A New Keynesian Preferred Habitat Model with Repo

Qian Wu*

September 3, 2024

Abstract

This paper develops a New Keynesian Preferred Habitat model with repo assets to account for financial and macroeconomic observations during the Global Financial Crisis (GFC) and the Covid-19 pandemic. The model features market segmentation, financial frictions, and quality preference. I show that there was a flight-to-liquidity demand for short-term Treasuries which causes endogenous convenience yields during the GFC and a flight-from-safety supply for long-term Treasuries which causes endogenous inconvenience yields during Covid-19. I then use the model to study the passthrough of monetary policies to asset prices and macroeconomic variables. The model equilibrium yields three key findings. Firstly, the excess return in the Treasury cash market involves risk premia and (in)convenience premia, while the excess return in the repo market includes only (in)convenience premia. Secondly, financial frictions attenuate the passthrough of conventional policy while strengthening that of QE and QT. Lastly, the efficacy of monetary policies is contingent upon the relative importance of the repo borrowing channel compared to the cash borrowing channel. These findings contribute to a deeper understanding of the monetary policy transmission mechanisms in the post-GFC era.

Keywords: monetary policy, quantitative easing, financial frictions, benchmark rate reform
JEL Codes: E52, E43, E44

*Indiana University, qwu2@iu.edu.

1 Introduction

In the two most recent recessions—the GFC and the Covid-19 pandemic—central banks employed various policy tools to mitigate the potential worsening of economic conditions. Following aggressive rate cuts, the zero lower bound constraint compelled monetary authorities to adopt unconventional tools such as Quantitative Easing (QE) and forward guidance, aiming to boost output and inflation by lowering market rates, particularly for longer tenors. Empirical evidence has shown that QE has significant and persistent effects, whereas forward guidance appears to be less effective in practice.

The New Keynesian model has served as a baseline for studying monetary policies to stablize the financial markets and the macroeconomy. Over the past decades, the evolving New Keynesian literature has made significant progress in modeling the non-neutral effects of unconventional policies such as QE and forward guidance [1]. Additionally, spurred by the financial market turmoil during the GFC, recent research has focused on extending the basic New Keynesian framework to incorporate financial frictions (for example, Christiano, Motto and Rostagno (2014); Guerrieri and Lorenzoni (2017)). These works share a key feature: the (occasionally) binding borrowing constraint. During economic distress, the borrower's net worth shrinks sharply, leading to limited credit flows and spiking interest rates. Consequently, output and price levels decline. This line of research has effectively explained many observations during the GFC, such as the concurrence of higher financing costs and lower net worth among financial intermediaries.

In this paper, I first present empirical observations indicating differing financial market activities during the GFC and the Covid-19 pandemic. I focus on two important financial markets that are both considered deep and safe: the Treasury cash and repo markets. I show that (1) although the term structure of the Treasury yield steepened in both periods, the GFC witnessed a decrease in yields for all maturities, whereas the Covid-19 pandemic saw an increase in long-term yields; (2) the Treasury-OIS spread was negative during the GFC but positive during Covid-19; (3) the repo-short rate spread dropped during the GFC but increased during Covid-19; and (4) primary dealers were reducing their net reverse repo positions during the GFC but expanding them during Covid-19. These observations cannot be fully explained by the current New Keynesian literature with financial frictions, which suggests that a recession should always be accompanied by spiking interest rates. These

[1] For example, Sims, Wu and Zhang (2023) introduce a leverage constraint to allow government purchase of long-term bond to relax financial institution's leverage constraint and thus decrease the interest rate. Angeletos and Lian (2018) rely on the lack of common knowledge to conger the forward guidance puzzle.

empirical findings prompt a reevaluation of the causes of financial market turmoil and their relationships to real economic activities during these two recessions.

I then propose a New Keynesian model with rich financial markets. Specifically, I model the Treasury bond and repo markets featuring market segmentation, limited risk-bearing capacity, and quality value of holding Treasuries. There are four key departures from the benchmark New Keynesian model. First, the market segmentation introduced by habitat investors allows demand and supply factors to have non-neutral effects. Second, arbitrageurs mitigate market segmentation by trading across different products; however, this trade is subject to costs, resulting in only a partial mitigation and causing deviations of market rates from the policy short rate. Third, as households invest in saving products indexed by various rates, the aggregate nominal rate becomes a function of the Treasury yields and repo rates, differing from the policy rate. This results in an imperfect mapping from the policy rate to market rates. Fourth, households receive extra utility from holding savings due to quality preferences, which allows for the concurrence of lower market rates and declining real economic activities.

The model equilibrium generates interesting insights. The time-varying excess return of Treasury cash investments encompasses two components: the risk premium and the (in)convenience premium. The former results from risk-averse arbitrageurs, while the latter arises from the non-pecuniary holding benefits (or costs) of Treasuries. Conversely, the time-varying excess return of Treasury repo investments encompasses just one component: the (in)convenience premium. Repo cash lenders use Treasuries as collateral but do not own the securities. Therefore, they are not directly exposed to the risk factors impacting bond prices, and their first-order conditions do not reflect the risk prices.

I show that the model can account for observations during the GFC and Covid-19. The changes in private habitat investors' net demand for Treasuries cause endogenous (in)convenience yields which explain the discrepancies in key financial market variables. During the GFC, there was a significant net demand shock for short-term Treasuries, leading to a decrease in short-term yields and a steepening of the Treasury yield curve. This surge in demand generated net convenience yields in both the Treasury cash and repo markets, as holding Treasuries enabled dealers to avoid costs associated with meeting the heightened market demand. The increased demand for Treasuries among private habitat investors also triggered a quality preference among households, enhancing the marginal benefit of saving. As a result, households did not increase consumption despite lower interest rates, which led to simultaneous drops in output and inflation along with declining interest rates. In contrast, during the Covid-19 period, the market experienced significant selling pressure for long-term Treasuries, causing an increase in long-term yields and a steepening of the yield curve. This excess supply created net inconvenience yields in the Treasury cash and repo markets, as dealers faced higher costs associated with expanding their balance sheets to absorb the additional supply. The resulting higher market rates made saving more attractive, discouraging aggregate consumption.

After calibrating the model to match key financial and macroeconomic moments, I use it to explore monetary policy implications. I show that market imperfections hinder the transmission of conventional expansionary policies. In a purely segmented economy without arbitrageurs, all bond supply is absorbed by habitat investors, disconnecting bond yields and repo rates from the short rate, rendering conventional policy ineffective. Allowing arbitrageurs to participate in the market alleviates this disconnection, as they link products through carry trades. However, when arbitrageurs are risk-averse or face non-pecuniary cost, these carry trades are imperfect, only partially overcoming market segmentation. In this scenario, conventional policy leads to under-reactions in asset prices and macro variables. Removing risk aversion and non-pecuniary cost allows for a perfect overcoming of market segmentation, enabling conventional expansion to achieve perfect transmission, as predicted by the Expectations Hypothesis. The model does not exhibit the forward guidance puzzle. Based on the calibration, forward guidance generates minimal responses in asset prices and macroeconomic variables compared to conventional expansion.

On the other hand, the effectiveness of QE increases with market imperfections. In the absence of market imperfections, investors are indifferent to different products, and any demand or supply shocks are smoothed out, so QE does not affect asset prices or macroeconomic variables. Financial frictions, such as risk aversion and non-pecuniary cost, create the conditions necessary for QE to have an impact. The additional demand from the central bank alleviates the risk exposure of private investors and reduces their balance sheet size, leading to a decrease in required excess returns. In the most frictional case, where no arbitrageurs are present, asset prices are determined solely by supply and demand for each product, and QE achieves its greatest effects. Additionally, the magnitude of the QE shock depends on the targeted maturity. Longer-term bond prices are more sensitive to demand factors than to short rates, so purchasing long-term bonds helps to offload more risk from private investors. Thus, the efficiency of QE increases with the targeted maturity.

I then use the model to examine how the benchmark rate reform affects the passthrough of monetary policies. I define the yield regime and repo regime as scenarios where the aggregated nominal rate is indexed solely by bond yields or repo rates, respectively. The passthrough of a conventional expansion achieves larger responses in asset prices and macroeconomic variables in the repo regime compared to the yield regime. Conversely, the impact of QE on boosting the economy is weakened in the repo regime. More interestingly, the forward guidance, which is stimulative in an yield economy, surprisingly becomes counterproductive in the repo economy. Since the short rate is transmitted to the economy more efficiently in the repo regime, the central bank can adopt a less aggressive stance to achieve a zero output gap in the long run. Consequently, the optimal long-run policy rate target is lower in the repo regime.

This paper contributes to the literature in several important ways. Primarily, it addresses the shortcomings of current New Keynesian models in capturing the dynamics of financial markets during periods of turmoil, which significantly impact macroeconomic variables. By integrating a detailed financial sector, the model adeptly captures the pivotal role of market frictions in shaping the divergent dynamics of Treasury cash and repo markets. This framework uniquely facilitates the examination of both conventional and unconventional monetary policies across different institutional contexts, pre- and post-GFC.

Moreover, this paper explores the potential impacts of the ongoing transition in the benchmark interest rate. The important ongoing transition from LIBOR to SOFR in the U.S. has raised a demand to better understand connections between repo rates and other financial and economic variables. Policymakers need to be alerted about the potential new transmission mechanisms to better operate policy tools; Market participants need to be aware of these new mechanisms to be more prepared for market reactions. The existing literature provides limited insights into these mechanisms, whereas this study offers a robust framework to analyze such transitions, considering the intricate interactions between the financial sector and the broader economy.

Methodologically, by incorporating repo assets into the New Keynesian Preferred Habitat framework, this work introduces a layer of complexity to general equilibrium modeling. It extends the models by Vayanos and Vila (2021) and Ray (2019) through a generalized numerical method that accommodates these new features.

I focus on preferred habitat as the mechanism for short rate and demand/supply factors to affect asset prices, aligning with works such as Vayanos and Vila (2021), Greenwood and Vayanos (2014), and He, Nagel and Song (2022). These papers emphasize the role of arbitrageurs' carry trades in partly overcoming the market segmentation introduced by habitat investors, allowing both short rate and demand/supply factor to be impactful. By integrating macroeconomic dynamics, this paper extends the preferred habitat literature, introducing interactions between financial markets and the broader economy. It shares similarities with Ray (2019) and Ray and Kamdar (2024). Yet this paper introduces a quality preference channel with a similar flavor as recent works such as Bodenstein et al. (2023) and Kekre and Lenel (2024). Because the marginal quality value is connected to private demand factor, this channel capture the real effect of flight-to-quality activities prevalent during economic distresses. The general equilibrium of this model can be described as a financial market-accommodated monetary policy rule, or vice versa, a monetary policy-consistent financial market optimization, in the similar spirit to Gallmeyer, Hollifield and Zin (2005) and Gallmeyer et al. (2007).

