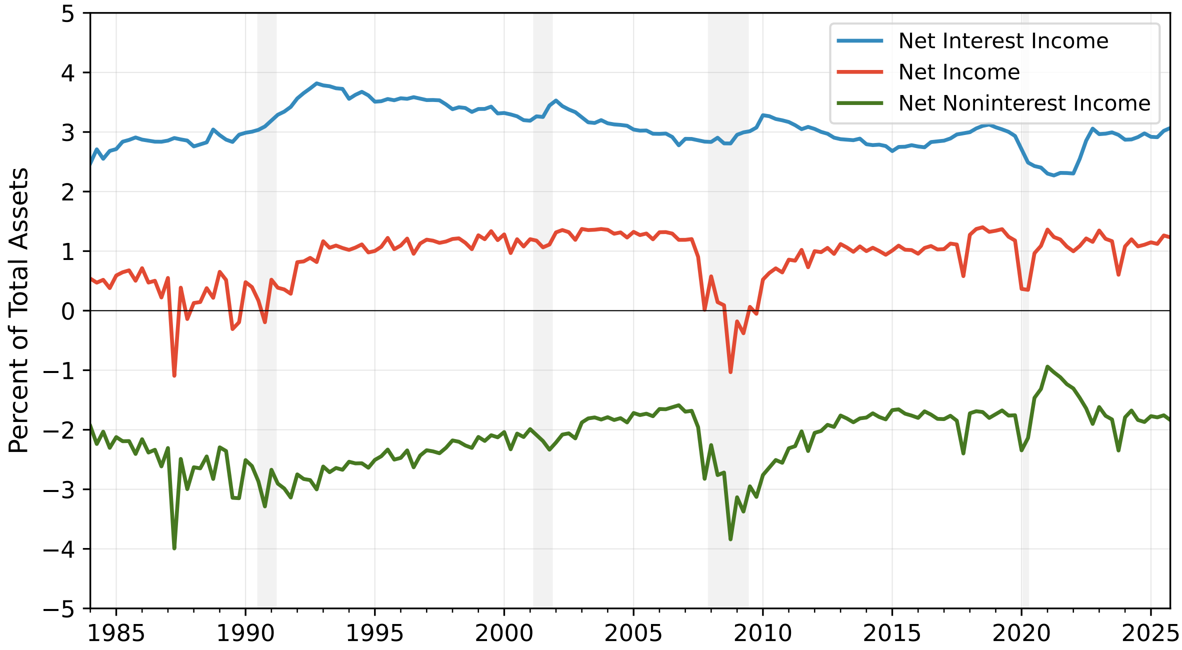
Figure 11: Maturity Gap by Bank Asset Size



Notes: This figure plots the maturity gap, defined as the difference between the average maturity of total assets and total liabilities (in years), for three buckets defined according to bank asset size (top 1%, top 50% to top 1%, bottom 50%). The maturity gap is computed as a value-weighted measure across all banks within a bucket at every point in time. Shaded areas denote NBER recession dates.

C Historical Bank Income Measures

Figure 12: Aggregate Bank Income Measures



Notes: Based on aggregate data for all FDIC-insured institutions, this figure plots net interest income (interest income minus interest expenses), net income, and net noninterest income (noninterest income minus noninterest expenses plus all other net income), all scaled by total assets. Shaded areas denote NBER recession dates.

D Endogenous Deposits

This section provides the derivations for the endogenous-deposit analysis summarized in Section 3.6. At $\bar{\alpha} = 1$, every expression below collapses to its Section 3 counterpart.

D.1 Environment

The setting is that of Section 3 with two choice variables: total assets $A \in [e, \bar{\alpha}]$ and the long-bond share $\alpha \in [0, 1]$. Deposits are $d \equiv A - e$, the dollar long-bond position is $L \equiv \alpha A$, and the long bond yields $r_\ell = r_{s,0} + \tau$ per period. The bank pays the per-period operating cost $c > 0$ of the baseline, independent of d. Net income in period 1 is

$$\mathrm{NI}_1 \;=\; A\big(r_{s,0} + \alpha\tau\big) - (A - e)\,\beta r_{s,0} - c \;=\; e\,r_{s,0} + (1 - \beta) r_{s,0}\,d + \tau L - c. \tag{28}$$

Wholesale borrowing, as in the portfolio-restriction discussion of Section 3.6, is available but strictly dominated within the capacity $\bar{\alpha}$: a wholesale dollar pays $r_{s,0}$ and reprices fully, while a deposit dollar pays $\beta r_{s,0}$ and reprices only with fraction β, at no additional cost. We therefore omit it. Repeating the steps of the baseline derivation at scale A and applying the small-rate approximation, the terminal payout net of initial equity and principal payments is

$$\Pi_2 \;=\; 2\,\mathrm{NI}_1 + b\,\varepsilon, \qquad b \;=\; (1-\alpha)A - \beta(A-e) \;=\; e + (1-\beta)d - L, \tag{29}$$

which generalizes equation (2): at $A = 1$, the exposure reduces to $b = \lambda - \alpha$. The hedging portfolio is $L^{\text{hedge}}(d) = e + (1-\beta)d$. One object recurs throughout: the *non-repricing funding base* at capacity,

$$\lambda_{\bar{\alpha}} \;\equiv\; e + (1-\beta)(\bar{\alpha} - e), \tag{30}$$

that is, equity plus the sticky fraction of deposits, with associated non-repricing income $\lambda_{\bar{\alpha}}\, r_{s,0}$, which generalizes the baseline's $\lambda r_{s,0}$; at $\bar{\alpha} = 1$, $\lambda_{\bar{\alpha}} = \lambda$. With the mean-variance preferences (4) and $\gamma \equiv \rho\sigma^2$ as in Section 3.3, the bank maximizes $U = \mathrm{NI}_1 - (\gamma/2)b^2$, subject to the profitability constraint $\mathrm{NI}_1 \geq \kappa$ where applicable.

D.2 Unconstrained Problem

Proposition 3 (Unconstrained Scale and Portfolio Choice)**.** *Suppose the profitability constraint is slack and $\bar{\alpha} \geq e + \tau/(\gamma\beta)$, so that the portfolio is interior.*[26] *Then*

$$A^* = \bar{\alpha}, \qquad b^* = -\frac{\tau}{\gamma}, \qquad L^* = \lambda_{\bar{\alpha}} + \frac{\tau}{\gamma}, \qquad \alpha^* = \alpha^{\text{hedge}}(\bar{\alpha}) + \frac{\tau}{\gamma\bar{\alpha}}, \tag{31}$$

with $\alpha^{\text{hedge}}(\bar{\alpha}) = \lambda_{\bar{\alpha}}/\bar{\alpha}$. The comparative statics of the share are

$$\frac{\partial\alpha^*}{\partial\tau} = \frac{1}{\gamma\bar{\alpha}} > 0, \qquad\qquad \frac{\partial\alpha^*}{\partial r_{s,0}} = 0. \tag{32}$$

At $\bar{\alpha} = 1$, we have $\alpha^* = \lambda + \tau/\gamma$, the interior optimum in equation (6).

Proof. The first-order condition in L gives $b^* = -\tau/\gamma$. The marginal value of a deposit dollar, holding exposure fixed, is obtained by splitting it β into the short asset and $1-\beta$ into the long bond: it earns $\beta r_{s,0} + (1-\beta)r_\ell - \beta r_{s,0} = (1-\beta)r_\ell$, which is strictly positive in every state. The bank therefore expands to its capacity regardless of $r_{s,0}$ and τ. The dollar position follows from the definition of b, interiority $L^* \leq \bar{\alpha}$ rearranges to $\tau/\gamma \leq \beta(\bar{\alpha} - e)$, and the comparative statics follow because $A^* = \bar{\alpha}$ is constant: the

[26] For a tighter capacity, the corner $\alpha^* = 1$ obtains with $A^* = \bar{\alpha}$: on the face $L = A$, the objective is increasing in A throughout, since deposits add income $(1-\beta)r_{s,0} + \tau$ per unit at exposure cost β.

myopic component is the only moving part, so the share rises one-for-one with $\tau/(\gamma\bar{\alpha})$ and does not respond to the short rate. Income at the optimum is

$$\mathrm{NI}_1^u \;=\; \lambda_{\bar{\alpha}}\, r_\ell + \frac{\tau^2}{\gamma} - c, \tag{33}$$

the long rate earned on the non-repricing funding base, plus myopic component income, net of operating costs. $\square$

D.3 Constrained Problem

The constraint binds if and only if $\kappa > \mathrm{NI}_1^u$, that is, $\kappa + c > \lambda_{\bar{\alpha}} r_\ell + \tau^2/\gamma$. Solving this condition with equality for τ gives the threshold

$$\tilde{\tau} \;=\; \frac{1}{2}\left(-\lambda_{\bar{\alpha}}\gamma + \sqrt{\lambda_{\bar{\alpha}}^2\gamma^2 + 4\gamma\big(\kappa + c - \lambda_{\bar{\alpha}} r_{s,0}\big)}\right), \tag{34}$$

which at $\bar{\alpha} = 1$ is the threshold in equation (10). For a fixed τ, the analogous threshold in the short rate is $\tilde{r} = (\kappa + c - \tau^2/\gamma)/\lambda_{\bar{\alpha}} - \tau$.

Proposition 4 (Constrained Portfolio at Capacity)**.** *Suppose the constraint binds. Then $A^* = \bar{\alpha}$ and*

$$L_{\text{bind}} \;=\; \frac{\kappa + c - \lambda_{\bar{\alpha}} r_{s,0}}{\tau}, \qquad \alpha_{\text{bind}} \;=\; \frac{L_{\text{bind}}}{\bar{\alpha}}, \tag{35}$$

with

$$\frac{\partial L_{\text{bind}}}{\partial \tau} = -\frac{L_{\text{bind}}}{\tau} < 0, \qquad \frac{\partial^2 L_{\text{bind}}}{\partial \tau^2} = \frac{2L_{\text{bind}}}{\tau^2} > 0, \qquad \frac{\partial L_{\text{bind}}}{\partial r_{s,0}} = -\frac{\lambda_{\bar{\alpha}}}{\tau} < 0. \tag{36}$$

A feasible solution exists if and only if $L_{\text{bind}} \le \bar{\alpha}$, that is, $\tau \ge (\kappa + c - \lambda_{\bar{\alpha}} r_{s,0})/\bar{\alpha}$; below this threshold the bank is unprofitable—the unprofitability condition of the portfolio-restriction discussion in Section 3.6, with λ replaced by $\lambda_{\bar{\alpha}} \ge \lambda$. The share α_{bind} reaches one exactly at the feasibility boundary, so there is no separate leverage regime: with deposits already at capacity, full duration and unprofitability coincide, as in the baseline.

Proof. Because NI_1 is deterministic and enters the objective one-for-one, a binding constraint pins income at κ and the bank minimizes b^2 subject to $\mathrm{NI}_1 = \kappa$. The deposit margin is unambiguous: in any binding solution the bank overshoots its hedge ($b < 0$, since $L_{\text{bind}} > \lambda_{\bar{\alpha}} + \tau/\gamma > \lambda_{\bar{\alpha}}$), and a marginal hedged deposit dollar both *adds* income at rate $(1-\beta) r_{s,0} \ge 0$ and *reduces* variance, as it raises b toward zero by $(1-\beta)$. Deposits therefore stay at capacity; formally, with $\theta = |b|/\tau > 0$ the multiplier on the income constraint, the multiplier on $A \le \bar{\alpha}$ is $\nu = |b|(1-\beta) r_\ell/\tau > 0$. The portfolio then solves the binding constraint, $\lambda_{\bar{\alpha}}\, r_{s,0} + \tau L = \kappa + c$, which is (35); the comparative statics follow by differentiating, using that a binding constraint implies $\kappa + c - \lambda_{\bar{\alpha}} r_{s,0} = \tau L_{\text{bind}} > 0$. $\square$

E Reaching for Credit Risk

This section extends the baseline model of Section 3 by allowing the bank to invest in risky loans alongside long-term bonds. Loans earn a credit spread above bonds but expose the bank to credit losses. The central question is how a decline in the term premium—for example, driven by QE—affects the bank's allocation between bonds and loans.

The answer depends on the nature of the friction that limits lending. When the bank expands its loan portfolio, it may face a downward-sloping demand curve for credit, requiring it to lower lending rates and accept a smaller spread to attract additional borrowers or raise loan volumes of previous borrowers. Alternatively, the bank may be able to sustain previous spreads but must extend credit to borrowers with higher default risk, taking on additional credit risk. We model both margins and show that the strength of the loan portfolio response to a decline in the term premium depends on their relative importance. When credit risk is the dominant friction, the bank actively substitutes loans for bonds as τ falls; when diminishing spreads dominate, the loan book is approximately fixed and all adjustment flows through the bond portfolio.

E.1 Environment

The setting is identical to Section 3, except that the bank now chooses between three assets at $t = 0$: a short-term asset (share $1 - \alpha$), a long-term bond (share α_B), and a risky loan (share α_L), with

$$\alpha_B + \alpha_L = \alpha, \qquad \alpha_B, \alpha_L \geq 0. \tag{37}$$

Bonds. As in the baseline, long-term bonds yield $r_\ell = r_{s,0} + \tau$ per period and are held to maturity.

Loans. Loans yield $r_\ell + \sigma(\alpha_L)$ per period, where $\sigma(\alpha_L) \geq 0$ is the credit spread above the bond yield. The spread is decreasing in the volume of lending:

$$\sigma(\alpha_L) = \sigma_0 - \frac{\chi}{2}\alpha_L, \qquad \chi > 0. \tag{38}$$

The parameter $\sigma_0 > 0$ is the marginal spread on the first unit of lending, and χ captures the slope of the credit demand curve: the bank faces a downward-sloping demand for credit and must lower lending rates to attract additional borrowers, compressing the marginal spread. Total spread income is $S(\alpha_L) \equiv \alpha_L\, \sigma(\alpha_L) = \alpha_L(\sigma_0 - \frac{\chi}{2}\, \alpha_L)$, which is maximized at $\alpha_L = \sigma_0/\chi$. Note that $\sigma(\alpha_L)$ is the average spread, while the marginal spread income is $S'(\alpha_L) = \sigma_0 - \chi\, \alpha_L$.

Credit risk. In period 2, after the interest rate shock ε is realized, loans suffer a credit loss $\alpha_L\, \eta$, where η is a mean-zero random variable with variance σ_η^2, independent of ε. The shock η captures aggregate credit risk—defaults and recovery shortfalls—that is orthogonal to interest rate movements.

E.2 The Bank's Problem

The bank earns the term premium on all long assets and the credit spread on loans. Period net income generalizes the baseline by the spread income term,

$$\mathrm{NI}_1 = \lambda r_{s,0} + \alpha\tau + S(\alpha_L) - c, \tag{39}$$

and, following the same small-rate approximation as in Section 3, the terminal payout net of initial equity and principal payments is:

$$\Pi_2 = \underbrace{2\,\mathrm{NI}_1}_{\text{expected payout}} + \underbrace{(\lambda - \alpha)\,\varepsilon}_{\text{interest rate risk}} - \underbrace{\alpha_L\, \eta}_{\text{credit risk}}. \tag{40}$$

The first term is the deterministic component: cumulative net income over both periods, now augmented by spread income $S(\alpha_L)$. The second is interest rate risk, exactly as in the baseline: the bank's exposure to ε depends on total duration $\alpha = \alpha_B + \alpha_L$, with both bonds and loans contributing equally. The third is credit risk, which loads only on the loan portfolio.

Since ε and η are independent, the variance decomposes as:

$$\mathrm{Var}(\Pi_2) = \underbrace{(\lambda - \alpha)^2\,\sigma^2}_{\text{interest rate variance}} + \underbrace{\alpha_L^2\,\sigma_\eta^2}_{\text{credit variance}}. \tag{41}$$

The two risk factors enter additively and are controlled by different portfolio choices: interest rate variance depends on total duration α, while credit variance depends only on the loan share α_L.