This work is also related to a group of papers investigating how the benchmark rate reform affects the broader economy. For example, Jermann (2019) and Cooperman et al. (2023) both show that the new indexing regime - SOFR regime - can exacerbate the negative funding cost shock, since the SOFR is a risk-free rate and does not reflect banks' financing cost. This paper, however, studies the monetary policy implications of such a reform. I utilize frictions on the Treasury cash and repo market (e.g., the scarcity premium as in d'Amico, Fan and Kitsul (2018); Arrata et al. (2020); Corradin and Maddaloni (2020), and the regulatory cost as in Duffie (2018); Cochran et al. (2023)) for an imperfect mapping from short rate to bond yields and repo rates [2]. In equilibrium, the power of the monetary policies depends on the relative importance of the indexing rates.

The rest of the paper is organized as follows. Section 2 outlines the institutional background for the Treasury cash and repo markets and their behaviors during the GFC and Covid-19. Section 3 introduces the theoretical model, with the general equilibrium explained in Section 4. Section 5 uses the model to explain observations during the two recent recessions. Section 6 presents the empirical methodology and data used to calibrate the model. In Section 7, I evaluate the model's performance by replicating key results from the literature. An analysis of the shock transmission mechanisms is included in Section 8. Section 9 presents policy analysis and Section 10 concludes the paper.

[2] Imperfect competition, not directly captured by my model, is also a well documented friction on the market. See Huber (2023) and Eisenschmidt, Ma and Zhang (2024) for examples

2 Institutional Background

In this section, I first outline the features of the Treasury cash and repo markets, focusing on the various types of frictions related to demand and supply factors that lead to convenience and inconvenience yields. I then discuss market performance during the 2008 GFC and the Covid-19 pandemic.

2.1 Treasury Cash Market

US Treasuries have been a cornerstone of premier safe assets globally for several decades. As of June 2024, the total outstanding Treasuries held by the public reached $27 trillion[3]. The US Treasury market features a deep and broad secondary cash market, with daily trading volumes exceeding $884 billion as of June 2024 [4]. Approximately 70% of this trading volume is concentrated in on-the-run securities—those most recently auctioned within a given tenor—while the remaining 30% involves off-the-run securities, which include all previously issued securities (He, Nagel and Song (2022)).

Broker-dealers play crucial roles in both the primary and secondary markets for US Treasuries. Specifically, primary dealers are expected to submit bids for all issuance auctions at reasonable prices [5]. Additionally, broker-dealers are key participants in the Treasury secondary cash market. Figure 1 summarizes the main components of the Treasury secondary cash market. In the Dealer-To-Client venue, dealers act as market makers, transacting with a variety of end-investors. In the Dealer-To-Dealer venue, dealers trade directly with each other. In the Interdealer-Broker venue, dealers trade anonymously with each other, with end-investors such as hedge funds, and with Principal Trading Firms, primarily on electronic platforms (Harkrader and Weitz (2020)). Notably, dealers are active in all three main segments of the secondary market, trading with clients, principal trading firms, and other broker-dealers. According to calculations by the New York Fed, the Treasury cash market trading volume is roughly split between the interdealer brokers and dealer-to-client sectors, with minimal dealer-to-dealer volume.

US Treasuries, particularly those with short tenors, have been studied to reflect convenience yields in the cash market due to their safety and liquidity (Bansal, Coleman and Lundblad (2010); Krishnamurthy and Vissing-Jorgensen (2012); Nagel (2016); Adrian, Crump and

[3] Data source: https://fiscaldata.treasury.gov/datasets/monthly-statement-public-debt/summary-of-treasury-securities-outstanding.

[4] Data source: https://www.sifma.org/resources/research/us-treasury-securities-statistics/.

[5] See https://www.newyorkfed.org/markets/primarydealers.

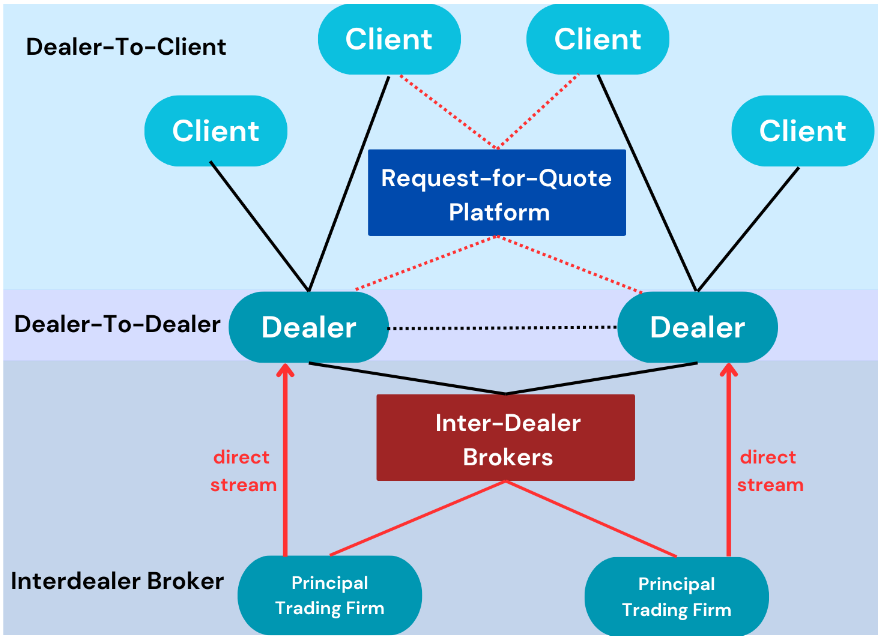


Figure 1: Structure of the Treasury secondary cash market.

Source: Brain et al. "Unlocking the Treasury Market through TRACE," Federal Reserve Bank of New York Liberty Street Economics (blog), September 28, 2018, http://libertystreeteconomics.newyorkfed.org/2018/09/unlocking-the-treasury-market-through-trace.html.

Vogt (2019)). The literature suggests that the high liquidity and safety of Treasuries drive up their price and consequently drive down their yield relative to assets that do not share these attributes to the same extent. The yield spread between two assets with identical attributes, except for differing liquidity or safety, is considered the liquidity or safety premium. Such premia are sensitive to demand and supply factors due to incomplete markets (Acharya and Laarits (2023)). For instance, heightened market demand for safe assets can prompt broker-dealers to increase short selling to meet this demand. If short selling incurs significant costs (e.g., Banerjee and Graveline (2013)), holding safe assets like Treasuries can offer non-pecuniary benefits by mitigating potential costs associated with short selling.

Despite the widely accepted convenience value of holding Treasuries, the post-GFC regulatory reforms imposed on financial institutions may increase the holding cost of Treasuries. Among these regulations, the most relevant is the Supplementary Leverage Ratio (SLR). US regulators proposed the SLR in 2012 and finalized the rule for the "enhanced" SLR in April 2014, with final implementations mostly completed by January 2018. As a non-risk-weighted capital regulation, the SLR requires US globally systemically important Bank Holding Companies (BHCs) to maintain capital equal to or greater than 5% of their total assets, regardless of the risk composition of those assets. Duffie (2018) argues that "the SLR increases the 'rental cost' for space on a bank's balance sheet." Given the lack of regulatory differentiation of asset risk, a same-size balance sheet expansion will incur higher balance sheet costs if the bank expands its holdings of safer assets, as risky assets typically come with higher returns [6].

2.2 Treasury Repo Market

In addition to selling them outright in the Treasury cash market, investors often use Treasuries as collateral to borrow cash on a short-term basis, particularly in the repurchase agreement (repo) market. A repo is a transaction in which one party sells an asset to another party with a promise to repurchase the asset later. The difference between the sale and repurchase price specified in a repo contract reflects the implied interest rate. For example, if a firm sells $9 million in Treasuries today and agrees to repurchase them for $9.09 million in a year, the implied interest rate is 1 percent. The securities serve as collateral to protect the cash investor against the risk that the collateral provider cannot repurchase the securities at the later date. The US Treasury repo market is crucial because it is a key source of short-term funding for securities dealers and some of their clients. As of the end of 2021, the average daily outstanding US Treasury repo was $1.7 trillion, with over 70% being overnight repos and less than 30% term repos [7] [8].

Figure 2 describes the main players and their relations in the U.S. Treasury repo market. This market comprises two segments, differentiated by their settlement processes: triparty repo and bilateral repo. In a triparty repo, a third party, typically a clearing bank, is involved. The clearing bank provides back-office support to both parties in the trade, including settling the repo on its books and ensuring that the terms of the agreement are met. In contrast, in a bilateral repo, each counterparty's custodian bank is responsible for clearing and settling the trade. The US Treasury repo market is roughly split between these two segments (Copeland et al. (2014); Baklanova et al. (2019)). Furthermore, within the triparty

[6] An implicit assumption here is the SLR is binding. For most of the period since the start of 2018 when the SLR was implemented, most of the big six U.S. BHCs maintained SLRs well above the 5 percent minimum level. However, the SLRs have trended down since 2021, approaching the 5 percent minimum level. See Cochran et al. (2023) for more details.

[7] Data source: https://www.sifma.org/wp-content/uploads/2022/02/SIFMA-Research-US-Repo-Markets-Chart-Book-2022.pdf.

[8] For comparison, in 2018, the average amount of outstanding securities lending arrangements against cash was around $700 billion. The size of the commercial paper market, a source of unsecured short-term funding for firms, was around $1 trillion. See Baklanova et al. (2019) for more details.

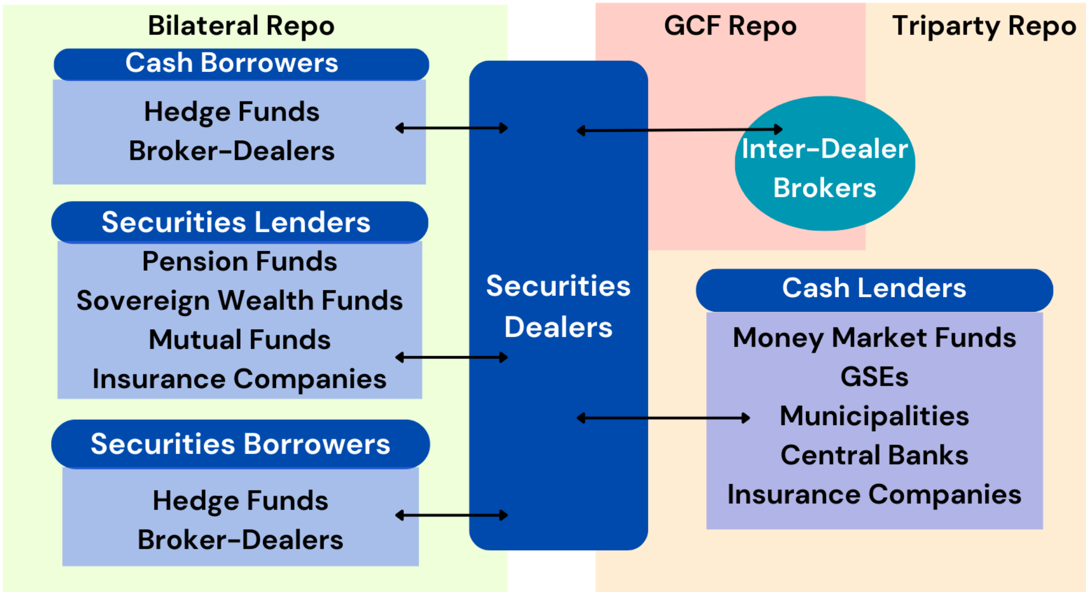


Figure 2: Structure of the Treasury repo market.

Source: Baklanova et al. "Reference Guide to U.S. Repo and Securities Lending Markets," Federal Reserve Bank of New York Staff Reports, no. 740 September 2015, https://www.newyorkfed.org/medialibrary/media/research/staff_reports/sr740.pdf.

repo market, a special general collateral financing repo service (GCF Repo) allows securities dealers registered with the Fixed Income Clearing Corporation (FICC) as netting members to trade repos among themselves. Thus, the GCF repo primarily functions as an inter-dealer market. Hereafter, I refer to non-GCF triparty repo simply as triparty repo. Triparty repo typically involves transactions with "general collateral," where the cash investor agrees to accept any securities within a specified asset class. In contrast, bilateral repos usually require specific securities, identified at the CUSIP level, to be agreed upon at the time of the trade.

As in the cash Treasury market, broker-dealers are also key intermediaries in the repo market, facilitating transactions between cash lenders (securities borrowers) and cash borrowers (securities lenders). Since triparty repo features "general collateral," cash lenders often use this platform to securely invest cash. The most important cash lenders in the triparty repo market are Money Market Funds (MMFs). As of September 30, 2020, the Financial Accounts of the United States show that MMFs accounted for close to 22% of total repo assets (Baklanova, Kuznits and Tatumetal (2021)). Broker-dealers are the main securities borrowers in the triparty repo market, sourcing cash from conservative cash investors such as MMFs. Unlike the triparty sector, the bilateral repo allows for the specific designation of securities when negotiating the terms of trade. Thus, the bilateral sector is heavily used by securities borrowers who seek specific types of assets. Common reasons to borrow specific securities include covering short sales, remedying failures to deliver securities to settle a transaction, or covering a hedge of a position. Firms managing large portfolios of securities, such as pension funds, central banks, or insurance companies, are the main providers of specified collateral securities. Hedge funds, a type of riskier investor, engage in both cash borrowing and securities borrowing in the bilateral repo market. First, hedge funds use bilateral repo intensively to obtain leverage to finance their Treasury cash positions, with funds transmitted by dealers from the triparty market (Baklanova, Copeland and McCaughrin (2015); He, Nagel and Song (2022)). Second, hedge funds also use the bilateral market to borrow securities to cover short sales or hedge positions.

U.S. Treasury repos are considered very low risk, as the transactions are collateralized by U.S. Treasuries. Consequently, investors often regard the U.S. Treasury repo rate as an almost risk-free rate. However, in practice, demand and supply factors can cause the Treasury repo rate to deviate somewhat from the risk-free rate. During periods of financial market tension, the shortage of safe and liquid assets often bestows Treasuries with a scarcity premium in the repo market (Jordan and Jordan (1997); d'Amico, Fan and Kitsul (2018)). This is particularly true in the bilateral repo sector, where securities borrowers seek specific securities. Due to high demand and low availability, Treasury lenders can use their securities to borrow cash at lower rates in the repo market. The rate spread between contracts with identical terms but different collateral is considered the scarcity premium. This phenomenon can be seen as a spillover of the liquidity/safety premium that Treasuries enjoy in the cash market (Duffie (1996); Jordan and Jordan (1997); d'Amico, Fan and Kitsul (2018)).

The SLR affects broker-dealers' intermediary activities not only in the Treasury cash market but also in the repo market. Like Treasury cash investments, Treasury repos are also characterized by very low risk but are subject to the same 5% capital requirement. Compared to riskier assets, the balance sheet expansion of Treasury repos implies a wealth transfer from the intermediary to the creditor. The interest payment must be large enough to offset this wealth transfer, resulting in a higher repo rate.

The Treasury cash and repo markets are crucial financial markets globally. Private participants heavily rely on these markets for low-cost financing of cash and securities. Additionally, both markets play significant roles in the implementation of monetary policy by authorities. The efficient functioning of these markets is essential for supporting financial stability and price discovery. In the remainder of this section, I analyze the performance of the Treasury cash and repo markets during the recent recessions. The aim is to provide empirical evidence on the differing causes of financial market turmoil during these periods. This evidence motivates the development of a general equilibrium model designed to simultaneously account for the varying financial market observations and the declining macroeconomic variables observed during these recessions.

2.3 Observations during GFC

The 2008 GFC was triggered by turmoil in the financial markets. During this period of economic distress, the subprime mortgage crisis led to diminished confidence in risky assets and a heightened preference for safe and liquid assets, such as U.S. Treasuries, resulting in what is known as a flight-to-liquidity and flight-to-safety (Longstaff (2002); Goldreich, Hanke and Nath (2005); Beber, Brandt and Kavajecz (2009)). Consequently, it is reasonable to expect a convenience yield on both direct and repo holdings of Treasuries. Additionally, since the SLR was not in place during this time, the inconvenience cost associated with holding Treasuries was relatively subtle.

Figure 3 illustrates key financial and macroeconomic variables around two significant events during the early stages of the GFC: the Bear Stearns liquidation on July 31, 2007, and the Lehman Brothers bankruptcy on September 15, 2008. The Treasury yields are daily series of constant maturity Treasury (CMT) yields sourced from FRED. The Repo wedge is a monthly series representing the spread between the GCF repo rate and the Effective Federal Funds Rate (EFFR); GCF data is obtained from DTCC, while EFFR data is from FRED. OIS rates are daily series downloaded from Bloomberg. Primary dealers' repo positions are reported weekly and summarized from the Fed Board FR2004 database. The output gap is calculated as the difference between real GDP and potential GDP, both of which are sourced from FRED. The PCE price index is a monthly series taken from FRED.

The Treasury cash market experienced declining yields across most tenors, with larger decreases observed in longer maturities (see Panel A). The 3-month Treasury yield dropped by more than one percentage point after each event, while the declines in the 5-year and 10-year Treasury yields were relatively smaller. This expansion in the term structure is consistent with the flight-to-liquidity theory. To isolate the (in)convenience yield, I use the Overnight Index Swap (OIS) rate, which reflects the interest rate of a contract with cash flows similar to those of Treasuries. The OIS is a fully collateralized interest rate swap contract that exchanges a constant cash flow for a floating payment indexed to the geometric average of the daily effective federal funds rate. The spread between Treasury yields and the OIS for the same maturity provides insight into the extra benefit or cost associated

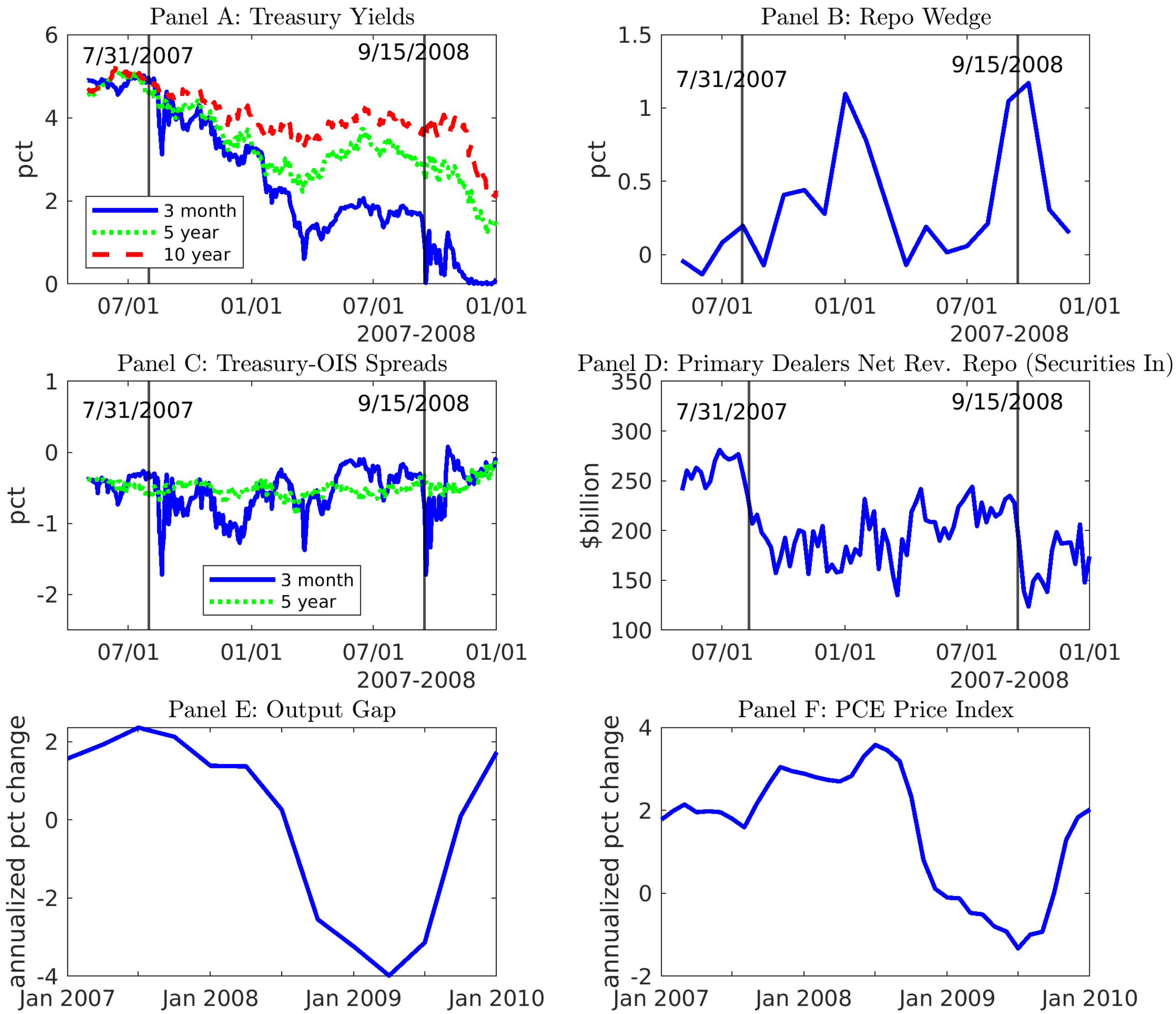


Figure 3: What happened during the early stage of GFC?

with holding Treasuries [9]. A negative Treasury-OIS spread indicates a convenience yield, where an increase in Treasury prices and a decrease in yields suggest that Treasuries offer additional value. Panel C shows that the Treasury-OIS spread was predominantly negative, indicating a convenience yield of holding Treasuries. This finding aligns with the flight-to-safety theory, as Treasuries, issued by the US federal government, feature very low credit risk. The term structure of the spread steepened following the two events, with the spread for short maturities decreasing more significantly than that for longer maturities. This reflects a flight-to-liquidity effect, where short-term Treasuries appreciated more than their longer-term counterparts.