The bank has the mean-variance preferences (4) over the terminal payout, subject to the profitability constraint $\mathrm{NI}_1 \geq \kappa$ and non-negative portfolio shares. With $\gamma \equiv \rho\sigma^2$

as in Section 3.3, define the analogous credit-risk cost:

$$\gamma_\eta \equiv \rho\, \sigma_\eta^2. \tag{42}$$

The parameter γ prices interest rate risk (as in the baseline), and γ_η prices credit risk. Substituting (40) into (4), the objective is:

$$U = \lambda r_{s,0} + \alpha\tau + S(\alpha_L) - c - \frac{\gamma}{2}(\lambda - \alpha)^2 - \frac{\gamma_\eta}{2}\,\alpha_L^2, \tag{43}$$

which is the baseline mean-variance problem (5) augmented by two credit terms: spread income net of credit risk.

E.3 Unconstrained Regime

When the profitability constraint is slack, the bank optimizes freely over α and α_L (with $\alpha_B = \alpha - \alpha_L$). The first-order conditions are:

$$\frac{\partial U}{\partial \alpha} = \tau + \gamma(\lambda - \alpha) = 0, \tag{44}$$

$$\frac{\partial U}{\partial \alpha_L} = \sigma_0 - \chi\,\alpha_L - \gamma_\eta\,\alpha_L = 0. \tag{45}$$

The first condition—identical to the baseline—pins down total duration at the interior optimum of equation (6):

$$\alpha^* = \lambda + \frac{\tau}{\gamma}. \tag{46}$$

The second directly determines the loan portfolio:

$$\alpha_L^* = \frac{\sigma_0}{\chi + \gamma_\eta}. \tag{47}$$

Proposition 5 (Portfolio Separation)**.** *In the unconstrained regime, the bank's portfolio separates into a duration decision and a lending decision:*

$$\alpha^* = \lambda + \frac{\tau}{\gamma}, \qquad \alpha_L^* = \frac{\sigma_0}{\chi + \gamma_\eta}. \tag{48}$$

Total duration is the same as in the baseline (hedge plus overshoot), while the loan portfolio balances the marginal spread income $S'(\alpha_L) = \sigma_0 - \chi\,\alpha_L$ against the marginal credit risk cost:

$$\underbrace{\sigma_0 - \chi\,\alpha_L^*}_{S'(\alpha_L^*)} = \underbrace{\gamma_\eta\,\alpha_L^*}_{\text{marginal credit risk cost}}. \tag{49}$$

The bond allocation absorbs all duration adjustment: $\alpha_B^* = \alpha^* - \alpha_L^*$.

Proposition 5 characterizes the interior solution. It requires a nonnegative bond position, $\alpha_L^* \leq \alpha^*$, that is, $\sigma_0/(\chi + \gamma_\eta) \leq \lambda + \tau/\gamma$, and that total duration fits on the balance sheet, $\lambda + \tau/\gamma \leq 1$. Both conditions hold with ample slack in the calibrations of Section E.6, where $\alpha_L^* = 0.20$ and $\alpha^* \geq \lambda = 0.586$. If credit demand were strong enough to violate the first condition, the bank would hold no bonds, total duration would be pinned by the loan margin, and duration adjustment would run through loans rather than bonds; the separation result applies to the interior region.

The two frictions contribute additively to pinning down the loan portfolio. When $\chi \to 0$ (constant spread), the loan allocation $\alpha_L^* = \sigma_0/\gamma_\eta$ is determined entirely by the credit-risk Sharpe ratio. When $\gamma_\eta \to 0$ (no credit risk), $\alpha_L^* = \sigma_0/\chi$ is determined entirely by the diminishing-returns technology.

The key observation is that α_L^* is *independent of* τ. The term premium governs total duration through the myopic component τ/γ. As the term premium falls, the bank's loan book is unaffected; only the bond portfolio shrinks.

E.4 Constrained Regime

When the term premium falls below a threshold, the profitability constraint binds. Because unconstrained income now includes the spread income $S(\alpha_L^*)$, this threshold lies below the baseline $\tilde{\tau}$ of equation (10): loans delay the onset of the constrained regime. When the constraint binds,

$$\lambda r_{s,0} + \alpha\tau + S(\alpha_L) - c = \kappa \qquad \Longrightarrow \qquad \alpha = \frac{\kappa + c - \lambda r_{s,0} - S(\alpha_L)}{\tau}. \tag{50}$$

Compared to the baseline (where $\alpha^*_{\text{bind}} = (\kappa + c - \lambda r_{s,0})/\tau$, equation (8)), spread income $S(\alpha_L)$ partially substitutes for duration income. Loans relax the profitability constraint: each dollar of spread income reduces the required duration by $1/\tau$.

With the constraint binding, the expected payout is pinned at $\mathbb{E}[\Pi_2] = 2\kappa$, and the bank's objective reduces to minimizing variance. The bank's single free variable is α_L, with α determined by (50) and $\alpha_B = \alpha - \alpha_L$. Define the *excess duration* as:

$$D(\alpha_L) \equiv \alpha - \lambda = \frac{\kappa + c - \lambda r_{s,0} - S(\alpha_L)}{\tau} - \lambda, \tag{51}$$

which measures how far the bank overshoots its interest rate hedge. In the baseline with $\alpha_L = 0$, $D_0 = (\kappa + c - \lambda r_{s,0})/\tau - \lambda > 0$: the constrained bank necessarily overshoots (see, for example, Figure 5). With loans, $D < D_0$ because spread income reduces required duration.

The bank solves:

$$\min_{\alpha_L} \; V(\alpha_L) = D(\alpha_L)^2\sigma^2 + \alpha_L^2\sigma_\eta^2. \tag{52}$$

The first-order condition equates the marginal reduction in interest rate variance to the marginal increase in credit variance:

$$\underbrace{D \cdot \frac{S'(\alpha_L)}{\tau} \cdot \sigma^2}_{\text{interest rate variance reduction}} = \underbrace{\alpha_L \cdot \sigma_\eta^2}_{\text{credit variance increase}}, \tag{53}$$

where $S'(\alpha_L) = \sigma_0 - \chi\,\alpha_L$ is the marginal spread income. The left side captures the benefit of substituting loans for bonds: each unit of loans generates $S'(\alpha_L)$ in additional spread income, which reduces required total duration by S'/τ and therefore lowers excess duration D by the same amount. Since interest rate variance is $D^2\sigma^2$, each unit reduction in D lowers it at rate $2D\,\sigma^2$, giving a total marginal benefit of $2D\,(S'/\tau)\,\sigma^2$ per unit of loans. The right side is the direct credit variance cost of expanding the loan book: since credit variance is $\alpha_L^2\sigma_\eta^2$, the marginal cost is $2\alpha_L\sigma_\eta^2$. The factors of two cancel, yielding (53).[27] Equation (53) characterizes the interior optimum: it requires $\alpha_L \geq 0$, a nonnegative bond position $\alpha_B = \alpha - \alpha_L \geq 0$, and total duration within the balance sheet, $\alpha \leq 1$; when a bound binds, the corresponding Kuhn–Tucker condition replaces the first-order condition. The limiting cases below and the illustrations in Section E.6 respect these bounds.

Limiting cases. The FOC (53) nests the two pure-friction models as special cases. When $\chi = 0$ (pure credit risk), $S' = \sigma_0$ is constant and the FOC gives α_L in closed form as a function of D; the loan book responds actively to τ because D depends on τ. When $\sigma_\eta^2 = 0$ (pure diminishing returns), the right side vanishes, forcing $S'(\alpha_L) = 0$ and hence $\alpha_L = \sigma_0/\chi$—the same as the unconstrained solution, independent of τ.

E.5 Loan Share Dynamics

How does a decline in the term premium affect the composition of the bank's long-asset portfolio? As τ falls, both loans and total duration respond, and the direction of the loan share $\ell \equiv \alpha_L/\alpha$ depends on which adjusts faster. We derive the comparative statics for total duration and the loan portfolio, then combine them to characterize the loan share response.

Total duration. Differentiating the profitability constraint (50) with respect to τ gives:

$$\frac{d\alpha}{d\tau} = -\frac{\alpha + S'\,d\alpha_L/d\tau}{\tau}. \tag{54}$$

[27]The second-order condition is satisfied: $d^2V/d\alpha_L^2 = 2\left[(S')^2\sigma^2/\tau^2 + \chi D\,\sigma^2/\tau + \sigma_\eta^2\right] > 0$, where positivity follows from $D > 0$ and $S' > 0$ at the interior optimum.

The first term, $-\alpha/\tau$, is the *mechanical effect*: holding loans fixed, a fall in τ forces the bank to hold more total duration to cover operating costs (this is the bonds-only response from the baseline, where $d\alpha/d\tau = -(\kappa + c - \lambda r_{s,0})/\tau^2$). The second term, $-S'(d\alpha_L/d\tau)/\tau$, is the *offset from loan expansion*: since $d\alpha_L/d\tau \leq 0$ (as we show next), more loans generate additional spread income, partially substituting for duration income and attenuating the required increase in α. In general, the mechanical effect dominates and $d\alpha/d\tau < 0$: total duration rises as τ falls.[28] A reversal ($d\alpha/d\tau > 0$) is theoretically possible only if the loan expansion is strong enough that the additional spread income more than compensates for the fall in τ, allowing the bank to reduce total duration.

When the loan market is deep enough that spread income can fully cover operating costs ($S(\sigma_0/\chi) = \sigma_0^2/(2\chi) \geq \kappa + c - \lambda r_{s,0}$), total duration remains bounded even as $\tau \to 0$. The profitability constraint still binds in this region—expected net income is pinned at the floor κ—but it no longer forces $\alpha \to \infty$. Instead, at the extreme ($\tau = 0$), the bank holds only loans with $\alpha_B = 0$ and $S(\alpha_L) = \kappa + c - \lambda r_{s,0}$, which pins α_L at a finite level.[29]

Loan portfolio response. Applying the implicit function theorem to the FOC (53) gives:

$$\frac{d\alpha_L}{d\tau} = -\frac{S'\,\sigma^2\,(D+\alpha)}{(S')^2\,\sigma^2 \;+\; \chi\,D\,\sigma^2\,\tau \;+\; \sigma_\eta^2\,\tau^2}, \tag{55}$$

where $D + \alpha = 2D + \lambda > 0$, $S' > 0$, and $D > 0$ at any interior constrained solution, so $d\alpha_L/d\tau < 0$: the bank expands its loan portfolio as the term premium falls.[30] The mechanism is as follows. As τ falls, the profitability constraint forces more total duration, which raises excess duration D and increases the marginal benefit of substituting loans for bonds (the left side of the FOC (53) grows). The bank expands lending until the marginal credit variance cost ($\alpha_L\sigma_\eta^2$) catches up.

The three terms in the denominator govern how aggressively the bank expands. The first, $(S')^2\sigma^2$, is a base term reflecting the curvature of spread income. The second, $\chi D\sigma^2\tau$, is the diminishing-returns channel (proportional to χ). The third, $\sigma_\eta^2\tau^2$, is the credit-risk channel (proportional to σ_η^2). When $\sigma_\eta^2 = 0$, the FOC forces $S'(\alpha_L) = 0$, pinning $\alpha_L = \sigma_0/\chi$ regardless of τ, so the loan portfolio does not respond: $d\alpha_L/d\tau = 0$.

[28] This holds in both limiting cases ($\chi = 0$ and $\sigma_\eta^2 = 0$).

[29] At the loans-only corner ($\alpha_B = 0$), the constraint $\alpha_L(\tau + \sigma_0 - (\chi/2)\alpha_L) = \kappa + c - \lambda r_{s,0}$ has a solution for all $\tau \geq 0$ provided $\sigma_0^2 \geq 2\chi(\kappa + c - \lambda r_{s,0})$. As $\tau \to 0$, $\alpha_L \to [\sigma_0 - \sqrt{\sigma_0^2 - 2\chi(\kappa + c - \lambda r_{s,0})}]/\chi$, which is finite.

[30] The numerator follows from $\partial D/\partial\tau = -\alpha/\tau$ and the product rule on $D \cdot S'/\tau$, using $\alpha + D = 2D + \lambda$. The denominator collects $-\partial f/\partial\alpha_L$ where $f = DS'\sigma^2/\tau - \alpha_L\sigma_\eta^2$.

Loan share. The loan share $\ell = \alpha_L/\alpha$ depends on both the loan and total duration responses derived above. Differentiating and substituting (54) yields:[31]

$$\frac{d\ell}{d\tau} = \frac{\ell}{\tau}(1-\Psi), \tag{56}$$

where the *substitution ratio* is

$$\Psi \equiv \left|\frac{d\alpha_L}{d\tau}\right| \cdot \frac{\tau + \ell S'}{\alpha_L}. \tag{57}$$

Since $\ell/\tau > 0$, the sign of $d\ell/d\tau$ is determined entirely by $1-\Psi$. When τ falls ($d\tau < 0$): if $\Psi > 1$ then $d\ell/d\tau < 0$, so ℓ rises; if $\Psi < 1$, ℓ falls. The substitution ratio compares the loan growth rate ($|d\alpha_L/d\tau|/\alpha_L$) to the total portfolio growth rate, with the factor $\tau + \ell S'$ measuring the effective marginal income per unit of total duration (bonds contribute τ, while loans contribute an additional $\ell S'$ through the credit spread).

Limiting cases and the combined model. Whether Ψ exceeds one depends on the parameters χ and σ_η^2.

Case 1: $\chi = 0$ (pure credit risk). Here $S' = \sigma_0$ is constant. From (55), the denominator reduces to $\sigma_0^2\sigma^2 + \sigma_\eta^2\tau^2$, and the FOC (53) gives $\sigma_\eta^2\tau^2 = D\sigma_0\sigma^2\tau/\alpha_L$. Substituting into the definition of Ψ:

$$\Psi = \frac{(D+\alpha)(\tau+\ell\sigma_0)}{\alpha_L\sigma_0 + D\tau}.$$

To verify $\Psi > 1$, expand the numerator using $\alpha\ell = \alpha_L$:

$$(D+\alpha)(\tau+\ell\sigma_0) = D\tau + D\ell\sigma_0 + \alpha\tau + \alpha_L\sigma_0 > \alpha_L\sigma_0 + D\tau,$$

where the inequality holds because $D\ell\sigma_0 + \alpha\tau > 0$. Hence $\Psi > 1$ and the loan share rises as τ falls. The bank substitutes aggressively into loans because the marginal comparison tilts in their favor: by the first-order condition (53), the variance relief each loan provides, $D\,\sigma^2\sigma_0/\tau$, grows as τ falls—required excess duration rises and each unit of spread income substitutes for more duration—while the marginal credit-risk cost $\alpha_L\sigma_\eta^2$ grows only in proportion to the loan share. The total credit variance cost $\alpha_L^2\sigma_\eta^2$ therefore rises along the path; it is the marginal trade-off, not the total cost, that favors loans, and each additional loan still generates constant spread income σ_0.

Case 2: $\sigma_\eta^2 = 0$ (pure diminishing returns). The FOC forces $S'(\alpha_L) = 0$, so $\alpha_L = \sigma_0/\chi$ is fixed independently of τ. Since $d\alpha_L/d\tau = 0$, we have $\Psi = 0 < 1$, and the loan share falls as τ falls: the loan book is pinned at the spread-maximizing level and cannot

[31] Differentiating $\ell = \alpha_L/\alpha$: $d\ell/d\tau = (1/\alpha)(d\alpha_L/d\tau) - (\alpha_L/\alpha^2)(d\alpha/d\tau)$. Substituting (54) for $d\alpha/d\tau$ and collecting terms gives $d\ell/d\tau = (d\alpha_L/d\tau)(\tau+\ell S')/(\alpha\tau) + \ell/\tau$. Writing $d\alpha_L/d\tau = -|d\alpha_L/d\tau|$ and using $\ell\alpha = \alpha_L$ yields $d\ell/d\tau = (\ell/\tau)(1-\Psi)$.

absorb any of the duration pressure, so all additional duration flows through bonds and $\ell \to 0$ as $\tau \to 0$.

Since all equilibrium quantities are continuous functions of χ and σ_η^2, and $\Psi > 1$ when $\chi = 0$ while $\Psi = 0$ when $\sigma_\eta^2 = 0$, by continuity there exists a boundary in (χ, σ_η^2) space separating the two regimes.

Proposition 6 (Loan Share Response)**.** *In the constrained regime:*

1. *The bank expands its loan portfolio as the term premium falls: $d\alpha_L/d\tau < 0$. When $\sigma_\eta^2 = 0$, loans are pinned at $\alpha_L = \sigma_0/\chi$ and do not respond: $d\alpha_L/d\tau = 0$.*
2. *When $\chi = 0$ (credit risk is the only friction), the loan share ℓ rises as τ falls. When $\sigma_\eta^2 = 0$ (diminishing returns is the only friction), the loan share falls as τ falls.*
3. *In the combined model ($\chi > 0$, $\sigma_\eta^2 > 0$), there exists a threshold in (χ, σ_η^2) space such that the loan share rises when credit risk is relatively more important and falls when diminishing returns are relatively more important.*

The credit risk extension delivers two main results. First, in the unconstrained regime the bank's portfolio separates: total duration is determined by the term premium (as in the baseline model), while the loan allocation is pinned by the balance of credit spreads against credit risk and diminishing returns. Loans do not respond to changes in term premia (such as a QE shock). Second, once the profitability constraint binds, loans become a tool for managing the NI shortfall. The bank expands lending as τ falls because loans generate spread income that reduces the required duration overshoot. A shallow credit demand curve (χ small) and volatile credit outcomes (σ_η^2 large) favor reallocation toward loans; a steep demand curve (χ large) and low credit risk (σ_η^2 small) favor reallocation toward bonds.

E.6 Graphical Illustration

Figures 13 and 14 illustrate the two polar cases, using the same baseline calibration as Figure 5, augmented with a credit spread $\sigma_0 = 0.01$. Both figures decompose the bank's long-asset portfolio into its loan component (teal shading) and bond component (blue shading), alongside total duration ($\alpha = \alpha_L + \alpha_B$, black line).

Pure credit risk ($\chi = 0$). Figure 13 shows the equilibrium when only credit risk disciplines lending ($\gamma_\eta = 0.05$, $\chi = 0$). In the unconstrained regime ($\tau > \tilde{\tau}$), the bank holds $\alpha_L = \sigma_0/\gamma_\eta = 0.20$ in loans, independent of τ, with bonds absorbing all duration adjustment. Once the profitability constraint binds, the bank actively substitutes out of bonds and into loans: α_L rises steeply while α_B falls. By $\tau = 0.025$, loans account for over 60 percent of the portfolio. Total duration grows as the term premium falls (the profitability constraint still tightens), but more slowly than in the bonds-only baseline because spread income absorbs part of the NI shortfall.

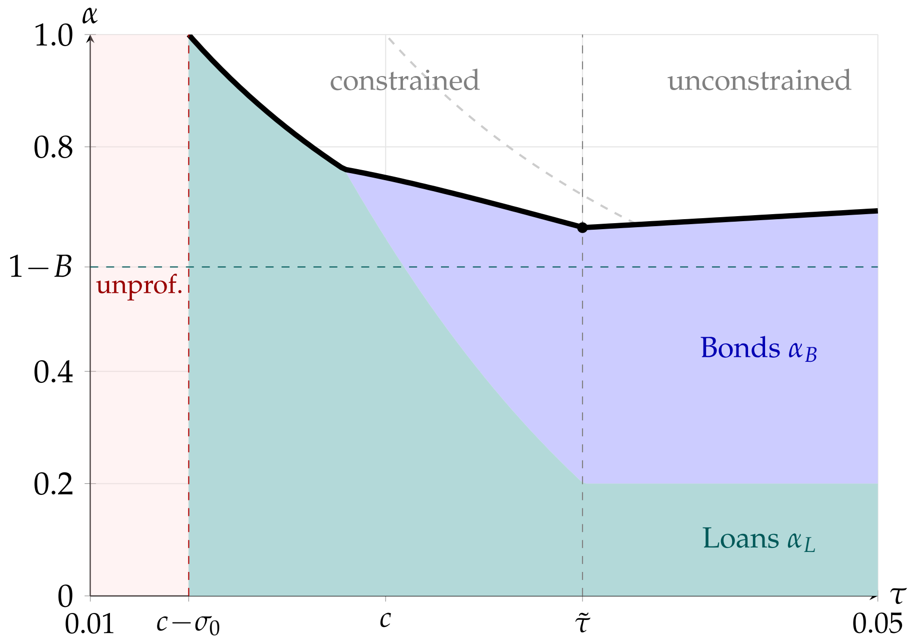


Figure 13: Portfolio composition with pure credit risk ($\chi = 0$). As τ falls below $\tilde{\tau}$, the bank substitutes bonds for loans, and the loan share rises. Total duration (black) grows more slowly than in the bonds-only baseline (gray dashed) because spread income partially offsets the NI shortfall. Parameters: $\sigma_0 = 0.01$, $\gamma_\eta = 0.05$, $\chi = 0$; all others as in Figure 5.

Pure diminishing returns ($\sigma_\eta^2 = 0$). Figure 14 shows the equilibrium when only the downward-sloping credit demand curve limits lending ($\chi = 0.05$, $\sigma_\eta^2 = 0$). The loan portfolio is fixed at $\alpha_L = \sigma_0/\chi = 0.20$ in both regimes: the bank maximizes spread income regardless of τ. All duration adjustment flows through bonds. In the constrained regime, α_B rises steeply as τ falls—the pure reach for duration of the baseline model, attenuated only by the constant contribution of spread income.

The contrast between the two figures is sharp. A constrained bank facing primarily credit risk (Figure 13) responds to QE by shifting from bonds to loans—a portfolio composition effect. A constrained bank facing primarily a downward-sloping demand curve (Figure 14) responds by extending the duration of its bond portfolio—the pure reach for duration of the baseline. Thus, the model can explain why constrained banks facing concentrated or saturated lending opportunities should exhibit rising bond duration with a stable loan allocation—and hence a declining loan share of the long-asset portfolio. This finding resonates with the experience of Silicon Valley Bank, which faced limited lending opportunities in the tech sector in the run-up to its failure, incentivizing the bank to increase its security holdings.

In the combined model ($\chi > 0$ and $\sigma_\eta^2 > 0$), the portfolio response is intermediate. The loan share rises or falls depending on the relative importance of credit risk versus diminishing returns, as characterized in Proposition 6. The key empirical object is whether the bank's lending technology exhibits more concavity in the spread (χ

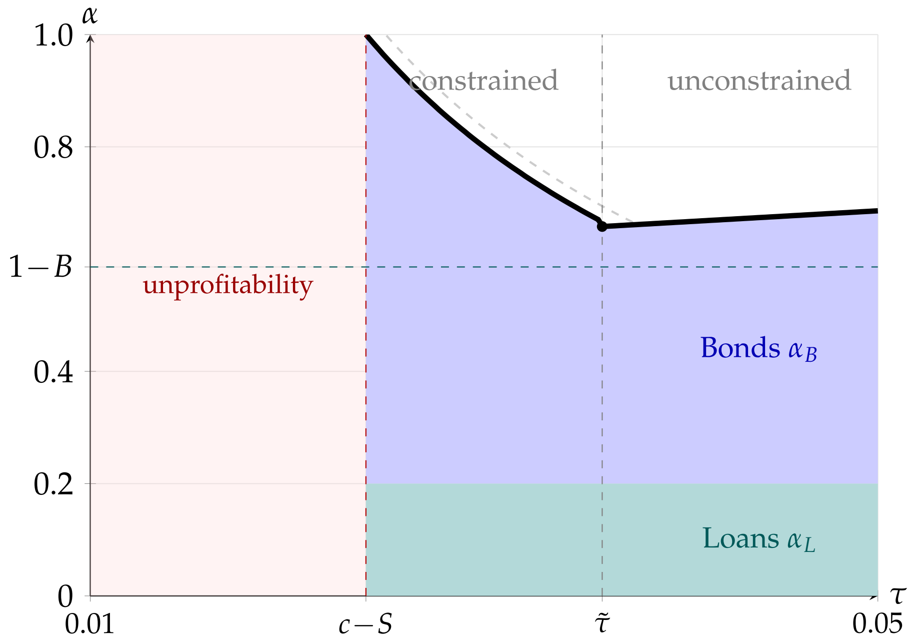


Figure 14: Portfolio composition with pure diminishing returns ($\sigma_\eta^2 = 0$). The loan portfolio is constant at $\alpha_L = \sigma_0/\chi$ in both regimes. All duration adjustment as τ falls flows through bonds. Total duration (black) lies below the bonds-only baseline (gray dashed) by the constant spread income, but follows the same convex shape. Parameters: $\sigma_0 = 0.01$, $\chi = 0.05$, $\sigma_\eta^2 = 0$; all others as in Figure 5.

large, demand curve steep) or more variance in credit outcomes (σ_η^2 large, credit risk dominant). The model provides a structural framework for distinguishing the two.

E.7 Discussion

Sections E.4–E.5 model two frictions that limit lending: a downward-sloping credit demand curve (parameterized by χ) and credit risk (parameterized by σ_η^2). This subsection considers alternative modeling choices and whether they would generate different predictions.

Alternatives within the existing framework. Several natural frictions produce predictions that are qualitatively equivalent to one of the two channels already in the model.

Regulatory or capacity constraints. If the bank faces a hard lending cap $\alpha_L \leq \bar{\alpha}_L$—due to regulatory limits, balance-sheet capacity, or a finite pool of eligible borrowers—the loan portfolio is fixed once the cap binds. All duration adjustment then flows through bonds, exactly as in the pure diminishing-returns case ($\sigma_\eta^2 = 0$).

Monitoring costs. If the bank must pay a convex monitoring cost $m(\alpha_L)$ to manage its loan portfolio, net spread income becomes $S(\alpha_L) - m(\alpha_L)$. This effectively steepens the declining marginal spread, increasing the curvature of the profitability constraint in α_L without introducing a qualitatively new channel. The prediction is again that of the diminishing-returns case: lending is limited by the flattening of net spread income.

Adverse selection and concentration risk. When the bank expands lending, it may attract riskier borrowers (adverse selection) or concentrate exposure in correlated sectors (concentration risk). Both frictions imply that credit risk scales more than proportionally with α_L—for instance, replacing $\alpha_L^2\sigma_\eta^2$ in (52) with a convex function $g(\alpha_L)$ satisfying $g'' > 2\sigma_\eta^2$. At the same time, adverse selection may raise ex-ante spreads (since riskier borrowers pay higher rates), so $S'(\alpha_L)$ need not be decreasing as our diminishing-returns specification assumes. This combination—rising spreads but accelerating risk—does not map cleanly into either polar case. However, the qualitative prediction is similar to the credit risk channel: the bank expands lending as τ falls, but the convexity of g provides stronger damping at high α_L, limiting the expansion more than in the linear-risk case.

Alternatives with different predictions. Two extensions would generate predictions not nested by the current model.

Correlated credit and interest rate risk. The model assumes $\eta \perp \varepsilon$, so the bank's variance decomposes as $D^2\sigma_\varepsilon^2 + \alpha_L^2\sigma_\eta^2$. If credit losses are positively correlated with rate increases ($\mathrm{Cov}(\varepsilon,\eta) > 0$)—as is plausible if rising rates trigger defaults—the variance includes a cross term $2D\alpha_L \mathrm{Cov}(\varepsilon,\eta)$. The covariance enters the constrained trade-off on both sides: holding more loans loads credit risk onto a portfolio that is already long duration (a marginal cost $D\,\mathrm{Cov}(\varepsilon,\eta)$), while the additional spread income reduces the required excess duration and hence the covariance exposure itself (a marginal benefit $\alpha_L S'(\alpha_L)\,\mathrm{Cov}(\varepsilon,\eta)/\tau$, since $dD/d\alpha_L = -S'(\alpha_L)/\tau$ along the binding constraint). Loan expansion is therefore damped relative to the independent case when $D > \alpha_L S'(\alpha_L)/\tau$, which holds in the deep-constrained region, where excess duration is large and the marginal spread is small. Negative correlation (credit losses fall when rates rise) reverses the comparison, making loans a partial hedge against interest rate risk and strengthening the substitution toward loans.

Heterogeneous loan duration. The model assumes loans and bonds contribute equally to total duration ($\alpha = \alpha_B + \alpha_L$). In practice, loans often have shorter effective duration than long-term bonds. If loans carry duration weight $\delta < 1$, so that $\alpha = \alpha_B + \delta\alpha_L$, the profitability constraint becomes $(\alpha_B + \delta\alpha_L)\tau + S(\alpha_L) = \kappa + c - \lambda r_{s,0}$. Because each unit of loans generates less duration income ($\delta\tau$ instead of τ), the bank must hold more total assets to cover costs, weakening loans as a substitute for bonds. Moreover, if term premia are increasing in duration, the relevant spread for loans would be $\delta\tau + \sigma(\alpha_L)$ rather than $\tau + \sigma(\alpha_L)$, further reducing the attractiveness of lending relative to bond holdings. The prediction is that shorter loan duration reduces the substitution toward loans in the constrained regime, pushing the model closer to the diminishing-returns case even when credit risk is the dominant friction.

Loan illiquidity and the nature of run risk. The model treats loans and bonds symmetrically as marketable claims, differing only in risk and return. In practice, loans are informationally sensitive and rarely trade in deep secondary markets, so a bank that responds to profitability pressure by reaching for credit risk tilts its portfolio toward illiquid assets. This changes the nature of its exposure to runs. The deposit franchise runs of Section 4 operate through insolvency and do not rely on asset illiquidity: the bank's bonds can be sold at market value to meet withdrawals. A loan-heavy bank, in contrast, could meet large withdrawals only by liquidating assets at a discount, so the classic mechanism of Diamond and Dybvig (1983) becomes operative in addition to the franchise channel.

Whether reaching for credit risk raises run risk on net is ambiguous within the model, since spread income raises current earnings and thereby the bank's solvency buffer, and credit shocks replace interest rate shocks as the relevant source of losses. The composition of fragility shifts, however: reaching for duration in liquid securities concentrates exposure to rising rates while preserving the bank's ability to meet withdrawals at market prices, whereas reaching for credit risk substitutes fragility through illiquidity for fragility through duration.

The two margins can also operate in combination: banks may extend illiquid loans at longer maturities, taking on duration and credit risk within the same asset. The historical evidence in Section 2 suggests that this margin is relevant—the average maturity of bank loan portfolios rose from around 2.5 years in the early 1990s to a peak of about 6 years in 2022 (Figure 4). A bank that reaches in this way combines the two sources of fragility: rising interest rates erode its solvency, while asset illiquidity exposes it to liquidation-driven runs.

F Endogenous Term Premium

In this section, we extend the baseline model to allow the term premium to respond endogenously to banks' aggregate demand for long-term assets. This extension is motivated by recent evidence that banks actively influence the pricing of interest rate risk. In particular, Haddad and Sraer (2020) show that the aggregate interest rate exposure of the banking sector is a strong predictor of excess Treasury bond returns, suggesting that banks' portfolio choices partly determine the term premium.

F.1 Setup and Equilibrium

The term premium now has two components:

$$\tau(\alpha) = \hat{\tau} \;-\; \mu\alpha, \tag{58}$$

where $\hat{\tau}$ is the exogenous component of the term premium and $\mu > 0$ measures the price impact of bank demand on the long-term bond market. Quantitative easing compresses $\hat{\tau}$ directly. When banks collectively tilt toward long-term assets (higher α), they bid up long-bond prices and compress the term premium further. The microfoundation follows standard preferred-habitat logic: long-term bonds are held by price-sensitive arbitrageurs (here, banks) and price-insensitive preferred-habitat investors. When banks increase demand, the market-clearing yield on the long bond falls.

Each bank is small and takes τ as given. A symmetric equilibrium is a pair (α^*, τ^*) such that (i) α^* solves the bank's problem given τ^*, and (ii) $\tau^* = \hat{\tau} - \mu\alpha^*$.

Unconstrained regime. When the profitability constraint is slack, the first-order condition in equation (6) gives $\alpha = \lambda + \tau/\gamma$. Substituting $\tau = \hat{\tau} - \mu\alpha$:

$$\alpha = \lambda + \frac{\hat{\tau} - \mu\alpha}{\gamma} \quad\Longrightarrow\quad \alpha^*_{\text{int}} = \frac{\lambda\gamma + \hat{\tau}}{\gamma + \mu}. \tag{59}$$

The equilibrium term premium is:

$$\tau^*_{\text{int}} = \frac{\gamma(\hat{\tau} - \mu\lambda)}{\gamma + \mu}. \tag{60}$$

Compared to the exogenous model (where $\alpha^*_{\text{int}} = \lambda + \hat{\tau}/\gamma$ and $\tau^* = \hat{\tau}$), the endogenous feedback rotates the FOC curve downward and compresses the equilibrium term premium to a fraction $\gamma/(\gamma + \mu)$ of the exogenous premium, net of a correction for the hedge. When $\mu = 0$, we recover the exogenous model.