The flight-to-liquidity and flight-to-safety premia also influenced the Treasury repo market. The ideal context to study the (in)convenience yield in the Treasury repo market would be the bilateral repo sector, where securities borrowers seek specific types of assets. However, due to data limitations, the bilateral repo rate for the studied period is unavailable. Instead, I use the spread between the GCF repo rate and the Effective Federal Funds Rate (EFFR) as a proxy for the (in)convenience yield in the repo market. The GCF repo market is an inter-dealer platform where smaller dealers obtain financing from larger dealers and then lend to cash borrowers, such as hedge funds, in the bilateral market. The EFFR, on the other hand, can be considered a risk-free rate. Thus, the GCF repo-EFFR spread measures the (in)convenience value for dealers holding Treasuries indirectly through the repo market [10]. Panel B shows that the repo wedge dropped sharply after the two events, indicating a decrease in the collateralized borrowing cost for dealers. This drop suggests that, following the event shocks, the market placed a higher value on the safety and liquidity of Treasury securities, providing Treasury holders with greater advantages when using their assets to finance cash in the repo market.

The increased demand for Treasuries due to economic distress was met by dealers through the repo market. I define the net repo holding of Treasury securities by primary dealers as the net reverse repo, which is calculated as reverse repos minus repos [11]. Panel D shows that primary dealers reduced their net repo Treasury holdings by more than $50

[9] Anther commonly used measure is the Treasury-LIBOR spread. I do not use this measure as OIS is available for more tenors, so the matching with Treasuries is better.

[10] A concern here is, EFFR is an uncollateralized rate while GCF repo rate is a collaterized rate. Alternatively I can take the GCF-triparty repo spread as the (in)convenience yield in the repo market, since triparty repo market involves a large number of cash-rich investors such as MMFs. However, during GFC, the Treasuries faced high extra demand due to their safety and liquidity. In this case, the triparty rate itself may contain liquidity and safety premia. Therefore I do not use this spread.

[11] I follow New York Fed to define repo as security out and reverse repo as security in from the stand point of dealers.

billion in response to the Bear Stearns liquidation and the Lehman Brothers bankruptcy. This reduction is consistent with market equilibrium under conditions of flight-to-liquidity and flight-to-safety: as the market experienced heightened demand for Treasuries from private participants, brokers reduced their holdings to meet this additional demand, with the adjustment occurring predominantly through the Treasury repo market.

Lastly, Panels E and F illustrate the behavior of the output gap and inflation rate during the early stages of the GFC. Both variables declined consistently during the market turmoil, reflecting typical recessionary trends. The output gap fell sharply, reaching a low of -4% by mid-2009, while the U.S. economy largely experienced deflation during this period. In the next subsection, I will demonstrate that despite both recessions showing declines in the output gap and inflation, the financial market behaviors during the Covid-19 recession differed significantly from those observed during the GFC.

2.4 Observations during Covid-19

In March 2020, the financial markets experienced one of the most dramatic upheavals in history [12]. Figure 4 depicts key financial and macroeconomic variables from this period [13]. Between March 9 and March 23, the 3-month Treasury yield rose by more than half a percentage point, while the 10-year Treasury yield fell by over half a percentage point. This led to a widening of the 10-year minus 3-month Treasury yield spread by more than one percentage point (see Panel A). Unlike the GFC, where most Treasury maturities saw price appreciation as investors sought safety, the mid- and long-term Treasuries depreciated during the Covid-19 period. This suggests that the market perceived a reduced value in long-term Treasuries despite their inherent safety features.

I use the Treasury-OIS spread to gauge the (in)convenience yield in the Treasury cash market. In contrast to the GFC, the spread was positive for most maturities during the Covid-19 pandemic (see Panel C). A positive Treasury-OIS spread indicates that holding Treasuries incurred a net non-pecuniary cost, suggesting a net inconvenience yield in the Treasury cash market. This outcome is consistent with the implementation of the SLR during this period, which introduced additional balance sheet costs for holding safe assets like Treasuries [14]. Moreover, the rise in the 10-year Treasury-OIS spread during this time signals

[12] Eisenbach and Phelan (2023) document the Treasury market distress; See Baker et al. (2020) and Mazur, Dang and Vega (2021) for the stock market crash.

[13] All variables are defined in the same way as for the GFC analysis expect otherwise explained.

[14] Klingler and Sundaresan (2023) find empirical evidence of diminishing Treasury convenience yields after GFC and identify the balance sheet constraint as a key reason.

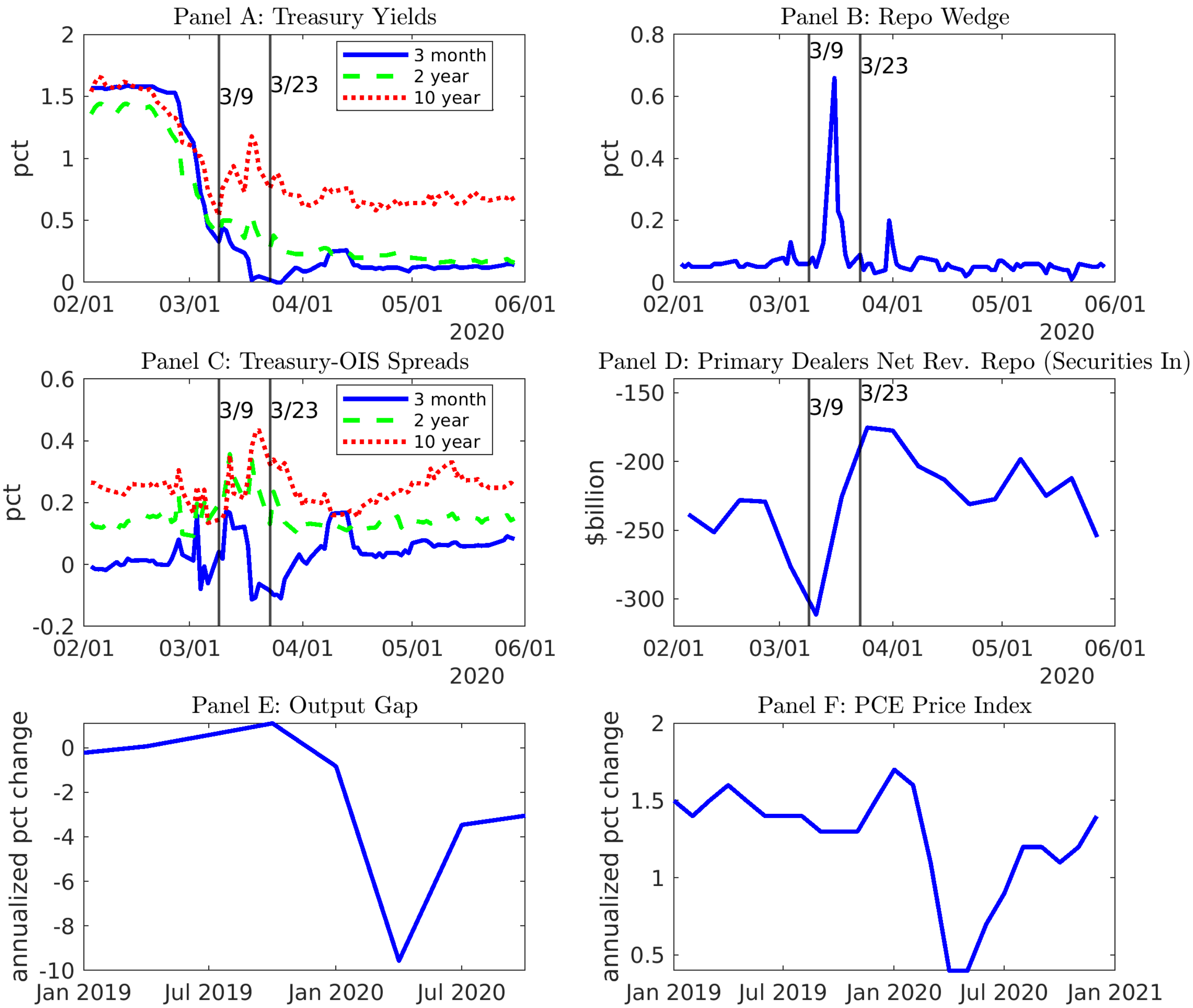


Figure 4: What happened during Covid-19?

a higher inconvenience associated with long-term Treasuries. This finding is particularly surprising given that economic distress typically drives down the Treasury-OIS spread for safe assets during recessions.

The Treasury repo market also experienced significant fluctuations during this period. To measure the (in)convenience yield in the Treasury repo market, I follow He, Nagel and Song (2022) and use the GCF-triparty repo spread. The triparty repo rate is used as a benchmark because it represents a collateralized market where cash-rich investors can lend cash with minimal risk, providing a more suitable comparison with the GCF repo rate. During the Covid-19 pandemic, the ample Treasury supply minimized concerns about the triparty repo rate reflecting a scarcity premium. As shown in Panel B, the repo wedge spiked by more than half a percentage point during this brief period. Complementing this, Panel D examines the changes in primary dealers' net reverse repo holdings, which surged by roughly $120 billion between March 9 and March 23. These observations contrast sharply with those during the GFC. The large increase in net reverse repo positions indicates substantial selling pressure on Treasuries, compelling primary dealers to absorb the excess supply through the repo market.

Finally, Panels E and F depict the output gap and inflation during the Covid-19 period. The output gap plummeted to its lowest level of -10% in the first half of 2020, while the inflation rate remained positive. The sharp drop in GDP was largely due to lockdowns imposed by major economies in response to the spreading virus. Despite the significant decline in actual GDP, potential GDP was not substantially affected, leading to a pronounced reduction in the output gap.

This section analyzes the Treasury cash and repo markets, highlighting their behavior during the two recent recessions. The evidence shows distinct patterns in these markets for the GFC and the Covid-19 pandemic, suggesting that different financial shocks triggered each recession despite similar macroeconomic patterns. In the next section, I introduce a modified Preferred-Habitat New Keynesian model with a detailed representation of financial markets to explain these observations during the two recessions.

3 The Model

I now develop a New Keynesian Preferred Habitat model integrated with a repo market to explain the observations from the GFC and Covid-19 documented in the previous section. This model also provides novel insights into the implications of monetary policies by accounting for various financial frictions. It builds upon the framework established by Ray (2019), which incorporates the entire term structure of Treasury yields into the aggregate borrowing cost. In his model, financial market disruptions cause long-term rates to deviate from the short rate, leading to equilibrium dynamics that differ from those predicted by a standard New Keynesian model.