Constrained regime. When the profitability constraint binds, $\lambda r_{s,0} + \alpha\tau - c = \kappa$ as in equation (7), and thus $\alpha = (\kappa + c - \lambda r_{s,0})/\tau$, the minimum required duration of equation (8). Substituting $\tau = \hat{\tau} - \mu\alpha$:

$$\alpha = \frac{\kappa + c - \lambda r_{s,0}}{\hat{\tau} - \mu\alpha} \quad\Longrightarrow\quad \mu\alpha^2 - \hat{\tau}\,\alpha + (\kappa + c - \lambda r_{s,0}) = 0. \tag{61}$$

By the quadratic formula, this has two roots:

$$\alpha_{\pm} = \frac{\hat{\tau} \pm \sqrt{\hat{\tau}^2 - 4\mu(\kappa + c - \lambda r_{s,0})}}{2\mu}, \tag{62}$$

with $\alpha_- \le \alpha_+$. Both roots satisfy the profitability constraint with equality, so the equilibrium-relevant root must be selected by an additional criterion.

Root selection. The regime boundary $\hat{\tau}^*$ is defined as the value of $\hat{\tau}$ at which the unconstrained solution first violates the profitability constraint as the exogenous term premium falls. That is, $\hat{\tau}^*$ is the largest $\hat{\tau}$ solving $\alpha^*_{\text{int}}(\hat{\tau}) \cdot \tau^*_{\text{int}}(\hat{\tau}) = \kappa + c - \lambda r_{s,0}$, where α^*_{int} and τ^*_{int} are the unconstrained solutions from (59)–(60).[32] The constrained solution must connect continuously to the unconstrained solution at this boundary, which pins down the root: at $\hat{\tau} = \hat{\tau}^*$, exactly one of $\alpha_{\pm}$ equals $\alpha^*_{\text{int}}(\hat{\tau}^*)$. We focus on the case $\mu < \mu_{\text{crit}}$ where the price impact is moderate.[33] The equilibrium-relevant (smaller) root is:

$$\alpha^*_{\text{bind}} = \alpha_- = \frac{\hat{\tau} - \sqrt{\hat{\tau}^2 - 4\mu(\kappa + c - \lambda r_{s,0})}}{2\mu}, \tag{63}$$

with corresponding equilibrium term premium $\tau^*_{\text{bind}} = \hat{\tau} - \mu\alpha^*_{\text{bind}}$, which simplifies to:

$$\tau^*_{\text{bind}} = \frac{\hat{\tau} + \sqrt{\hat{\tau}^2 - 4\mu(\kappa + c - \lambda r_{s,0})}}{2}. \tag{64}$$

Note the sign flip: the smaller root for α gives the *larger* equilibrium term premium, since less duration exposure means less demand compression. As the exogenous component $\hat{\tau}$ falls (due to QE), the discriminant shrinks and α^*_{bind} increases: the bank reaches for duration, just as in the exogenous model, but the feedback loop amplifies the effect. The equilibrium term premium τ^*_{bind} falls, reflecting the additional compression from banks' own demand.

[32]Since α^*_{int} and τ^*_{int} are both linear in $\hat{\tau}$, their product is a quadratic in $\hat{\tau}$. The equation $\alpha^*_{\text{int}} \cdot \tau^*_{\text{int}} = \kappa + c - \lambda r_{s,0}$ therefore has at most two solutions. For all economically relevant parameter values, only one solution is positive, so $\hat{\tau}^*$ is uniquely determined.

[33]The constrained region spans $\hat{\tau} \in [\hat{\tau}_{\text{crit}}, \hat{\tau}^*]$ (where $\hat{\tau}_{\text{crit}}$ is defined in Section F.2 below) and shrinks as μ increases. At $\mu = \mu_{\text{crit}}$, the region vanishes: $\hat{\tau}^* = \hat{\tau}_{\text{crit}}$ and the two roots merge at $\alpha_{\text{merge}} = \hat{\tau}_{\text{crit}}/(2\mu)$. Requiring this merged point to also satisfy the unconstrained FOC yields, after substitution, $u^2 + \gamma\lambda u - \gamma(\kappa + c - \lambda r_{s,0}) = 0$ with $u = \sqrt{\mu(\kappa + c - \lambda r_{s,0})}$—the same quadratic as condition (9) that defines $\tilde{\tau}$ in equation (10). Hence $\mu_{\text{crit}} = \tilde{\tau}^2/(\kappa + c - \lambda r_{s,0})$. When $\mu < \mu_{\text{crit}}$, continuity selects α_-: at $\hat{\tau} = \hat{\tau}^*$, the unconstrained solution coincides with the lower root, and α_- nests the exogenous model as $\mu \to 0$, with $\alpha_- \to (\kappa + c - \lambda r_{s,0})/\hat{\tau}$ while α_+ diverges. The bank reaches for duration as $\hat{\tau}$ falls. The upper root, in contrast, is off the equilibrium path and generally violates the balance-sheet bound: $\alpha_+ > 1$ for all $\hat{\tau} > \mu + \kappa + c - \lambda r_{s,0}$, so it satisfies $\alpha_+ \le 1$ at most in a band of width $(\sqrt{\mu} - \sqrt{\kappa + c - \lambda r_{s,0}})^2$ above $\hat{\tau}_{\text{crit}}$—about 10^{-5} under the calibration of Figure 15. When $\mu > \mu_{\text{crit}}$, continuity instead selects α_+, which *decreases* as $\hat{\tau}$ falls—the opposite of the reach-for-duration mechanism.

F.2 Instability and Amplification

The constrained equilibrium exists if and only if the discriminant in (61) is non-negative:

$$\hat{\tau}^2 \geq 4\mu(\kappa + c - \lambda r_{s,0}) \quad \Longleftrightarrow \quad \hat{\tau} \geq 2\sqrt{\mu(\kappa + c - \lambda r_{s,0})} \equiv \hat{\tau}_{\text{crit}}. \tag{65}$$

Proposition 7 (Financial Instability Threshold)**.** *Suppose the price impact satisfies $\mu \geq \kappa + c - \lambda r_{s,0}$, as in our calibration, so that the tangency lies within the feasible portfolio set. Then there exists a critical level of the exogenous term premium component*

$$\hat{\tau}_{\text{crit}} = 2\sqrt{\mu(\kappa + c - \lambda r_{s,0})}, \tag{66}$$

below which no equilibrium in which banks meet the profitability constraint exists. At this threshold, the feedback between bank demand for duration and term premium compression becomes self-reinforcing.[34]

Intuition. Below $\hat{\tau}_{\text{crit}}$, a feedback loop emerges: banks need duration income to survive $\rightarrow$ their demand compresses the term premium $\rightarrow$ which tightens the profitability constraint further $\rightarrow$ requiring yet more duration, and so on without bound. The two roots of the quadratic merge at $\hat{\tau}_{\text{crit}}$, and below it no real solution exists.

Amplification. In the constrained regime, the endogenous α^*_{bind} exceeds the exogenous benchmark $(\kappa + c - \lambda r_{s,0})/\hat{\tau}$ at every $\hat{\tau}$, and the amplification factor grows as $\hat{\tau}$ approaches $\hat{\tau}_{\text{crit}}$. The sensitivity of α^* to QE also diverges near the instability threshold. Differentiating (63) with respect to $\hat{\tau}$:

$$\frac{d\alpha^*_{\text{bind}}}{d\hat{\tau}} = \frac{1}{2\mu}\left(1 - \frac{\hat{\tau}}{\sqrt{\hat{\tau}^2 - 4\mu(\kappa + c - \lambda r_{s,0})}}\right) < 0, \tag{67}$$

so α^* increases as $\hat{\tau}$ falls. Since QE compresses $\hat{\tau}$ directly, the response of maturity mismatch to a QE-induced compression is $-d\alpha^*_{\text{bind}}/d\hat{\tau} > 0$, which diverges as $\hat{\tau} \to \hat{\tau}_{\text{crit}}$.

Regime characterization. The three regimes are summarized in Table A.11. The boundary $\hat{\tau}^*$ between the unconstrained and constrained regimes is defined by the condition that the unconstrained α^*_{int} just satisfies the profitability constraint with equality (see Section 3).

[34] The tangency at $\hat{\tau}_{\text{crit}}$ occurs at $\alpha_{\text{merge}} = \sqrt{(\kappa + c - \lambda r_{s,0})/\mu}$, which lies in the feasible set $\alpha \leq 1$ provided $\mu \geq \kappa + c - \lambda r_{s,0}$; otherwise the bank reaches the corner $\alpha = 1$—the unprofitability boundary—at $\hat{\tau} = \mu + \kappa + c - \lambda r_{s,0}$, before the tangency; in that case the no-equilibrium boundary is this corner rather than $\hat{\tau}_{\text{crit}}$. Under the leverage-capacity interpretation of Section 3.6 ($\alpha \in [0, \bar{\alpha}]$), the condition relaxes to

Table A.11: Three regimes of the endogenous term premium model.

Regime	**Condition**	**Behavior**
Unconstrained	$\hat{\tau} > \hat{\tau}^*$	α^* given by (59), increasing in $\hat{\tau}$
Constrained (stable)	$\hat{\tau}_{\text{crit}} \leq \hat{\tau} \leq \hat{\tau}^*$	α^* given by (63), decreasing and convex in $\hat{\tau}$, with feedback amplification
No equilibrium	$\hat{\tau} < \hat{\tau}_{\text{crit}}$	Feedback loop: profitability constraint infeasible

F.3 Graphical Illustration and Discussion

F.3.1 Equilibrium Maturity Mismatch

Figure 15 compares the equilibrium α^* in the exogenous and endogenous models, both plotted against the exogenous component $\hat{\tau}$. The price impact parameter is $\mu = 0.026$; all other parameters are as in Figure 5. The endogenous feedback rotates the FOC curve downward from $\alpha = \lambda + \hat{\tau}/\gamma$ to $\alpha = (\lambda\gamma + \hat{\tau})/(\gamma + \mu)$ (blue vs. gray solid). In the constrained regime, the endogenous binding-constraint curve (orange) lies above the exogenous one (gray dashed): the bank must hold *more* duration in the endogenous model because its own demand compresses the very term premium it is trying to earn. An instability threshold $\hat{\tau}_{\text{crit}} \approx 0.051$ emerges, well to the right of the exogenous unprofitability threshold at $c = 0.025$: the feedback loop causes failure at a much higher $\hat{\tau}$ than simple unprofitability would.

F.3.2 Equilibrium Term Premium

Figure 16 plots the equilibrium term premium τ^* against the exogenous component $\hat{\tau}$. The dashed 45-degree line represents the exogenous benchmark ($\tau^* = \hat{\tau}$ when $\mu = 0$). With the feedback, τ^* falls below the 45-degree line. The shaded wedge between the two represents the additional compression caused by banks' collective demand for duration.

In the *unconstrained regime* (right of $\hat{\tau}^*$), the equilibrium term premium is a linear function of $\hat{\tau}$ with slope $\gamma/(\gamma + \mu) < 1$: a marginal change in the exogenous component passes through at a constant rate, and the wedge $\hat{\tau} - \tau^*_{\text{int}} = \mu(\hat{\tau} + \gamma\lambda)/(\gamma + \mu)$ increases linearly in $\hat{\tau}$. In the *constrained regime* (between $\hat{\tau}_{\text{crit}}$ and $\hat{\tau}^*$), τ^*_{bind} given by (64) falls as $\hat{\tau}$ decreases, reflecting the additional compression from the feedback loop. At $\hat{\tau}_{\text{crit}}$, the equilibrium term premium is $\tau^*_{\text{crit}} = \hat{\tau}_{\text{crit}}/2 = \sqrt{\mu(\kappa + c - \lambda r_{s,0})}$, which lies below the baseline threshold $\tilde{\tau}$ of equation (10) precisely when $\mu < \mu_{\text{crit}}$.

$\mu \geq (\kappa + c - \lambda r_{s,0})/\bar{\alpha}^2$. The calibration of Figure 15 satisfies it: $\mu = 0.026 \geq 0.025$, so $\alpha_{\text{merge}} \approx 0.98$.

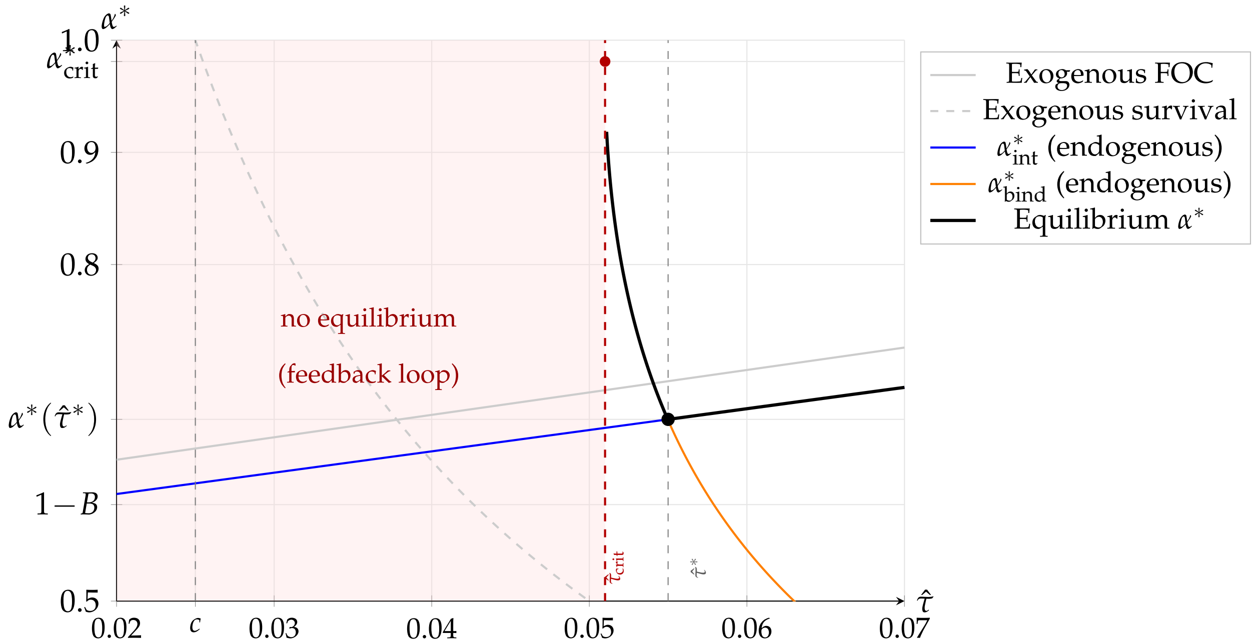


Figure 15: Equilibrium maturity mismatch: exogenous vs. endogenous term premium. The endogenous feedback ($\mu > 0$) rotates the FOC curve downward (blue vs. gray solid) and shifts the solvency curve upward and to the right (orange vs. gray dashed), amplifying the reach for duration. An instability threshold $\hat{\tau}_{\text{crit}}$ emerges well to the right of the exogenous unprofitability threshold c. Parameters: $\mu = 0.026$; all others as in Figure 5.

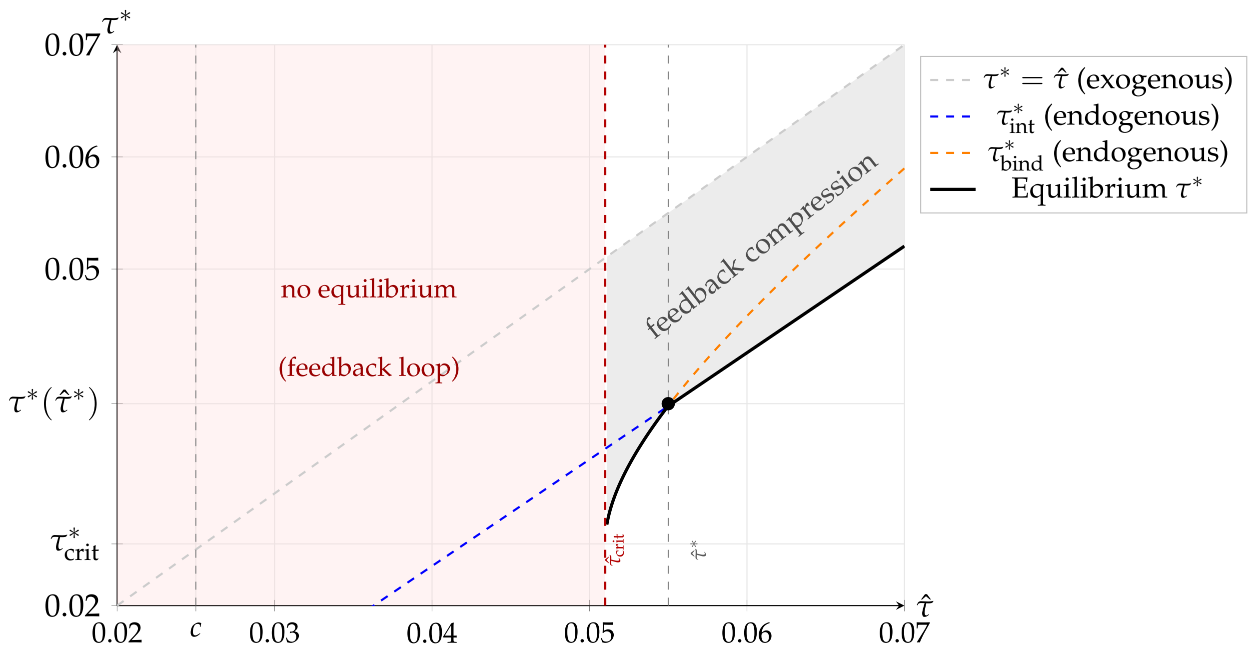


Figure 16: Equilibrium term premium with endogenous feedback. The shaded wedge between the 45-degree line and the equilibrium path represents the additional term premium compression caused by banks' collective demand for duration. The wedge widens as $\hat{\tau}$ falls toward $\hat{\tau}_{\text{crit}}$. Parameters: $\mu = 0.026$; all others as in Figure 5.

Discussion. The model delivers a feedback loop that links term premium compression to bank fragility. A decline in the term premium squeezes banks' NIs until the profitability constraint eventually binds. Banks respond by reaching for duration—increasing α—which further compresses the term premium through the demand channel ($\mu\alpha$). This additional compression tightens the profitability constraint further, requiring yet more duration exposure. Near the instability threshold, the system becomes fragile: small additional term premium compression triggers large increases in maturity mismatch and, ultimately, the possibility of bank failure.

This amplification mechanism is analogous to fire-sale spirals but operates through the *asset side* of the balance sheet via duration risk rather than through leverage and margin constraints (Brunnermeier and Pedersen, 2009). The implication is that term premium changes, possibly driven by QE, have a non-linear, potentially destabilizing effect on the banking sector that is absent from models with exogenous term premia.

G Derivations for the Deposit Franchise Run Model

This section derives the expressions in Section 4. Throughout, $\mathrm{NI}_1 = \lambda r_{s,0} + \alpha\tau - c$ is per-period net income and $\Pi_2 = 2\mathrm{NI}_1 + (\lambda - \alpha)\varepsilon$ is the terminal payout net of initial equity, as in equation (2).

G.1 Run Condition and Threshold

When uninsured depositors withdraw at $t = 1$, the bank loses the period-2 franchise on their deposits: the volume $(1 - e)u$ times the deposit spread $(1 - \beta)r_{s,1} = (1 - \beta)(r_{s,0} + \varepsilon)$, which is the franchise loss $\Delta(\varepsilon)$ in equation (13). A run leaves the bank insolvent if terminal net worth—equity plus cumulative net income—falls short of the franchise loss:

$$e + \Pi_2(\varepsilon, \alpha) \;=\; e + 2\mathrm{NI}_1 + (\lambda - \alpha)\varepsilon \;<\; \Delta(\varepsilon). \tag{68}$$

Collecting the terms in ε,

$$e + 2\mathrm{NI}_1 - (1 - e)(1 - \beta)r_{s,0}u \;<\; \left[\alpha - \lambda + (1 - e)(1 - \beta)u\right]\varepsilon \;=\; \left(\alpha - \tilde{\lambda}\right)\varepsilon, \tag{69}$$

with $\tilde{\lambda} \equiv \lambda - (1 - e)(1 - \beta)u$. Using $\lambda = e + (1 - e)(1 - \beta)$, the adjusted base is $\tilde{\lambda} = e + (1 - e)(1 - \beta)(1 - u)$, the non-repricing funding that survives a run. Since the equilibrium portfolio satisfies $\alpha^* \geq \lambda > \tilde{\lambda}$ in every regime, the coefficient on ε is positive and (69) is equivalent to

$$\varepsilon \;>\; \frac{e + 2\mathrm{NI}_1 - (1 - e)(1 - \beta)r_{s,0}u}{\alpha - \tilde{\lambda}}. \tag{70}$$

Dropping the constant $(1-e)(1-\beta)r_{s,0}u$—which is second order relative to equity and exactly zero at the zero-lower-bound calibration ($r_{s,0}=0$)—and substituting NI_1 yields the run threshold in equation (14).

G.2 Run Onset in the Constrained Regime

At the constrained portfolio $\alpha^*_{\text{bind}} = (\kappa + c - \lambda r_{s,0})/\tau$ of equation (8), net income is pinned at $\mathrm{NI}_1 = \kappa$ and

$$\alpha^*_{\text{bind}} - \tilde{\lambda} = \frac{\kappa + c - \lambda r_{s,0} - \tilde{\lambda}\tau}{\tau},$$

so the run threshold becomes equation (15). Differentiating,

$$\frac{d\varepsilon^*(\tau)}{d\tau} = \frac{(e+2\kappa)(\kappa + c - \lambda r_{s,0})}{\left(\kappa + c - \lambda r_{s,0} - \tilde{\lambda}\tau\right)^2} > 0,$$

so the bank grows more fragile as the term premium falls. Runs enter the shock support when $\varepsilon^*(\tau) = \bar{\varepsilon}$; solving this linear equation for τ gives the run onset threshold in equation (16). With $\varepsilon \sim U[-\bar{\varepsilon}, \bar{\varepsilon}]$, the run probability for $\tau \le \tau_{\text{run}}$ is $p(\tau) = \Pr\left[\varepsilon > \varepsilon^*(\tau)\right] = \left(\bar{\varepsilon} - \varepsilon^*(\tau)\right)/(2\bar{\varepsilon})$.

The comparative statics of τ_{run} follow directly from equation (16):

$$\frac{\partial \tau_{\text{run}}}{\partial \bar{\varepsilon}} = \frac{(\kappa + c - \lambda r_{s,0})(e+2\kappa)}{\left(e + 2\kappa + \tilde{\lambda}\bar{\varepsilon}\right)^2} > 0,$$

and higher equity lowers τ_{run} through both terms: the numerator weakly falls, since $\partial\lambda/\partial e = \beta$ raises non-repricing income, and the denominator rises, since $\partial\tilde{\lambda}/\partial e = 1 - (1-\beta)(1-u) = \psi > 0$.

G.3 Immunity Threshold

Within the constrained regime, $\varepsilon^*(\tau)$ is increasing in τ and the feasible range ends at $\alpha = 1$, where $\tau = \kappa + c - \lambda r_{s,0}$. The minimum of the run threshold is therefore

$$\min_{\tau} \varepsilon^*(\tau) = \frac{e+2\kappa}{1-\tilde{\lambda}}.$$

Runs are impossible at every term premium if and only if this minimum weakly exceeds $\bar{\varepsilon}$. Using $1 - \tilde{\lambda} = (1-e)\left[\beta + (1-\beta)u\right] = (1-e)\psi$ with $\psi \equiv 1 - (1-u)(1-\beta)$,

$$e + 2\kappa \ge \bar{\varepsilon}(1-e)\psi \quad \Longleftrightarrow \quad e \ge e^* \equiv \frac{\bar{\varepsilon}\psi - 2\kappa}{1 + \bar{\varepsilon}\psi},$$

the immunity threshold in Section 4.

G.4 Fundamental Insolvency

Absent a run, the bank's terminal net worth turns negative when $e + \Pi_2 < 0$, that is, for $\varepsilon > \tilde{\varepsilon}(\alpha) \equiv (e + 2\text{NI}_1)/(\alpha - \lambda)$ whenever $\alpha > \lambda$. Since $\tilde{\lambda} < \lambda$, we have $\alpha - \lambda < \alpha - \tilde{\lambda}$ and therefore $\tilde{\varepsilon}(\alpha) > \varepsilon^*(\alpha)$: the run threshold is crossed strictly first, so self-fulfilling runs—not fundamental insolvency—are the binding failure mode. At the most exposed constrained portfolio, $\alpha = 1$ and $\text{NI}_1 = \kappa$, the insolvency threshold is $\tilde{\varepsilon} = (e + 2\kappa)/(1 - \lambda)$. The calibration in Section 4 sets $\bar{\varepsilon} = 0.14$ below this value (≈ 0.145 at $\kappa = 0$), placing fundamental insolvency outside the shock support for all feasible portfolios.

G.5 Market-Value Formulation

At $t = 1$, the bank's remaining cash flow to shareholders is $e + \Pi_2$ at $t = 2$, so its market value is $V_1 = (e + \Pi_2)/(1 + r_{s,0} + \varepsilon)$. Dividing the run condition (68) by $1 + r_{s,0} + \varepsilon$ gives the equivalent present-value condition

$$V_1 \; < \; \Delta_V(\varepsilon) \equiv \frac{\Delta(\varepsilon)}{1 + r_{s,0} + \varepsilon},$$

the formulation in Drechsler et al. (2026). Terminal values are more convenient algebraically because Π_2 is linear in ε.

H Additional results from the Dynamic Model

We provide a proof of Proposition 2, show the solution from the polar cases to further build the intuition of the global solution, and, finally, map the dynamic model into the two-period model. Throughout, we write $r_t \equiv r_{s,t}$ for the short rate.

Proof of Proposition 2. Assume the constraint is always slack. Then, the HJB (23) is

$$\begin{aligned} \frac{\omega}{1-\rho} = \frac{\xi_t^{AU}}{\xi^{AU}} \frac{1}{1-\rho} + \max_{\alpha} \Bigg\{ & \alpha_t \tau + [1 + (1-\beta)\delta] r_t - c\delta \\ & - \frac{\rho}{2} (\alpha \sigma_\ell)^2 + \frac{\xi_r^{AU}}{\xi^{AU}} \frac{\kappa_r}{1-\rho} (\bar{r} - r_t) \\ & + \frac{1}{2} \frac{\xi_{rr}^{AU}}{\xi^{AU}} \frac{\sigma_r^2}{1-\rho} - \frac{\xi_r^{AU}}{\xi^{AU}} \alpha \sigma_\ell \sigma_r \Bigg\}. \end{aligned}$$

The first-order conditions (FOC) are

$$\tau - \alpha_t \rho \sigma_\ell^2 - \frac{\xi_r^{AU}}{\xi^{AU}} \sigma_\ell \sigma_r = 0.$$

Substituting the FOC into the HJB, we obtain a partial differential equation for $\xi^{AU}(t,r)$, which is exponentially affine,

$$\xi^{AU}(t,r) = \exp\left(\Omega^{AU}(t) + \Lambda^{AU}(t)\, r_t\right), \tag{71}$$

with boundary condition

$$\xi^{AU}(T,r) = \exp\left(\Omega^{AU}(T) + \Lambda^{AU}(T)\, r_t\right) = 1.$$

Substituting (71) into the HJB, we obtain ordinary differential equations for the functions $\Lambda^{AU}(t)$ and $\Omega^{AU}(t)$,

$$\frac{\Lambda_t^{AU}}{1-\rho} + [1 + (1-\beta)\,\delta] - \Lambda^{AU}\frac{\kappa_r}{1-\rho} = 0,$$
$$\frac{\Omega_t^{AU}}{1-\rho} - \frac{\omega}{1-\rho} - c\delta + \Lambda^{AU}\frac{\kappa_r}{1-\rho}\bar{r} + \frac{1}{2}\left(\Lambda^{AU}\right)^2\frac{\sigma_r^2}{1-\rho} + \frac{\left(\tau - \Lambda^{AU}\sigma_\ell\sigma_r\right)^2}{2\rho\sigma_\ell^2} = 0.$$

The solution for $\Lambda^{AU}(t)$ is

$$\Lambda^{AU}(t) = \frac{1-\rho}{\kappa_r}[1 + (1-\beta)\,\delta]\,(1 - \exp(-\kappa_r(T-t)))\,. \tag{72}$$

Equation (72) is key because $\Lambda^{AU}(t) = \frac{\xi_r^{AU}}{\xi^{AU}}$, the key term in banks' hedging demand for long-term assets. We substitute it into the optimal demand for the long-term asset to obtain

$$\alpha_t^{AU} = \frac{\tau}{\rho\sigma_\ell^2} - \frac{\sigma_r}{\rho\sigma_\ell}\frac{1-\rho}{\kappa_r}[1 + (1-\beta)\,\delta]\,(1 - \exp(-\kappa_r(T-t)))\,. \tag{73}$$

We now turn to the case in which the constraint is always binding. The binding profitability constraint is given by

$$\alpha_t^C\tau + \left[1 + (1-\beta)\delta\right]r_t - c\delta = \kappa.$$

Hence, assuming $\tau > 0$, the portfolio share in long-term assets is pinned down by

$$\alpha_t^C = \frac{\kappa + c\delta - \left[1 + (1-\beta)\delta\right]r_t}{\tau}. \tag{74}$$

Define

$$K \equiv 1 + (1-\beta)\delta, \qquad a_C \equiv \frac{\kappa + c\delta}{\tau}, \qquad b_C \equiv \frac{K}{\tau},$$

so that

$$\alpha_t^C = a_C - b_C r_t.$$

Therefore, once the profitability constraint binds, the bank no longer chooses α_t freely. Importantly, note that the term premium, τ, affects the sensitivity of α_t^C to the short rate. When the term premium is small, the α_t^C will be very sensitive to changes in the short rate, making the bank's ξ^C very sensitive to r.

Because α_t^C is affine in r_t, the reduced HJB is quadratic in the state variable. Hence, the solution to the HJB takes the following exponential-quadratic solution,

$$\xi^C(t, r_t) = \exp\Big(\Omega^C(t) + \Lambda^C(t) r_t + \Psi^C(t) r_t^2\Big),$$

with terminal condition

$$\xi^C(T, r) = 1 \qquad \Longrightarrow \qquad \Omega^C(T) = \Lambda^C(T) = \Psi^C(T) = 0.$$

Then, importantly, we have

$$\begin{aligned}
\frac{\xi_t^C}{\xi^C} &= \Omega_t^C + \Lambda_t^C r_t + \Psi_t^C r_t^2, \\
\frac{\xi_r^C}{\xi^C} &= \Lambda^C + 2\Psi^C r_t, \\
\frac{\xi_{rr}^C}{\xi^C} &= 2\Psi^C + (\Lambda^C + 2\Psi^C r_t)^2.
\end{aligned}$$

Substituting into the reduced HJB and collecting terms in powers of r_t gives the following system of ODEs.

First, the coefficient on r_t^2 implies

$$\Psi_t^C - 2\kappa_r \Psi^C + 2\sigma_r^2 (\Psi^C)^2 + 2(1-\rho) b_C \sigma_\ell \sigma_r \Psi^C - \frac{1}{2}\rho(1-\rho) b_C^2 \sigma_\ell^2 = 0.$$

Second, the coefficient on r_t implies

$$\Lambda_t^C + \kappa_r\big(2\bar{r}\,\Psi^C - \Lambda^C\big) + 2\sigma_r^2 \Lambda^C \Psi^C - (1-\rho)\sigma_\ell \sigma_r \big(2a_C \Psi^C - b_C \Lambda^C\big) + \rho(1-\rho) a_C b_C \sigma_\ell^2 = 0.$$

Finally, the constant term implies

$$\Omega_t^C - \omega + (1-\rho)\kappa + \kappa_r \bar{r}\,\Lambda^C + \sigma_r^2 \Psi^C + \frac{1}{2}\sigma_r^2 (\Lambda^C)^2 - (1-\rho) a_C \sigma_\ell \sigma_r \Lambda^C - \frac{1}{2}\rho(1-\rho) a_C^2 \sigma_\ell^2 = 0.$$

The boundary conditions are $\Lambda^C(T) = \Psi^C(T) = \Omega^C(T) = 0$.