Although the macroeconomic dynamics in this model are represented as a linearized equilibrium system, it is important to understand the underlying foundations. Figure 5 summarizes the key agents and their interactions within the model's framework. The model introduces innovative elements, such as incorporating a Treasury repo market in line with He, Nagel and Song (2022) and allowing for the (in)convenience (dis)utility of holding Treasuries for households. The government's Treasury supply is absorbed through the holdings of habitat investors and arbitrageurs. In equilibrium, Treasury prices (yields) are determined by the balance between bond supply and demand.

Arbitrageurs are further divided into dealers and hedge funds, integrating the Treasury repo market, where market clearing determines equilibrium Treasury repo rates. Households engage in saving and borrowing through bundles of savings products indexed to Treasury yields and repo rates, ensuring that any shifts in market rates are transmitted to the macroeconomic side. Innovatively, the model introduces a convenience component in households' marginal utility of saving. During economic distress, the marginal convenience utility for households to save increases, which encourages saving and discourages consumption.

3.1 Macroeconomic Dynamics

The macroeconomic framework of this model follows the approach of Werning (2011), where the equilibrium dynamics are captured by a three-equation New Keynesian system in continuous time. The first equation is the IS curve, expressed as:

$$\mathrm{d}x_t = \varsigma^{-1} \left(\tilde{r}_t + \varpi_t - \pi_t - \bar{r} \right) \mathrm{d}t, \tag{1}$$

where x_t represents the output gap—the logarithmic difference between actual output and the potential output that would exist under flexible prices. π_t denotes the inflation rate, and $\bar{r}$ is the natural borrowing rate, corresponding to a zero output gap in a scenario of perfect price flexibility and no financial frictions. The term ς^{-1} is the elasticity of intertemporal substitution [15].

[15] For example, in the CRRA utility function, $u(c) = \frac{c^{1-\varsigma}-1}{1-\varsigma}$.

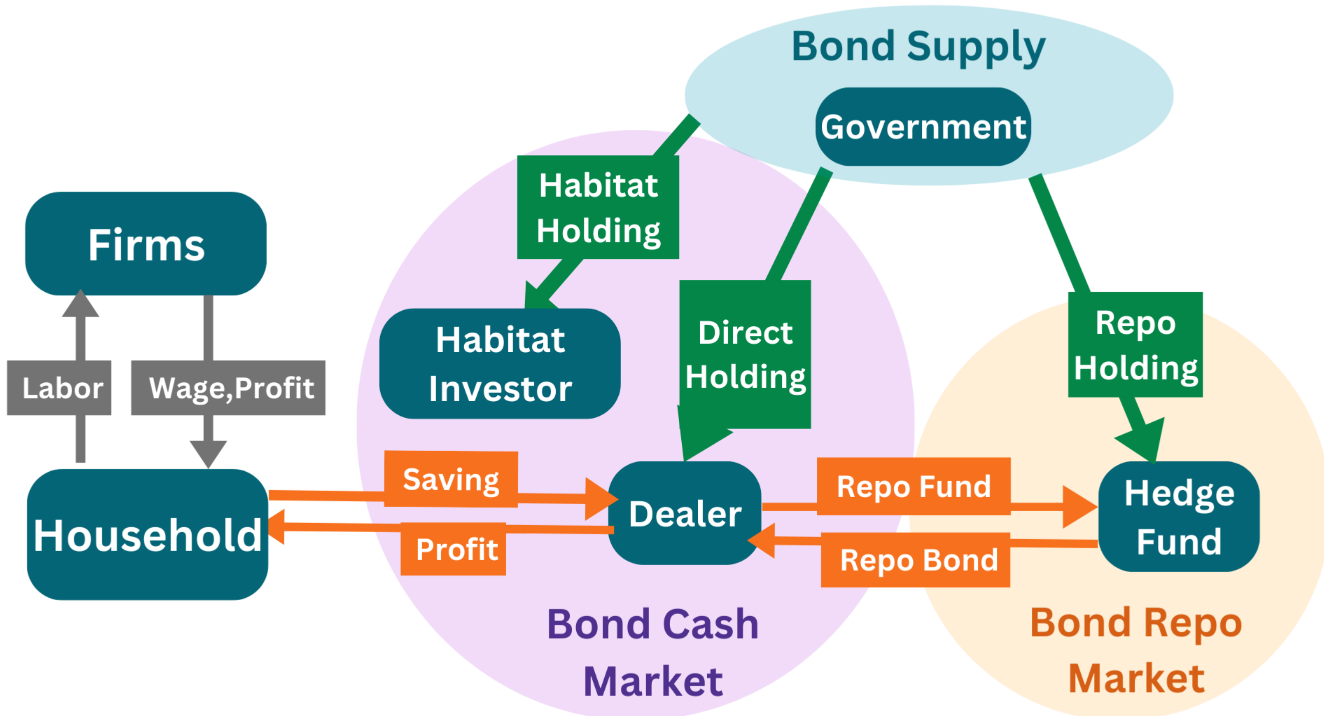


Figure 5: Overview of agents and flows in the model.

There are two key innovations in this IS curve. First, following Ray (2019), the model assumes that the nominal borrowing cost in the economy deviates from the policy rate. Instead, the overall borrowing cost is governed by $\tilde{r}_t$, the aggregate nominal rate, which is a combination of the term structures for bond yields and repo rates:

$$\tilde{r}_t \equiv \int_0^T \eta^i(\tau) i_t(\tau) \mathrm{d}\tau + \int_0^T \eta^R(\tau) R_t(\tau) \mathrm{d}\tau, \tag{2}$$

where $i_t(\tau)$ represents the bond yield with maturity τ at time t and $R_t(\tau)$ the repo rate. The weight functions $\eta^i(\tau)$ and $\eta^R(\tau)$ determine the relative importance of bond yields and repo rates in shaping the overall borrowing cost. Appendix E provides the micro foundation for the aggregate nominal rate $\tilde{r}_t$. Specifically, when households have heterogeneous access to capital markets and a "head of households" sets transfers to equalize wealth across households, the equilibrium aggregate nominal rate takes the form of equation (2). Expressing the aggregate nominal rate in this manner enables the model to capture aspects of investment and savings decisions that simpler models typically overlook. For example, in reality, durable consumption and housing are primarily influenced by long-term rates. Moreover, the wholesale secured funding market, such as the repo market, plays a crucial role in shaping the aggregate borrowing cost in the economy.

What's more, ϖ_t denotes the marginal quality value for households to invest in saving products, with a larger ϖ_t increasing the marginal benefit of saving. The quality value brings non-pecuniary benefit for households to save, similar to Bodenstein et al. (2023) and Kekre and Lenel (2024). I model this marginal quality value in relation to the exogenous private habitat demand shifter β_t using the following linear relationship:

$$\varpi_t = \varpi \beta_t, \tag{3}$$

where $\varpi < 0$ is a negative constant. Equation (3) states that when there is an exogenous demand shock from private habitat investors ($\beta_t < 0$), the marginal quality value of investing in saving products rises. This effect is particularly pronounced during periods of financial distress, when heightened uncertainties make safe assets like U.S. Treasuries more attractive due to their inherent safety and liquidity. Conversely, in the case of an exogenous supply shock ($\beta_t > 0$), the marginal quality value decreases as Treasuries become more abundant in the market. This innovation introduces a second channel, beyond the traditional interest rate channel, through which the demand and supply factors can impact the real economy. Specifically, during periods of economic distress, private investors typically increase their demand for safe assets. According to the traditional interest rate channel, this heightened demand raises Treasury prices and lowers yields, which would generally stimulate aggregate consumption. However, in this model, a second mechanism is at play. As demand for safe assets rises, households place greater value on the quality of these assets. This increased valuation leads to a rise in savings and a reduction in consumption, contrasting with the usual stimulative effect predicted by the interest rate channel. Thus, while higher demand for Treasuries lowers yields and might normally boost consumption, the model suggests that the enhanced focus on asset quality during distress periods can actually dampen consumption.

The Phillips curve in this economy is given by

$$\mathrm{d}\pi_t = (\chi \pi_t - \delta x_t)\, \mathrm{d}t, \tag{4}$$

where χ controls the discount rate and δ measures the stickiness of price. A smaller δ indicates greater price stickiness. As $\delta \to \infty$, the economy approaches a state of perfectly flexible prices. This forward-looking Phillips curve states that the inflation is proportional to the present value of future output gaps. The parameter χ adjusts the rate at which past inflation affects current inflation, while δ reflects how the output gap impacts inflation.

Finally, the policy rate is governed by the Taylor rule:

$$\mathrm{d}r_t = -\psi_r \left(r_t - \phi_\pi \pi_t - \phi_x x_t - r^* \right) \mathrm{d}t + \sigma_r \, \mathrm{d}B_{r,t}, \tag{5}$$

where ψ_r controls the mean-reversion rate of the policy rate, ϕ_π and ϕ_x are the weights assigned to inflation and output targets, r^* is the target rate that maintains a zero output gap in the steady state economy, and $dB_{r,t}$ is the policy uncertainty term with volatility governed by σ_r. The macroeconomic dynamics in this model are fully characterized by equations (1)-(5).

3.2 Term Structures Determination

The term structures for bonds and repo assets are determined according to a Preferred Habitat model that incorporates repo assets. There are two risk factors in this economy: the short rate and the private habitat demand shifter. The processes governing these risk factors are specified as follows:

$$d \begin{bmatrix} r_t \\ \beta_t \end{bmatrix} = - \begin{bmatrix} \kappa_r & \kappa_{r\beta} \\ 0 & \kappa_\beta \end{bmatrix} \left(\begin{bmatrix} r_t \\ \beta_t \end{bmatrix} - \begin{bmatrix} r^{ss} \\ \beta^{ss} \end{bmatrix} \right) dt + \begin{bmatrix} \sigma_r & 0 \\ 0 & \sigma_\beta \end{bmatrix} d \begin{bmatrix} B_{r,t} \\ B_{\beta,t} \end{bmatrix},$$

where r_t is the short rate and β_t is the demand shifter. $dB_{r,t}$ and $dB_{\beta,t}$ are Brownian motions associated with these risk factors. Collecting the two risk factors into a vector s_t, I can rewrite the processes in vector format as:

$$ds_t = -\Gamma(s_t - s^{ss})dt + \Sigma dB_t. \tag{6}$$

Equation (6) generates uncertainties into the model, with the long-run means of r_t and β_t being r^{ss} and β^{ss}, respectively. In the baseline specification, I assume $\beta^{ss} = 0$. Since Γ is non-diagonal, the drift terms of r_t and β_t are correlated. This correlation arises as an equilibrium result, influenced by the feedback mechanism of the Taylor rule. Specifically, when the demand factor affects bond prices, aggregate consumption adjusts according to the IS curve. Consequently, the Taylor rule requires the short rate to adjust in response to stabilize the economy.