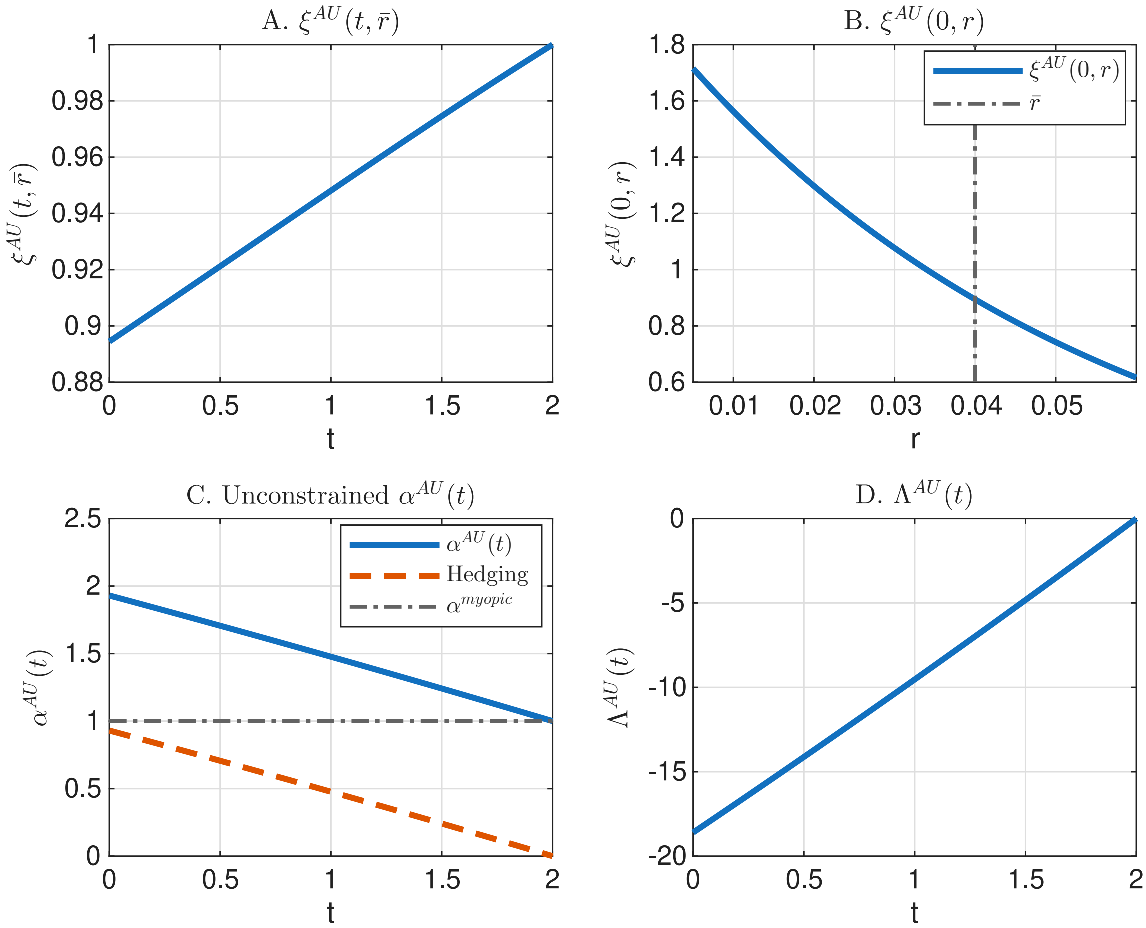
The key element to highlight is that, in the constrained solution, the semi-elasticity $\frac{\xi_r^C}{\xi^C}$ has a different form than in the unconstrained solution. Thus, in the global problem, when the bank is unconstrained but the probability that the constraint binds increases (due to changes in the level of the short rate or under unexpected changes in the term premium), the banker's value function will more closely resemble the constrained value function, so the semi-elasticity will change. As a result, the hedging motives in the unconstrained region will change as the probability of reaching the constraint increases, therefore affecting the optimal α_t in the unconstrained region. Finally, note that a lower term premium will make the semi-elasticity even more negative because α_t^C will be more sensitive to r. $\square$

Solution to the polar cases. Figure 17 shows the solution when the bank is always unconstrained (i.e., the profitability constraint is always slack). Panel A shows ξ^{AU} across time, while Panel B shows ξ^{AU} across r at time 0. Intuitively, ξ^{AU} is the bank's continuation-value index: it measures how valuable wealth is in a given state, taking into account future opportunities and risks. In Panel A, ξ^{AU} increases over time as the bank captures valuable profit opportunities, until it reaches its terminal value of 1. Panel B shows ξ^{AU} at time 0 for different values of the state variable r. Higher rates decrease ξ^{AU} and, therefore, reduce the marginal utility of wealth for a given level of wealth. This is because, at higher rates, deposits are very valuable and, as a result, the bank's marginal utility of wealth is lower. When rates are sufficiently low, however, bankers' marginal utility of wealth is higher than in the case without deposits because we assume $c > (1-\beta)\bar{r}$. Thus, at low rates, deposits represent a significant drag on banks' profits.

Panel C shows the unconstrained α over time and its components—the myopic component, shown in dotted black, and the hedging component, shown in dashed red. The blue line, the total demand for the long-term asset, is the sum of the dotted black and dashed red lines. Intuitively, the hedging demand is positive because banks are risk averse ($\rho > 1$) and dislike states of the world in which rates are low. By holding long-term assets, they can realize gains when rates decline and deposits become a drag. This is exactly the logic of Di Tella and Kurlat (2021) and Drechsler, Savov, and Schnabl (2017). Finally, Panel D shows $\Lambda^{AU}(t)$, which is the always-unconstrained semi-elasticity ξ_r^{AU}/ξ^{AU} and captures the strength of the hedging demand. This object provides a benchmark for the global solution. As the probability of hitting the constraint rises, the global semi-elasticity ξ_r^*/ξ^*, evaluated in the unconstrained region, becomes more negative relative to $\Lambda^{AU}(t)$. We discuss this further in the constrained solution next.

Figure 18 shows the solution in the constrained case. Panel A shows the solution for ξ^C. As opposed to the unconstrained case, ξ^C is, in general, greater than one, reflecting the fact that the banks' marginal utility of wealth is higher due to the fact that its portfolio policy generates a very volatile wealth path. Panel B shows how ξ^C changes for different values of the state variable r_t. Note that banks' continuation value is much more sensitive to changes in the interest rate in the constrained region than in the un-

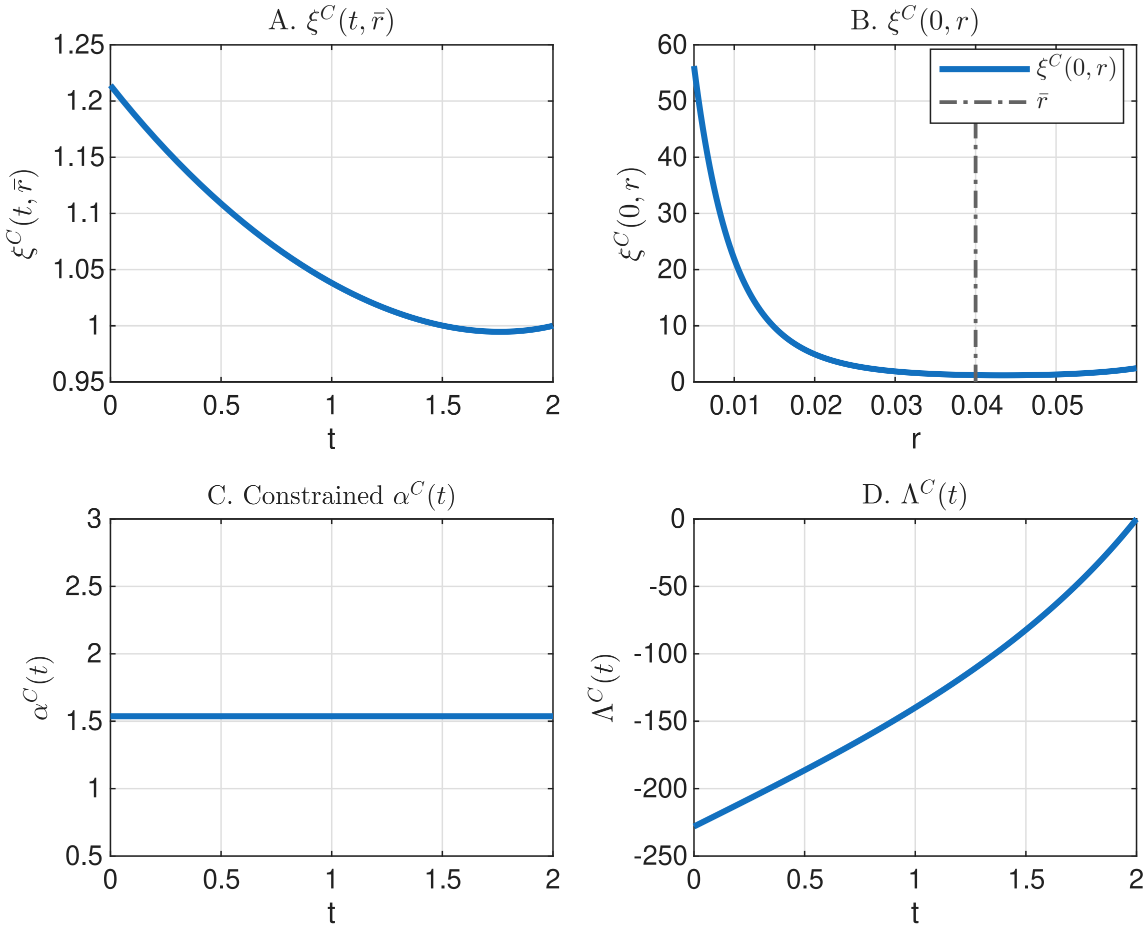
Figure 17: The unconstrained case



constrained (shown in Panel B of Figure 17). Panel C shows the constrained α, which is pinned down by the profitability constraint—which is assumed to be always binding. Panel D shows the intercept of semi-elasticity ξ_r^C/ξ^C, which is, consistent with panel B, notably more negative than in the unconstrained case. Adding the second term of ξ_r^C/ξ^C makes the semi-elasticity even more negative. The fact that banks' continuation value is very sensitive to interest rates in the constrained region is going to be key in the global solution. When unconstrained, the bank will try to anticipate the profitability constraint by adjusting its portfolio ex-ante.

Figure 19 shows the same exercise conducted in the main paper, but in the context of the dynamic model. We focus on the solution at $t = 0$, although any other point in the time dimension will show the same point. The figure shows banks' policy for different levels of the term premium. When the term premium is above approximately 1.6%, the bank is unconstrained and, as a result, its optimal duration exposure follows the unconstrained policy. However, when the term premium declines below 1.6%, the optimal policy is pinned down by the profitability constraint and therefore follows the constrained policy. In the full dynamic model discussed in Section 5, we show that the optimal global policy of the bank is not strictly the envelope of the policies in the polar

Figure 18: The constrained case



cases although it follows qualitatively that shape. Banks will internalize the constraint ex ante and their optimal global policy may follow a similar pattern to the piecewise policy discussed in Figure 19 even though the constraint is not binding.

Finally, Figure 20 shows a similar exercise to Figure 19 but emphasizing the dynamics of the optimal policy α_t. That is, we assume an exogenous "MIT" sequence of τ shocks that cause a gradual decline in the term premium (shown in the left panel). Banks' duration exposure is shown in the right panel. The decline in the term premium causes banks to hit the profitability constraint at approximately $t = 0.4$. From then on, α is pinned down by the profitability constraint, because at such low levels of term premium the bank will touch its constraint.

Map with the static model. In the model in the main text, we defined

$$e = \frac{N_t}{A_t},$$

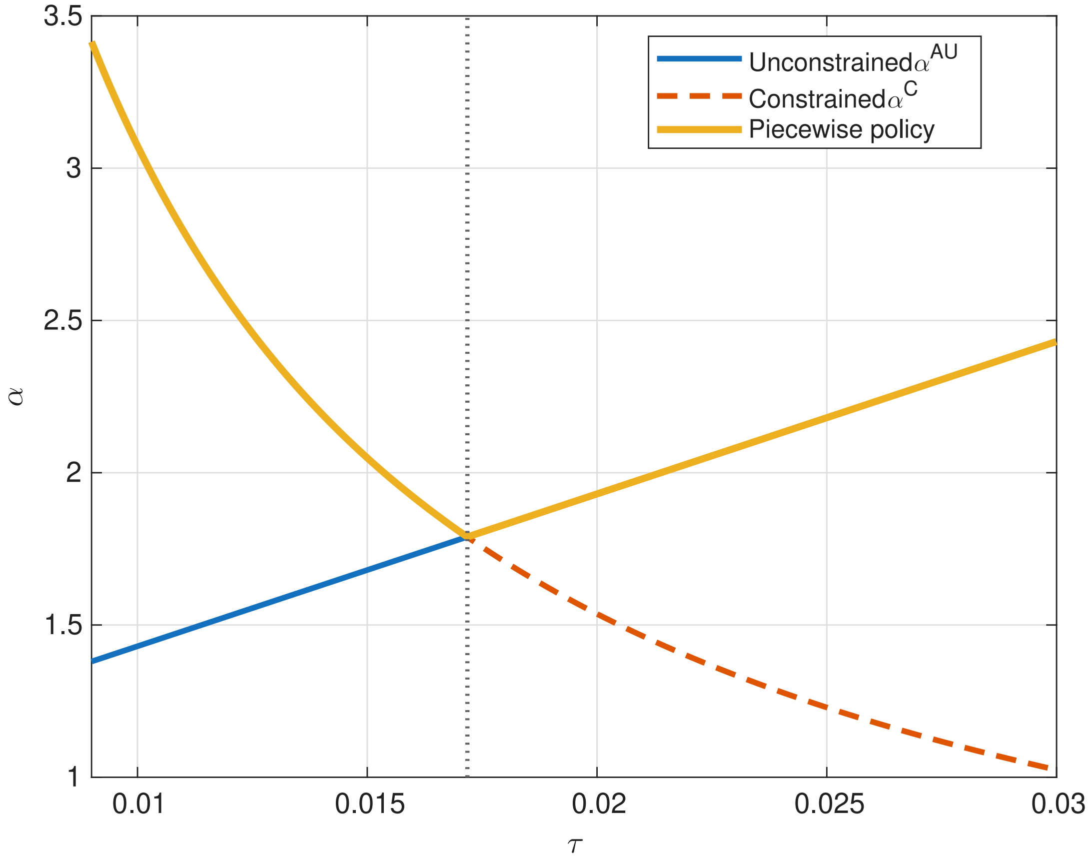


Figure 19: This figure shows banks' piecewise policy (yellow) as well as the solution assuming the constraint is always slack (blue) and assuming it is always binding (in red), for different values of τ

Thus, the deposit-to-wealth ratio, δ, can be written as

$$\delta = \frac{1-e}{e}.$$

Then, we can write

$$\begin{aligned} 1 + (1-\beta)\delta &= \frac{e + (1-e)(1-\beta)}{e} \\ &= \frac{1-(1-e)\beta}{e} \\ &= \frac{\lambda}{e}, \end{aligned}$$

where $\lambda \equiv 1-(1-e)\beta$ as in the two-period model. Thus, the interior optimum in the dynamic model using the notation from the main text is given by

$$\alpha_t = \frac{\tau}{\rho\sigma_\ell^2} - \frac{\lambda}{e}\frac{\sigma_r}{\rho\sigma_\ell}(1-\rho)\frac{1-\exp\left(-\kappa_r(T-t)\right)}{\kappa_r}. \tag{75}$$

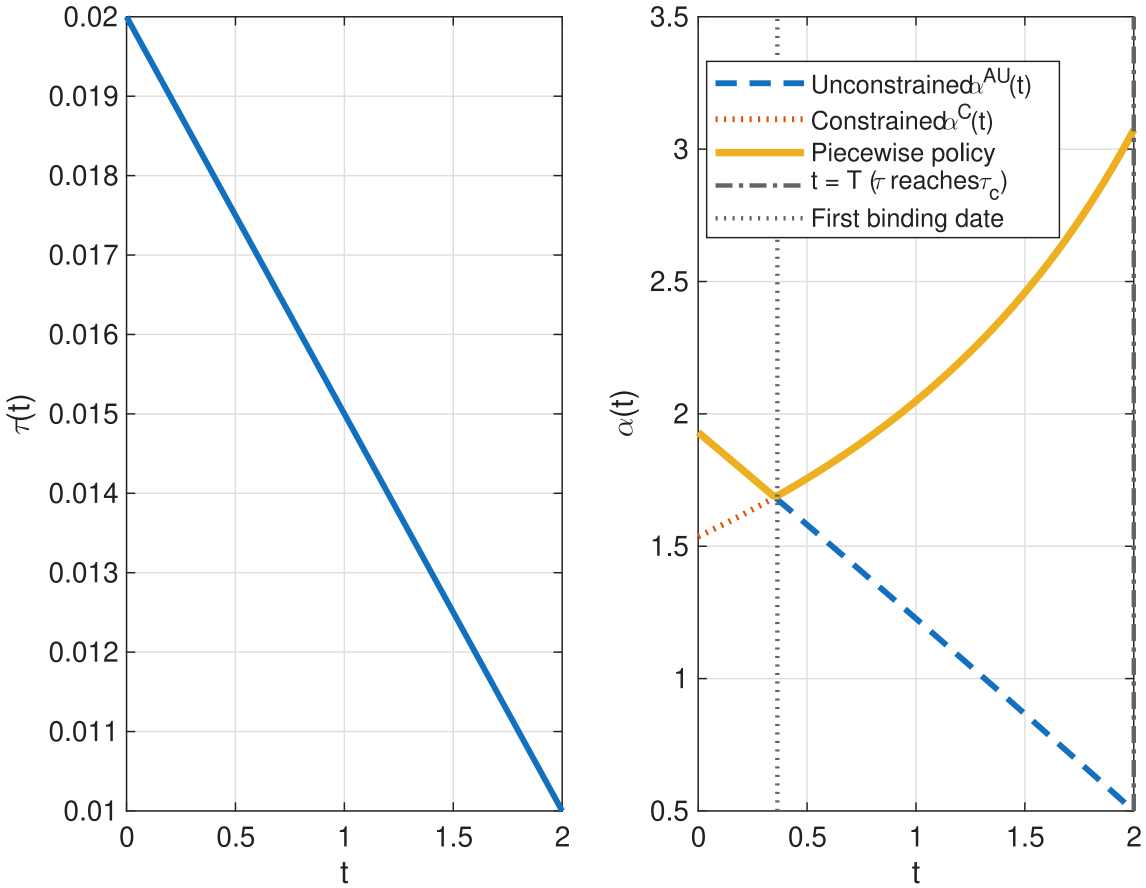


Figure 20: The left panel shows an exogenous decline in term premium from 2% to 1%. The right panel shows the piecewise policy (yellow line) as well as the always unconstrained (blue) and always constrained (red).

Finally, in the two-period model we defined $\alpha = L/A$. Thus,

$$\alpha_t^{\text{two-period}} = \frac{L}{A} = \frac{L/N}{A/N} = \frac{\alpha_t}{1+\delta}.$$

The mapping extends to the remaining parameters. The dynamic profitability floor applies per unit of net worth, $\mathbb{E}_t[\mathrm{d}N_t/N_t] \geq \kappa\,\mathrm{d}t$, while the floor in the two-period model applies per unit of assets, so the two are related by $\kappa^{\text{two-period}} = e\kappa$. Similarly, the cost c applies per unit of deposits in the dynamic model but per unit of assets in the two-period model, so $c^{\text{two-period}} = c\,(1-e)$. With these mappings, the constrained portfolio (74) nests the static solution exactly:

$$\frac{\alpha_t^C}{1+\delta} = e\,\frac{\kappa + c\delta - [1+(1-\beta)\delta]\,r_t}{\tau} = \frac{e\kappa + c\,(1-e) - \lambda r_t}{\tau} = \frac{\kappa^{\text{two-period}} + c^{\text{two-period}} - \lambda r_t}{\tau},$$

which is the minimum required duration α^*_{bind} of the baseline model in equation (8).

The hedging portfolio. A key result presented in the two-period model is that banks can choose a portfolio under which their payoffs would be immune to interest rate shocks. We show how to generalize this result in the dynamic model. We define the hedging portfolio as the one in which the value of the bank has a diffusion term equal to zero. That is, we compute the α such that the optimal value of the bank is immune to rate shocks. From the optimal value function

$$V^*(t,N,r) = \xi^*(t,r)\frac{N^{1-\rho}}{1-\rho}, \tag{76}$$

we apply Itô's lemma to (76), and the diffusion component is

$$\begin{aligned} \sigma_V &= \sigma_\xi + (1-\rho)\,\sigma_N \\ &= \frac{\xi_r^*}{\xi^*}\sigma_r - (1-\rho)\,\alpha\sigma_\ell. \end{aligned}$$

Then, if we make the diffusion component of $\mathrm{d}V/V$ equal to zero (i.e., the value of shareholders' welfare is not affected by interest rate shocks),

$$\alpha^{H,*}(t,r) = \frac{\xi_r^*(t,r)/\xi^*(t,r)}{1-\rho}\frac{\sigma_r}{\sigma_\ell}.$$

In the always-unconstrained case, using the closed-form expression for the semi-elasticity, the hedging portfolio is

$$\begin{aligned} \alpha_t^{H,AU} &= \frac{\frac{1}{\kappa_r}\left[1+(1-\beta)\,\delta\right]\left(1-\exp\left(-\kappa_r(T-t)\right)\right)\sigma_r}{\sigma_\ell} \\ &= \frac{\frac{1}{\kappa_r}\frac{\lambda}{e}\left(1-\exp\left(-\kappa_r(T-t)\right)\right)\sigma_r}{\sigma_\ell}, \end{aligned}$$

where $\lambda = 1-(1-e)\beta$, as in the two-period model. This expression is positive, as expected, given that $\sigma_\ell > 0$. Note that, again, banks consider the deposit leverage ratio, their deposit market power, the persistence of the interest rate, as well as the exposure of the long-term asset to rate shocks, to construct the hedging portfolio. So, we can write the always-unconstrained optimal portfolio as the sum of the myopic component plus the hedging portfolio (which is different from the hedging component),

$$\alpha_t^{AU} = \frac{\tau}{\rho\sigma_\ell^2} + \frac{\rho-1}{\rho}\alpha_t^{H,AU}.$$

Numerical Solution of the Global HJB. We solve the global finite-horizon HJB after applying the transformation $q(t,r) \equiv \log\xi(t,r)$. We discretize time and the short-rate state on uniform grids and solve backward from the terminal condition $q(T,r) = 0$. At each time step, we use an implicit Euler discretization,

$$\frac{q(t+\Delta t,r)-q(t,r)}{\Delta t}+\mathcal{H}(r,q(t,r),q_r(t,r),q_{rr}(t,r))=0,$$

and solve the resulting nonlinear system jointly across all short-rate grid points using a nonlinear root finder. Spatial derivatives are approximated using second-order centered finite differences at interior grid points and second-order one-sided differences at the boundaries.

At every iteration, the unconstrained candidate is computed using the endogenous derivative of the global continuation value,

$$\alpha^U(t,r)=\frac{\tau}{\rho\sigma_\ell^2}-\frac{\sigma_r}{\rho\sigma_\ell}q_r(t,r),$$

while the portfolio required by the profitability constraint is

$$\alpha^C(r)=\frac{\kappa+c\delta-[1+\delta(1-\beta)]\,r}{\tau}.$$

The global policy is therefore $\alpha^*(t,r)=\max\left\{\alpha^U(t,r),\alpha^C(r)\right\}$.

Additional results from the quantitative analysis. In Figure 9, we showed the model-implied maturity gap for two alternative values of the short rate and the term premium. In Figure 21, we compute the maturity gap assuming the bank is always unconstrained (AU), the maturity gap implied by the profitability constraint, and the maturity gap implied by the global solution (α^*).

The key takeaway is that, in the dynamic model, the increase in the maturity gap caused by the decrease in the term premium and the short rate is driven entirely by the bank's concerns about profitability. After the term premium and the short rate decline, the bank is not constrained, but its investment opportunity set is very sensitive to changes in the short rate. As a result, the hedging motives, driven by the semi-elasticity ξ_r/ξ, are strong enough to increase the bank's demand for long-term assets.

Another plausible hypothesis for why banks may have increased their maturity gap over the last 40 years it that the lower level of rates caused a switch in household demands toward saving deposits, which ultimately translates into a higher deposit market power (Supera, 2021). As a result of higher deposit market power, banks extended the duration of their assets to match the interest rate sensitivity of assets with liabilities, as predicted by Drechsler et al. (2021). We study this hypothesis through the lens of our model.

Figure 22 shows the same model-implied maturity gaps as in the previous figures, in blue, and compares them with the model-implied maturity gaps when the deposit

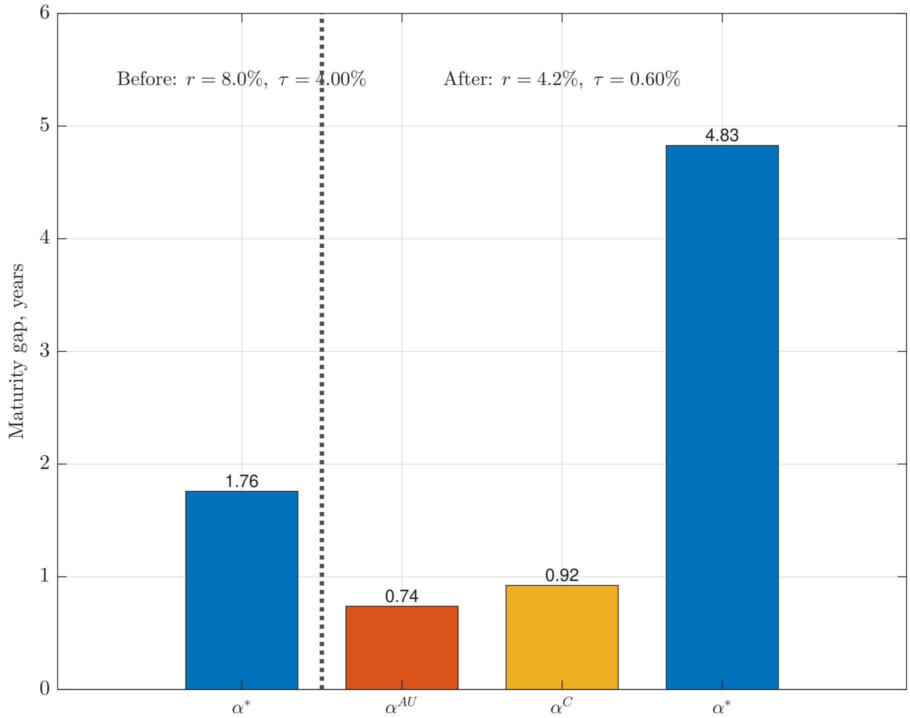


Figure 21: The figure shows the model-implied maturity gap for when the short rate is at 8% (4.2%) and the term premium is set at 4% (0.6%), as in 1985 (2025). See Section 5 for more details.

beta is 0.22 instead of 0.44, in red. Reducing the deposit beta by half has two implications for banks' demand for duration, and these forces work in opposite directions. On the one hand, a lower deposit beta improves banks' profits, thereby alleviating banks' concerns about profitability. This force pushes demand for duration down. On the other hand, a lower deposit beta increases banks' hedging motives, which pushes demand for duration up.

When the term premium and the level of rates are high, the stronger hedging motive increases the maturity gap somewhat. This can be seen by comparing the first bar with the second bar in the figure. However, when both the level of rates and the term premium decline, higher deposit market power decreases the maturity gap. This can be seen by comparing the third bar, in blue, with the fourth bar, in red. The reason is that, at low levels of rates and term premia, profitability concerns are strong, and the effect of greater deposit market power in alleviating profitability concerns dominates the stronger hedging motive.

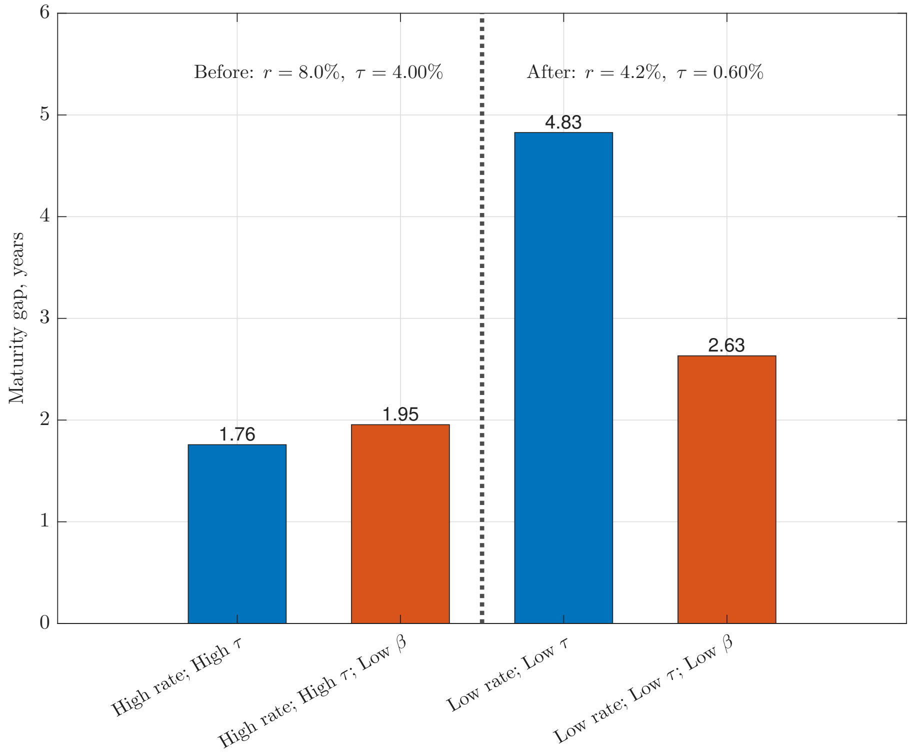


Figure 22: The figure shows the model-implied maturity gap for when the short rate is at 8% (4.2%) and the term premium is set at 4% (0.6%), as in 1985 (2025). The red bars show the results when the deposit beta is cut in half from the baseline calibration. See Section 5 for more details.

I Additional Bank-Level Evidence

I.1 Variable Definitions

Table A.12 lists the Call Report (MDRM) codes used to construct the non-maturity variables in the bank-level analysis of Section 6. Balance-sheet items are from Schedule RC and the regulatory-capital and deposit schedules; income-statement items are from Schedule RI, which is reported year-to-date and which we convert to single-quarter flows. Balance-sheet items are reported on a domestic-office basis (RCON) by domestic filers; for banks with foreign offices, which report on a consolidated basis (RCFD), we scale the consolidated values to a domestic-office basis using the ratio of domestic to consolidated total assets. For income-statement items, we adjust for foreign-office activity using the additional codes noted in the table. The maturity and repricing variables that enter the maturity gap, including cash and federal funds, are documented in Tables A.6 and A.10.

Table A.12: Non-Maturity Call Report Variables Used in Section 6

Variable	Code	Description
Panel A: Balance-Sheet Variables (Schedule RC)		
Total assets	RCON/RCFD 2170	Total assets.
Total liabilities	RCON/RCFD 2950	Total liabilities.
Total deposits	RCON/RCFD 2200	Total deposits.
Total loans and leases	RCON/RCFD 2122	Total loans and leases.
Securities, held-to-maturity	RCON/RCFD 1754	Held-to-maturity securities (amortized cost).
Securities, available-for-sale	RCON/RCFD 1773	Available-for-sale securities (fair value).
Panel B: Income-Statement Variables (Schedule RI)		
Net income	RIAD 4340	Less foreign-office adjustment RIAD 4341 (through 2005) or RIAD C914 (from 2006).
Net interest income	RIAD 4074	Foreign-office adjusted: less RIAD 4842 (through 2000) or RIAD B525 (2001–2005); from 2006, less RIAD C899 plus RIAD C900.
Noninterest income	RIAD 4079	Less foreign-office adjustment RIAD 4097 (through 2005); from 2006, less RIAD C902 through RIAD C905.
Noninterest expense	RIAD 4093	Less foreign-office adjustment RIAD 4239 (through 2005) or RIAD C907 (from 2006).
Total interest income	RIAD 4107	Total interest income.
Total interest expense	RIAD 4073	Total interest expense.
Interest income, loans	RIAD 4010	Total interest and fee income on loans and leases.
Interest income, securities	RIAD B488	From 2001, RIAD B488 (Treasury/agency) plus RIAD B489 (MBS) plus RIAD 4060 (equity dividends); through 2000, RIAD 4027 plus RIAD 4506 plus RIAD 4507 plus RIAD 3657 plus RIAD 3658 plus RIAD 3659.