3.2.1 Habitat Agents

There are agents demanding Treasuries as habitat investors in the market. Habitat agents with maturity preference τ hold only bonds with maturity τ and no other maturities [16]. In reality, these agents can be pension funds and insurance companies who typically hold bonds with maturities that match their liabilities. Additionally, central banks also act as habitat investors, aligning with the observation that their asset purchases or sales often focus on particular products and maturities.

In this section, I consider only habitat demand from private agents and postpone modeling habitat demand from the central bank to later sections. The aggregate demand from private habitat investors is given by:

$$H_t(\tau) = -\alpha(\tau) log P_t(\tau) - \theta(\tau)\beta_t. \tag{7}$$

This demand depends on bond prices $P_t(\tau)$ and an exogenous demand shifter β_t. Note that the one-dimensional demand shock β_t can achieve heterogeneous effect on maturities by the location function $\theta(\tau)$. In the default specification, $\theta_\tau > 0$ for all maturities, although this can be adjusted if needed. The flexibility of $\theta(\tau)$ allows for the analysis of non-standard demand shocks such as operation twist, where the short maturities are sold and long maturities are purchased. In this setup, a negative β_t represents a demand shock, while a positive β_t represents a supply shock. The parameter $\alpha(\tau) \geq 0$ fensures that habitat investors decrease their demand if bond prices rise, reflecting a typical inverse relationship between price and demand.

3.2.2 Arbitrageurs

To incorporate the Treasury repo market, I build on the framework of He, Nagel and Song (2022) by distinguishing between hedge funds and dealers. Hedge funds use repo transactions to finance their bond purchases, borrowing from dealers in the repo market[17]. I assume there is a representative hedge fund whose optimization problem is formulated as follows:

[16] Vayanos and Vila (2021) provide a micro optimization foundation for this specification.

[17] "Hedge funds typically finance their cash Treasury holdings with repo, with the vast majority of hedge fund repo borrowing taking place in the bilateral repo market...Hedge fund repo borrowing doubled over the two years preceding the pandemic...reaching \$1.2 trillions..." See Banegas, Monin and Petrasek (2021) for hedge funds' repo holdings.

$$
\max_{Q_t^h(\tau)} E_t\left[dW_t^h\right] - \frac{1}{2\rho_h}\,\mathrm{Var}_t\left[dW_t^h\right],
$$
$$
s.t.
$$
$$
dW_t^h - W_t^h r_t dt = \int_0^T \underbrace{Q_t^h(\tau)}_{\text{repo demand}} \underbrace{\left(\frac{dP_t(\tau)}{P_t(\tau)} - R_t(\tau)dt\right)}_{\text{trading profit}} d\tau. \tag{8}
$$

The hedge fund maximizes expected trading profits while minimizing risk. The trading profit is defined as the difference between the bond's return and the repo rate, reflecting the cost of financing bonds through repo borrowing. The parameter ρ_h controls the risk-bearing capacity of hedge funds, influencing its optimization problem.

On the other side of the repo market, dealers play a dual role: they provide secured funding to hedge funds and directly hold the remaining bonds on the market. Both their direct holdings and repo operations are financed through the risk-free market. For this setup, I assume a representative dealer whose optimization problem is formulated as follows:

$$
\max_{X_t(\tau),Q_t^d(\tau)} E_t\left[dW_t^d\right] - \frac{1}{2\rho_d}\,\mathrm{Var}_t\left[dW_t^d\right],
$$
$$
s.t.
$$
$$
\begin{aligned}
dW_t^d - W_t^d r_t d_t = {} & \int_0^T \underbrace{X_t(\tau)}_{\text{direct holdings}} \Big(\underbrace{\frac{dP_t(\tau)}{P_t(\tau)} - r_t dt}_{\text{excess return}} - \underbrace{\Lambda_t(\tau)dt}_{\text{non-pecuniary cost}}\Big) d\tau \\
& + \int_0^T \underbrace{Q_t^d(\tau)}_{\text{repo supply}} \Big(\underbrace{R_t(\tau) - r_t}_{\text{repo wedge}} - \underbrace{\Lambda_t(\tau)}_{\text{non-pecuniary cost}}\Big) dt d\tau.
\end{aligned} \tag{9}
$$

Note that the representative dealer shares the same maximization objective as the representative hedge fund except the risk aversion. Dealers hold bonds in two formats: the direct holdings and the indirect holdings through repo. This approach aligns with observed dealer behaviors in the Treasury markets, particularly during periods of market stress, such as the pandemic. During the first quarter of 2020, for instance, significant adjustments in government bond holdings by various investors were largely absorbed by broker-dealers, with a substantial portion accommodated through repo financing. As detailed by He, Nagel and Song (2022), broker-dealers temporarily absorbed \$400 billion through repo financing, illustrating their critical intermediary role in the market. [18]

A non-pecuniary cost (or benefit, if negative) is incorporated for both direct and repo holdings of bonds to represent the ”inconvenience yield”—the wedge between the total yield and the pecuniary yield. Treasuries are typically associated with significant convenience yields before the GFC due to their safety and liquidity features (Longstaff (2002); Krishnamurthy and Vissing-Jorgensen (2012); Nagel (2016)). During financial distress, the demand for safe assets like Treasuries surges, making them highly desirable due to their lower transaction costs and higher ease of sale. Holding Treasuries provides a buffer against negative funding shocks, and for dealers, having Treasuries on their balance sheets reduces the costs associated with asset financing for short-selling. This convenience yield often extends to the Treasury repo market because of the collateral's scarcity. Conversely, in the post-crisis period, U.S. Treasury yields have frequently exceeded other risk-free benchmark rates (Du, Im and Schreger (2018); Klingler and Sundaresan (2019); J Jermann (2020); Augustin et al. (2021)). Research suggests that this disparity results from dealers' balance sheet constraints, particularly those imposed by the Supplementary Leverage Ratio (SLR) requirement under Basel III regulations. The SLR mandates that bank holding companies maintain capital equal to or greater than 5% of their total assets, irrespective of the asset risk profile. Since the SLR does not differentiate between risky assets (e.g., unsecured loans) and safe assets (e.g., Treasury securities and repos backed by Treasuries), it effectively imposes additional holding costs on safe assets. This is because risky assets typically offer higher returns than safe assets, making the opportunity cost of holding safe assets higher. In the following discussion, I follow He, Nagel and Song (2022) to show that the same functional format for $\Lambda_t(\tau)$ can capture the mechanisms for both pre-and post-GFC periods.

For the post-GFC era, due to the SLR, dealers incur a balance sheet cost when increasing their Treasury holdings, either directly or through repo cash lending. Assuming a linear structure for the marginal holding cost, I define this cost as follows:

$$\Lambda_t(\tau) = \lambda B_t(\tau), \text{ where} \tag{10}$$

$$B_t(\tau) = X_t(\tau) + Q_t^d(\tau). \tag{11}$$

[18] "During the first quarter of 2020, foreign investors sold about $270 billion worth of Treasuries; mutual funds sold around $240 billion; hedge funds sold more than $30 billion...Much of this supply was temporarily accommodated by broker-dealers, partly through somewhat higher direct holdings (about $50 billion), but also indirectly through a massive expansion of $400 billion in repo financing... " See He, Nagel and Song (2022) for details.

Here, $\Lambda_t(\tau)$represents the marginal non-pecuniary cost of holding Treasuries, which increases proportionally with the size of the dealer's balance sheet $B_t(\tau)$. This balance sheet size is the sum of the dealer's direct holdings of bonds $X_t(\tau)$ and their repo supply $Q_t^d(\tau)$. This linear structure simplifies the solution for financial market equilibria and allows for affine bond prices. It aligns with empirical evidence indicating that balance sheet size impacts non-pecuniary costs. For instance, Klingler and Sundaresan (2023) document a well-fitted linear relationship between the Treasury yield-OIS spread and the relative size of holdings. Similar linear relationships are also observed in Moskowitz et al. (2024).

The pre-GFC era is documented with significant convenience yield of holding Treasuries. From equation (9), the marginal non-pecuniary benefit is represented by $-\Lambda_t(\tau)$. To relate this marginal benefit to the asset positions, a reasonable assumption is that the marginal benefit is proportional to the uncovered repo security lending position:

$$-\Lambda_t(\tau) = \lambda(-Q_t^d(\tau) - X_t(\tau)), \tag{12}$$

where $-Q_t(\tau)$ represents the quantity of net repo security lending for Treasuries with maturity τ. implies that when dealers need to deliver a large amount of securities that cannot be fully covered by their direct holdings, the non-pecuniary benefit of holding Treasuries increases. This is because, in the presence of high demand for Treasuries, dealers are willing to pay a premium to hold these assets for their safety and liquidity features and to avoid failures in security delivery. Essentially, equation (12) aligns with the structure of equation (10), meaning that the same specification for $\Lambda_t(\tau)$ can be applied to both pre- and post-GFC scenarios. In the pre-GFC context, the marginal benefit reflects the convenience yield associated with holding Treasuries during periods of high demand, whereas in the post-GFC context, it reflects the balance sheet constraints imposed by regulatory requirements.

4 General Equilibrium

This section outlines the general equilibrium of the model in a structured manner. I will start by defining the equilibrium conditions in the financial markets, specifically focusing on the Treasury cash and repo markets. Following this, I will incorporate the macroeconomic dynamics to complete the discussion of the economy's general equilibrium. To facilitate a clearer understanding, I will conclude with a simple case study that illustrates how the general equilibrium is determined.

4.1 Financial Markets Equilibrium

I follow the literature to normalize the aggregate bond supply for each tenor τ to be zero. The financial market equilibrium is defined in the following way:

DEFINITION 1. *(Financial Markets Equilibrium) The financial markets equilibrium is a collection of quantities $\{Q_t^h(\tau)\}$ by hedge fund, $\{Q_t^d(\tau), X_t(\tau)\}$ by dealer, $\{H_t(\tau)\}$ by habitat investor, and prices $\{P_t(\tau), R_t(\tau)\}$, such that*

1. *The representative hedge fund solves its optimization problem.*
2. *The representative dealer solves its optimization problem.*
3. *Both Treasury cash and Repo market clear for all maturities, i.e.,*

$$Q_t^h(\tau) = Q_t^d(\tau) \equiv Q_t(\tau), \tag{13}$$

$$H_t(\tau) + X_t(\tau) + Q_t(\tau) = 0. \tag{14}$$

To solve for the financial markets equilibrium, I rely on the structure of bond prices, leveraging results from the literature, particularly the affine structure established by Vayanos and Vila (2021)

$$P_t(\tau) = exp\Big[-\Big(A(\tau)' s_t + C(\tau)\Big)\Big], \tag{15}$$

where $A(\tau) = [A_r(\tau)\ A_\beta(\tau)]'$ is a two-dimensional vector capturing the sensitivities of the bond price to the state variables. $C(\tau)$ is a scalar function of maturity τ. Using Ito's Lemma, the instantaneous return on Treasury is derived as

$$\frac{dP_t(\tau)}{P_t(\tau)} = \mu_t(\tau)dt - A(\tau)'\Sigma dB_t, \tag{16}$$

where $\mu_t(\tau)$ is the instantaneous expected return, which can be expressed as:

$$\mu_t(\tau) \equiv A'(\tau)' s_t + C'(\tau) + A(\tau)'\Gamma\left(s_t - r^{ss}\epsilon\right) + \frac{1}{2}A(\tau)'\Sigma\Sigma' A(\tau). \tag{17}$$

The remaining term in the instantaneous return $A(\tau)'\Sigma dB_t$ captures the bond's exposure to risk. Replacing the instantaneous return (16) in the representative hedge fund's problem

(8), its FOC can be derived as

$$\mu_t(\tau) = \underbrace{R_t(\tau)}_{\text{financing cost}} + \underbrace{\frac{1}{\rho_h} A(\tau)' \Sigma \Sigma' \left[\int_0^T Q_t^h(\tau) A(\tau) d\tau \right]}_{\text{risk premium}}. \tag{18}$$

The hedge fund's optimization problem requires balancing the expected return against the financing cost and risk premium. The left-hand side of equation (18) gives the marginal benefit of shifting one unit of wealth from the short rate r_t to the Treasury bond with maturity τ, while the right-hand side gives the marginal cost as a sum of financing cost and risk premium. The financing cost arises because hedge fund finances its bond holdings through the repo market, incurring interest costs. Second, the risk premium reflects the additional cost due to the bond's exposure to uncertainties (captured by $A(\tau)'\Sigma dB_t$) and the hedge fund's risk aversion.

With the same replacing process, I can write the representative dealer's FOC with respect to direct holdings as

$$\mu_t(\tau) = \overbrace{r_t}^{\text{financing cost}} + \underbrace{\Lambda_t(\tau)}_{\text{non-pecuniary cost}} + \underbrace{\frac{1}{\rho_d} A(\tau)' \Sigma \Sigma' \left[\int_0^T X_t(\tau) A(\tau) d\tau \right]}_{\text{risk premium}}. \tag{19}$$

The left-hand side of equation (19) represents the expected return from holding an additional unit of the Treasury bond. This return is influenced by the bond's price sensitivity to the underlying risk factors and the overall state of the economy. The right-hand side represents the total cost of holding the bond, broken down into three components. The financing cost r_t is the cost the dealer incurs for financing the purchase of the bond. It reflects the return the dealer could earn on an alternative risk-free investment. The non-pecuniary cost $\Lambda_t(\tau)$ arises from regulatory constraints or other factors that make holding larger quantities of bonds more costly. Finally, the risk premium compensates the dealer for the uncertainty in bond prices due to exposure to market risks.

In addition to direct holdings, dealers also choose the quantity of repo assets $Q_t^d(\tau)$. The FOC with respect to repo assets yields the expression $R_t(\tau) - r_t - \Lambda_t(\tau)$, which represents the difference between the repo wedge and the non-pecuniary cost. For an inner solution, where the dealer holds both direct and repo assets, these two terms must be equivalent:

$$R_t(\tau) - r_t = \Lambda_t(\tau). \tag{20}$$

The left-hand side $R_t(\tau) - r_t$ reflects the cost of financing through the repo market relative to a risk-free alternative. The repo wedge represents the cost for the hedge fund to finance bond holdings through the repo market. It also measures the the additional return required by the dealer to justify holding bonds through the repo market instead of directly. The right-hand side $\Lambda_t(\tau)$ captures the cost or benefit associated with directly holding Treasuries. It could represent regulatory costs, liquidity premiums, or other market frictions. If the repo wedge is greater than the non-pecuniary cost, financing through the repo market is more expensive than holding bonds directly. In this case, the dealer would prefer to hold all bonds directly. On the other hand, if the repo wedge is less than the non-pecuniary cost, the repo financing is cheaper than holding bonds directly. In this case, the dealer would prefer to hold all bonds indirectly through repo. An inner solution occurs when the repo wedge equals the non-pecuniary cost. In this case, the dealer holds a balanced portfolio of direct and repo assets, as neither option is more expensive on the margin.

Equation (20) provides insight into the relationship between the repo wedge and the sign of $B_t(\tau)$, which is defined as the total bond position held by the dealer. When $B_t(\tau) > 0$, market clearing implies that the habitat holding is smaller than arbitrageurs' holding $H_t(\tau) < 0 < B_t(\tau)$. This indicates that Treasuries are abundant in the market. Since there is an oversupply of Treasuries, the dealer faces a higher cost to absorb these bonds. As a result, the dealer compensates for this by charging a higher repo rate than the risk-free rate, leading to a positive repo wedge. In contrast, when $H_t(\tau) > 0 > B_t(\tau)$, the Treasuries are scarce in the market. In this situation, dealers are under pressure to obtain more Treasuries to meet the high demand from habitat investors. This urgency causes the dealer to accept a repo rate that is lower than the risk-free rate. Consequently, the repo wedge becomes negative [19].

Imposing equation (20) on equation (18), I find that the FOCs for the dealer and the hedge fund only differ in risk aversion. Recall that the total balance sheet size $B_t(\tau)$ is a sum of the direct holding $X_t(\tau)$ and the repo holding $Q_t(\tau)$, equations (11), (13), and (14) imply that

$$H_t(\tau) = -B_t(\tau). \tag{21}$$

[19] Such observation is also called "scarcity premium" in the repo market. See for example, d'Amico, Fan and Kitsul (2018) and Arrata et al. (2020) for empirical evidence.

The optimal risk sharing quantities are

$$X_t(\tau) = -\frac{\rho_d}{\rho_d + \rho_h} H_t(\tau), \tag{22}$$

$$Q_t(\tau) = -\frac{\rho_h}{\rho_d + \rho_h} H_t(\tau). \tag{23}$$

In equilibrium, whether dealer or hedge fund holds a larger amount of bonds depends on their risk-bearing capacities. The entity with a lower risk aversion will take a larger position.

To understand the equilibrium in financial markets, I substitute the quantities in the dealer's FOC using equations (10), (21), and (22):

$$\underbrace{\mu_t(\tau) - r_t + \lambda H_t(\tau)}_{\text{net expected excess return}} = -A(\tau)' \underbrace{\frac{1}{\rho_h + \rho_d} \Sigma\Sigma' \left[\int_0^T H_t(\tau) A(\tau) d\tau \right]}_{\text{risk price}}. \tag{24}$$

Equation (24) differs from the standard preferred-habitat literature as the portfolio adjustment induces non-pecuniary cost. In equilibrium, the expected excess return of holding one more unit of maturity τ net of the non-pecuniary cost is equal to the risk premium that compensates arbitrageurs for bearing the short rate risk and the demand risk. The portfolio risk exposure increases by the covariance between the portfolio's risk factor sensitivity $\int_0^T H_t(\tau)A(\tau)d\tau$ and the additional position's risk factor sensitivity $A(\tau)$. Increasing bond position for maturity τ can expose arbitrageurs to higher risk in two ways. First, a larger aggregate holding holding makes investors more sensitive to volatility in unit prices, which is the external margin of the marginal cost ($\int_0^T H_t(\tau)A(\tau)d\tau$). The internal margin comes from the relative adjustment of positions across different maturities. If a portfolio adjusts to hold more bonds with maturities that are particularly sensitive to risk factors, the internal margin increases the portfolio's overall risk. However, when factor sensitivity $A(\tau)$ is constant across maturities, this internal margin vanishes, leaving only the external margin to impact the risk.

This model ensures an arbitrage-free equilibrium, as can bee seen from equation (24). The net expected excess return $\mu_t(\tau) - r_t + \lambda H_t(\tau)$ for any maturity is characterized by the product of that maturity's factor sensitivity and a common factor price for short rate and demand factor. In other words, the net expected excess return per unit of factor sensitivity is uniform across all assets (maturities). This uniformity prevents the possibility of constructing arbitrage portfolios, which would otherwise exploit discrepancies in returns across different maturities. The risk price involves equilibrium terms $H_t(\tau)$ and $A(\tau)$. In equation (24), I can express the instantaneous expected return $\mu_t(\tau)$ and the habitat demand $H_t(\tau)$ as functions of the risk factors. Collecting terms with related to the risk factors result in a two-equation ODE system to solve for $A(\tau$ and a scalar equation to solve for $C(\tau)$. The detailed derivations for the affine coefficients $A(\tau)$ and $C(\tau)$are provided in Appendix B1.

Given $A(\tau)$ and $C(\tau)$, the bond yields are derived as

$$\begin{aligned} i_t(\tau) &= -\frac{1}{\tau} logP_t(\tau) \\ &= \frac{1}{\tau}\Big(A(\tau)' s_t + C(\tau)\Big), \end{aligned} \tag{25}$$

and the equilibrium repo rates are derived as

$$\begin{aligned} R_t(\tau) &= \Delta_t(\tau) + r_t \\ &= \Lambda_t(\tau) + r_t \\ &= \lambda\Big[\alpha(\tau) logP_t(\tau) + \theta(\tau)\beta_t\Big] + r_t \\ &= \lambda\Big[\theta(\tau)\beta_t - \alpha(\tau)\Big(A(\tau)' s_t + C(\tau)\Big)\Big] + r_t. \end{aligned} \tag{26}$$

4.2 General Equilibrium

DEFINITION 2. *(General Equilibrium) The general equilibrium is a pair of financial intermediary efficiency $\hat{A}$ and policy persistence κ_r such that*

1. *Given policy persistence κ_r, financial markets (Treasry cash market, Treasury repo market) equilibrium results in a financial intermediary efficiency equal to $\hat{A}$.*

2. *Given the financial intermediary efficiency $\hat{A}$, the macroeconomy equilibrium results in a policy persistence equal to κ_r.*

With the asset prices solved from financial markets equilibrium, I now revisit the macroeconomic dynamics. First, I can rewrite the effective nominal rate as a linear function of state variables

$$\begin{aligned} \tilde{r}_t &= \int_0^T \eta^i(\tau) i_t(\tau) \mathrm{d}\tau + \int_0^T \eta^R(\tau) R_t(\tau) \mathrm{d}\tau \\ &\equiv \hat{A}' s_t + \hat{C}, \end{aligned} \tag{27}$$

where

$$\hat{A}' \equiv \int_0^T \frac{1}{\tau}\eta^i(\tau)A(\tau)'d\tau + \int_0^T \eta^R(\tau)\Big[\lambda\Big(\theta(\tau)\gamma' - \alpha(\tau)A(\tau)'\Big) + \epsilon'\Big]d\tau, \tag{28}$$

$$\hat{C} \equiv \int_0^T \frac{1}{\tau}\eta^i(\tau)C(\tau)d\tau - \lambda\int_0^T \eta^R(\tau)\alpha(\tau)C(\tau)d\tau. \tag{29}$$

The financial markets' equilibrium, influenced by the persistence of policy κ_r, feeds into the determination of the financial intermediary efficiency $\hat{A}$. This means that the stability and predictability of monetary policy impact how efficiently intermediaries can function in the markets. Conversely, the efficiency of financial intermediaries $\hat{A}$ influences the broader macroeconomic equilibrium, which in turn determines the level of policy persistence κ_r. This reflects the feedback loop where the functioning of financial markets and intermediaries affects the overall economy, which then guides central bank policy.