Notes: This table lists the Call Report (FFIEC 031–034 and 041) variable codes used to construct the non-maturity variables in the Section 6 bank-level analysis. Income-statement items (Schedule RI) are reported year-to-date and converted to single-quarter flows. "Through" and "from" dates indicate the reporting periods over which each code applies. For banks with foreign offices, balance-sheet items use the consolidated RCFD value, scaled to a domestic-office basis by the ratio of domestic to consolidated total assets. The maturity and repricing variables underlying the maturity gap are listed in Tables A.6 and A.10.

I.2 Dynamic Response of the Maturity Gap

To trace out the dynamic response of the maturity gap to bank profitability, we estimate horizon-by-horizon local projections of the form:

$$y_{i,t+h} = \phi_t^h + \beta^h Z_{i,t-1} + \theta_1^h y_{i,t-1} + \theta_2^h Z_{i,t-2} + u_{i,t}^h, \quad h = 0, 1, 2, \ldots, 10 \qquad (77)$$

where $y_{i,t+h}$ is the outcome at horizon h ahead, ϕ_t^h are time fixed effects, and we control for one lag of the dependent variable and a further lag of the profitability indicator. To account for asset renewal over horizons longer than one quarter, we rescale the maturity gap at horizon h by $(2M-1)/(1+h)$, where M is the average maturity of bank assets over the sample (in quarters), under the same asset-side renewal assumptions as in Section 6. Standard errors are two-way clustered by bank and time.

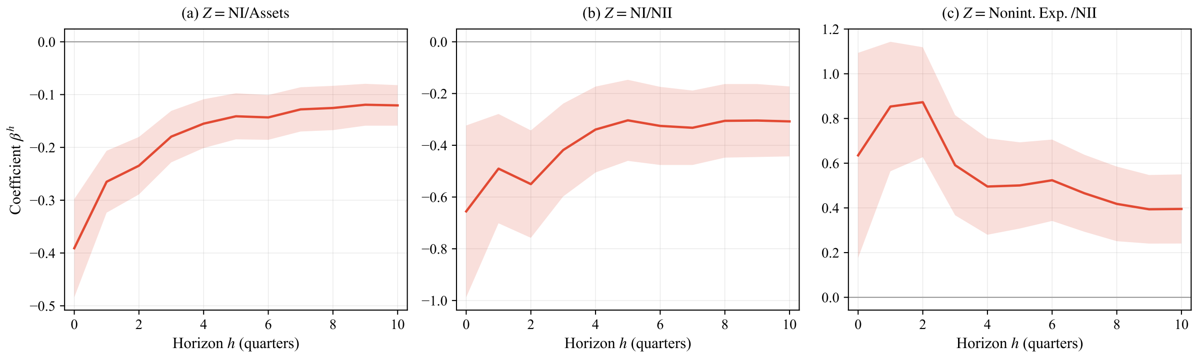


Figure 23: Dynamic response of the maturity gap. Estimated coefficients β^h from local projection (77) for $h = 0, \ldots, 10$, with 95 percent confidence bands. Panels (a)–(c) use Z = NetIncome/Assets, Z = NetIncome/NII, and Z = NoninterestExpenses/NII, respectively. Standard errors are two-way clustered by bank and time.

Figure 23 reports the estimated coefficients β^h for the three profitability measures. Consistent with the static results in Table 2, the coefficients on Z = NetIncome/Assets (panel a) and Z = NetIncome/NII (panel b) are negative on impact and remain negative over the entire impulse response horizon, leveling off over time: a decline in profitability predicts a persistent extension of the maturity gap. The coefficient on Z = NoninterestExpenses/NII (panel c), for which an increase represents greater profitability pressure, is positive throughout, peaking after roughly two quarters, and slowly declining thereafter. Overall, these estimates show persistent effects even when we allow for a growing fraction of assets to be renewed.

I.3 Decomposition of the Maturity Response

To pinpoint where the maturity gap response originates, we decompose it into its components. Table A.13 reports estimates of regression (77) at $h = 0$ and for Z =

NetIncome/Assets, replacing the dependent variable in turn by the maturity of total assets, the maturity of total liabilities, the maturity of securities, and the maturity of loans. Each component is rescaled by $(2M - 1)$, where M is the average maturity of the corresponding component (in quarters); for the maturity gap in column (i), M is the average maturity of total assets, as in Section 6. Columns (i)–(iii) are estimated on a common sample—the bank-quarters for which the maturity gap is available—so that the asset- and liability-maturity responses decompose the maturity gap response; columns (iv) and (v) use all observations for which the respective maturity is available.

Table A.13: Decomposition of the Maturity Response.

	(i) Maturity Gap	(ii) Mat.Assets	(iii) Mat.Liab.	(iv) Mat.Sec.	(v) Mat.Loans
NI/Assets	-0.39***	-0.39***	-0.00	-0.34***	-0.19***
	(0.05)	(0.05)	(0.00)	(0.12)	(0.04)
Time FE	✓	✓	✓	✓	✓
Partial R^2	0.04	0.04	0.03	0.04	0.02
Observations	516,792	516,792	516,792	687,538	717,054
Number of Banks	10,989	10,989	10,989	11,696	11,873

Notes: Estimation results for regression (77) at $h = 0$ for the baseline profitability measure $Z =$ NetIncome/Assets. The dependent variable in column (i) is the maturity gap; in columns (ii)–(v), the dependent variable is the maturity of total assets, total liabilities, securities, and loans, respectively. Each dependent variable is rescaled by $(2M - 1)$, where M is the average maturity of the corresponding component in the sample (in quarters); for the maturity gap in column (i), M is the average maturity of total assets. Columns (i)–(iii) are estimated on a common sample, restricted to bank-quarters for which the maturity gap is available. The partial R^2 conditions on the variation explained by the lagged dependent variable. Standard errors in parentheses are two-way clustered by bank and time. Sample: 1997:Q2–2026:Q1. Notation: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Three patterns stand out. First, the entire response of the maturity gap is concentrated on the asset side: the coefficient on the maturity of liabilities is zero to two decimal places and statistically insignificant, while the coefficient on the maturity of assets (-0.39) coincides with the coefficient on the maturity gap itself (-0.39). Second, both securities and loans contribute to the asset response, with coefficients of -0.34 and -0.19, respectively, both statistically significant at the 1 percent level. Third, the securities response is roughly twice as large as the loan response. The credit-risk extension of the model rationalizes this pattern (Appendix E): when the loan portfolio is constrained, adjustments concentrate on the more flexible margin.

Table A.14 examines the income and balance sheet consequences of the maturity extension. We use the same specification but replace the dependent variable in turn by the average interest income on securities, the average interest on loans, the bank's overall net interest margin (each normalized in the same way as the corresponding maturity measures), and the share of securities and loans in total assets. The first three columns deliver uniformly negative coefficients: when profitability declines, banks subsequently earn higher interest on securities and loans, and a higher overall net interest margin. Consistent with the mechanism, banks extend the duration of their long-maturity assets to capture additional interest income, and the additional income manifests in their net interest margin. Columns (iv) and (v) further document that banks tilt their balance sheets toward securities and loans when profitability declines, although the magnitudes are relatively small. The dominant adjustment margin in our data is therefore the maturity of existing asset classes rather than the composition of the balance sheet across asset classes.

Table A.14: Income and Asset-Composition Responses.

	(i) Int.Sec.	(ii) Int.Loans	(iii) NIM	(iv) Sec./Assets	(v) Loans/Assets
NI/Assets	-0.53*	-0.70***	-0.54***	-0.02**	-0.04**
	(0.28)	(0.10)	(0.08)	(0.01)	(0.02)
Time FE	✓	✓	✓	✓	✓
Partial R^2	0.28	0.17	0.1	0.03	0.06
Observations	699,625	708,519	712,110	725,254	725,254
Number of Banks	11,825	11,809	11,870	11,946	11,946

Notes: Estimation results for regression (77) at $h = 0$ for the baseline profitability measure Z = NetIncome/Assets. The dependent variable in columns (i)–(iii) is the bank's average interest income on securities, average interest income on loans, and overall net interest margin, respectively, each rescaled by $(2M - 1)$ as in Table A.13, with M the average maturity of securities, loans, and total assets, respectively (in quarters). The dependent variable in columns (iv) and (v) is the share of securities and loans in total assets. The partial R^2 conditions on the variation explained by the lagged dependent variable. Standard errors in parentheses are two-way clustered by bank and by time. Sample: 1997:Q2–2026:Q1. Notation: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

I.4 Interaction with the Term Premium and Federal Reserve QE

As a final test of the mechanism, we ask whether the reach-for-duration response is amplified in periods of low term premia. While our baseline regression (27) already accounts for the level of the term premium through bank net income, we explore whether the documented patterns intensify or weaken depending on the level of term premia.

To this end, we estimate an interaction regression of the form:

$$y_{i,t+h} = \phi_t^h + \beta_1^h Z_{i,t-1} + \beta_2^h Z_{i,t-1} \cdot S_{t-1} + \theta_1^h y_{i,t-1} + \theta_2^h Z_{i,t-2} + u_{i,t}^h \quad \text{for} \quad h = 0, 1, \ldots, 10, \tag{78}$$

where S_{t-1} is an aggregate state variable, normalized so that zero corresponds to its sample minimum. Coefficient β_1^h then captures the response of the maturity gap to a change in profitability Z when S is at its sample minimum, and β_2^h captures how this response varies as S rises. We again use the baseline measure $Z = \text{NetIncome}/\text{Assets}$.

We consider two state variables. The first is the term premium TP_{t-1}, for which we use the 10-year estimate from Adrian, Crump, and Moench (2013). The second is the Federal Reserve's holdings of Treasury and agency securities relative to GDP, $\text{SOMA}_{t-1}/\text{GDP}_{t-1}$, as an indicator of quantitative easing. Figures 24 and 25 report the two sets of coefficient sequences.

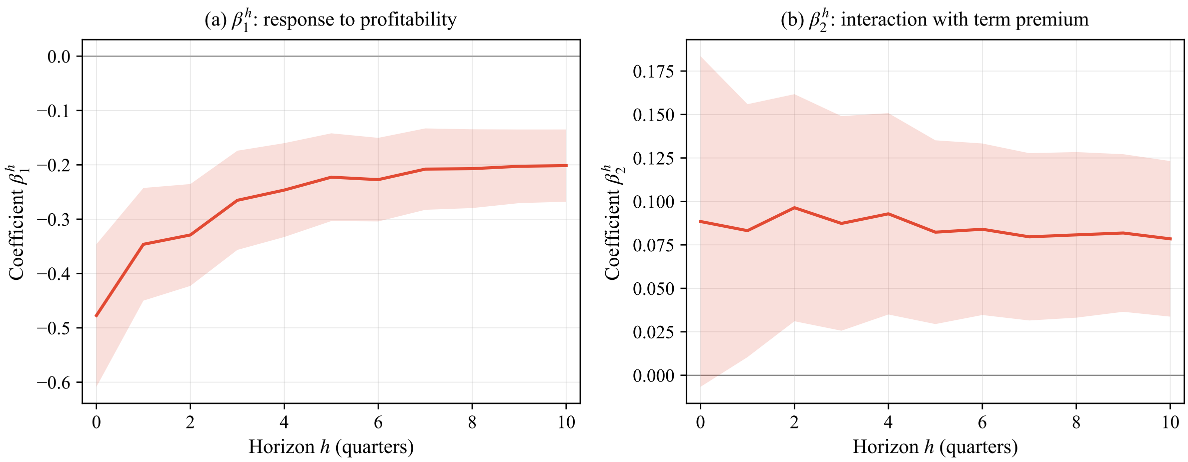


Figure 24: Dynamic response of the maturity gap, interacted with the term premium. Estimated coefficients β_1^h (left panel) and β_2^h (right panel) from local projection (78) with $S_{t-1} = TP_{t-1}$, for $h = 0, \ldots, 10$, with 95 percent confidence bands. The term premium is the 10-year estimate from Adrian, Crump, and Moench (2013), normalized so that zero corresponds to its sample minimum. Standard errors are two-way clustered by bank and time.

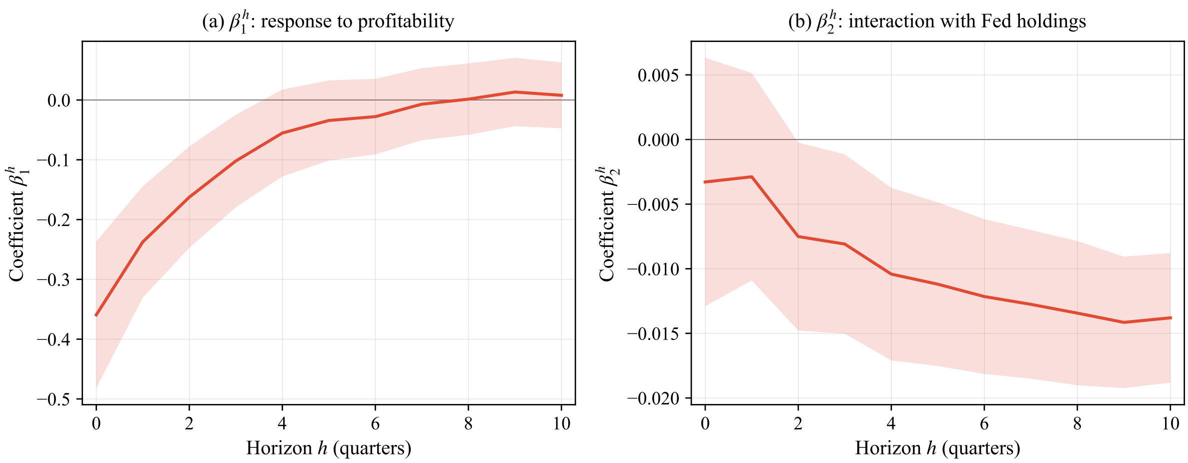


Figure 25: Dynamic response of the maturity gap, interacted with Federal Reserve holdings. Estimated coefficients β_1^h (left panel) and β_2^h (right panel) from local projection (78) with $S_{t-1} = \text{SOMA}_{t-1}/\text{GDP}_{t-1}$, for $h = 0, \ldots, 10$, with 95 percent confidence bands. $\text{SOMA}_{t-1}/\text{GDP}_{t-1}$ is the Federal Reserve's holdings of Treasury and agency securities as a share of GDP, normalized so that zero corresponds to its sample minimum. Standard errors are two-way clustered by bank and time.

For the term premium (Figure 24), the coefficient on profitability is around -0.48 on impact when the term premium is at its sample minimum, declining in magnitude toward -0.2 at longer horizons—slightly larger in magnitude than the unconditional coefficient in Figure 23; the interaction coefficient β_2^h is positive and stable around 0.08, so the strength of the reach-for-duration response attenuates as the term premium rises. These results are not specific to the term premium estimate we employ: we obtain similar results with the 2-year or 5-year estimate from Adrian, Crump, and Moench (2013) or with estimates based on Kim and Wright (2005).

For the Federal Reserve's holdings (Figure 25), the coefficient on profitability is around -0.36 on impact at the smallest, pre-QE balance sheet and decays toward zero within roughly two years. The interaction coefficient is negative and grows in magnitude with the horizon, from near zero on impact to about -0.014 after ten quarters—so the reach-for-duration response both strengthens and becomes more persistent as the central bank's holdings expand. Both findings are consistent with our mechanism and show that the reach-for-duration effect strengthens when term premia are compressed, which may be the result of quantitative easing.