It follows that the IS curve in general equilibrium is expressed as

$$dx_t = \varsigma^{-1}(\hat{A}'s_t + \hat{C} + \varpi_t - \pi_t - \bar{r})dt. \tag{30}$$

The other two macroeconomic equations, the Phillips curve and the Taylor rule, remain unchanged. The macroeconomic dynamics in general equilibrium are fully characterized by equations (4), (5), and (30). These three equations form a system that describes the behavior of the economy with two state variables r_t and β_t, and two jump variables x_t and π_t. The rational expectation general equilibrium is summarized by the following four-equation system:

$$d\begin{bmatrix} r_t \\ \beta_t \\ x_t \\ \pi_t \end{bmatrix} = -\underbrace{\begin{bmatrix} \psi_r & 0 & -\psi_r\phi_x & -\psi_r\phi_\pi \\ 0 & \kappa_\beta & 0 & 0 \\ -\varsigma^{-1}\hat{A}_r & -\varsigma^{-1}(\hat{A}_\beta + \varpi) & 0 & \varsigma^{-1} \\ 0 & 0 & \delta & -\chi \end{bmatrix}}_{\Upsilon}\left(\begin{bmatrix} r_t \\ \beta_t \\ x_t \\ \pi_t \end{bmatrix} - \begin{bmatrix} r^{ss} \\ \beta^{ss} \\ x^{ss} \\ \pi_{ss} \end{bmatrix}\right)dt + \underbrace{\begin{bmatrix} \sigma_r & 0 \\ 0 & \sigma_\beta \\ 0 & 0 \\ 0 & 0 \end{bmatrix}}_{\begin{bmatrix} S \\ 0 * I_2 \end{bmatrix}} d\underbrace{\begin{bmatrix} B_{r,t} \\ B_{\beta,t} \end{bmatrix}}_{B_t}. \tag{31}$$

The solution to this system is derived from the eigenvalues and eigenvectors of the dynamic matrix Υ:

$$d \begin{bmatrix} r_t \\ \beta_t \end{bmatrix} = -\Gamma \left(\begin{bmatrix} r_t \\ \beta_t \end{bmatrix} - \begin{bmatrix} r^{ss} \\ \beta^{ss} \end{bmatrix} \right) dt + S dB_t, \tag{32}$$

$$\begin{bmatrix} x_t \\ \pi_t \end{bmatrix} = \Omega \left(\begin{bmatrix} r_t \\ \beta_t \end{bmatrix} - \begin{bmatrix} r^{ss} \\ \beta^{ss} \end{bmatrix} \right), \tag{33}$$

where Γ and Ω are derived from eigenvalues and eigenvectors of Υ [20]. The derivation is included in Appendix B2. In general equilibrium, the transition matrix Γ is an upper-triangular since the short rate will respond to the economic results caused by demand shocks.

The difficulty of solving this system arises from the fact that in this model, the dynamic matrix Υ is an equilibrium objective. To see this point, note that $\hat{A}_r$ and $\hat{A}_\beta$ are both endogenous terms and must be solved from the asset pricing side. The general equilibrium $\hat{A}$ must ensure that Γ in equation (32) coincides with that in equation (6). Appendix B3 provides a description of the solution algorithm for the general equilibrium. The process involves ensuring that the dynamic matrix Υ correctly reflects both the macroeconomic and financial market dynamics. The solution algorithm typically involves iterating between the asset pricing side (solving for $\hat{A}$) and the macroeconomic side (solving for Γ and Ω) until consistency is achieved.

The general equilibrium of this model can be described as a financial market-accommodated monetary policy rule, or vice versa, a monetary policy-consistent financial market optimization, in the similar spirit to Gallmeyer, Hollifield and Zin (2005) and Gallmeyer et al. (2007). When the central bank adjusts the short rate, the transmission to market rates is not perfect. The extent of this passthrough depends on the efficiency of financial intermediaries in adjusting their portfolios. On the other hand, when financial intermediaries optimize their portfolio choices, they take into account the impact of their behavior on the short rate passthrough to the real economy, which in turn influences future short rate movements.

To grasp the concept of general equilibrium in this model, it is essential to explore the relationship between the bond price sensitivity $\hat{A}$ and policy persistence κ_r across both the macroeconomic and financial markets equilibrium. In the macroeconomic equilibrium, this relationship is positive. When $\hat{A}$ increases, it indicates that bond prices become more sensitive to risk factors. A higher $\hat{A}$ makes the short-term interest rate more effective in influencing economic output through the Euler equation, which links interest rates to con-

[20] A stable solution requires that Υ has two positive eigenvalues and two negative eigenvalues. Ray (2019) provides a complex discussion about determinacy of this system.

sumption decisions. As a result, monetary policy becomes more potent in stimulating the economy, leading to a faster reversion of the policy rate to its mean. This means that κ_r, the speed at which the policy rate reverts to its mean, is larger when $\hat{A}$ is larger. In contrast, the financial markets equilibrium suggests a negative relationship between $\hat{A}$ and κ_r. Arbitrageurs profit by exploiting changes in risk factors, such as interest rate fluctuations. When κ_r is larger, it implies that short rate shocks are more transitory, meaning that such shocks quickly dissipate and return to their mean. The opportunities for arbitrageurs to profit from carry trades (strategies that exploit differences between short rate and other rates) diminish. As a result, arbitrageurs engage in fewer portfolio adjustments, which weakens the connection between bond prices and risk factors. Consequently, $\hat{A}$ is smaller when κ_r is larger.

The general equilibrium is achieved when there is a pair of $\hat{A}$ and κ_r that satisfies of both macroeconomic and financial markets equilibrium. The macroeconomic consistency states that in general equilibrium, the Taylor rule must be compatible with the dynamics of financial markets. This means that monetary policy should evolve in a manner consistent with the optimal portfolio decisions of investors, reflecting their expectations about the economy and interest rates. On the other hand, arbitrageurs must internalize the effects of their portfolio adjustments on the broader macroeconomy. Their actions influence bond prices, which in turn affect the short rate and broader economic conditions. In general equilibrium, bond prices and the resulting short rate dynamics must align with what arbitrageurs expect.

4.3 A Simple Case

In this section, I present a simplified version of the model where analytic solutions for the affine coefficients are attainable. This simplified case provides insight into how the general equilibrium is determined and highlights the effects of key model innovations. All proofs and derivations are in Appendix A. The assumptions made to simplify the model are:

1. *Price is fully rigid so that there is no inflation, i.e.,* $\pi_t = 0$.
2. *The habitat demand is price inelastic, i.e.,* $\alpha(\tau) = 0$ *for all* τ.
3. *There is no demand risk, i.e., short rate is the only state variable.*
4. *The aggregate nominal interest rate has only a yield component, with uniform weight across the maturities, i.e.,* $\eta^i(\tau) = \eta^i = 1/T$ *and* $\eta^R(\tau) = 0$.

In this simplified case, we derive a closed-form solution for $A_r(\tau)$. While $C(\tau)$ can potentially be solved in closed form if the loading function of the demand shifter is specified with certain formats, it is not central to the analysis here. Therefore, we abstract from solving $C(\tau)$ in this discussion.

LEMMA 1. *(Affine coefficient, simple case). In the simple case, given κ_r, the affine coefficient $A_r(\tau)$ can be expressed as:*

$$A_r(\tau) = \frac{1 - e^{-\kappa_r \tau}}{\kappa_r}.$$

Without inflation, the linearized macroeconomic system simplifies to:

$$\Upsilon = \begin{bmatrix} \psi_r & -\psi_r \phi_x \\ -\varsigma^{-1} \hat{A}_r & 0 \end{bmatrix},$$

where

$$\hat{A}_r = \frac{1}{T} \int_0^T \frac{1 - e^{-\kappa_r \tau}}{\kappa_r \tau} d\tau. \tag{34}$$

Equation (34) shows that $\hat{A}_r$ and κ_r are negatively correlated. When $\kappa_r = 0$ (i.e., the short rate shocks are permanent), $\hat{A}_r = 1$. This represents a scenario where the aggregate interest rate perfectly reflects the short rate, indicating perfect monetary policy transmission. As $\kappa_r > 0$, indicating that the short rate shock reverts to its long-run average, $\hat{A}_r$ becomes smaller than 1. This means that long-term yields and the aggregate nominal interest rate under-react to short-term rate changes. In other words, with a higher κ_r , the short rate shocks are more transitory, leading to less pronounced changes in the expected future short rate and, consequently, a less responsive aggregate nominal interest rate. The negative correlation between $\hat{A}_r$ and κ_r reflects the diminished effect of monetary policy on long-term yields when the policy rate changes are more transient. As κ_r increases, the term structure of Treasury yields becomes less responsive to short-term rate movements, thereby affecting the overall transmission of monetary policy to the aggregate nominal rate.

LEMMA 2. *(Macroeconomic equilibrium solution, simple case). Given $\hat{A}_r$, the solution* *of the macroeconomic equilibrium is*

$$dr_t = -\kappa_r(r_t - r^{ss})dt + \sigma_r dB_{r,t},$$
$$x_t = \omega_x(r_t - r^{ss}),$$

where

$$\frac{\kappa_r(\kappa_r - \psi_r)}{\psi_r\phi_x\varsigma^{-1}} = \hat{A}_r, \tag{35}$$
$$\omega_x = \frac{\psi_r - \kappa_r}{\psi_r\phi_x}.$$

From equation (35), $\hat{A}_r$ and κ_r are positively related once κ_r exceeds $\psi_r/2$. When monetary policy is effective (high $\hat{A}_r$), a conventional expansion (such as a decrease in the short rate) has a powerful impact on output. This effectiveness causes the policy to be in place for a shorter duration because the output responds strongly, prompting a quicker adjustment of the policy rate to stabilize the economy. This mechanism can be easily seen from the Taylor rule (5). According to the Taylor rule, if an expansionary policy significantly boosts output, there is a concern of overheating. This concern leads to a prompt response by increasing the short rate to stabilize the economy, which in turn decreases the policy persistence (i.e., κ_r increases). Thus, a higher κ_r corresponds to a higher $\hat{A}_r$ in equilibrium.

LEMMA 3. *(General equilibrium, simple case). The general equilibrium is a pair of $\hat{A}_r$ and κ_r that satisfies both equations (34) and (35).*

Equation (34) characterizes a negative relation between $\hat{A}_r$ and κ_r when the financial markets are in equilibrium. Conversely, equation (35) characterizes a positive relation between $\hat{A}_r$ and κ_r when the macroeconomy is in equilibrium. The general equilibrium is obtained when both the financial markets and the macroeconomy are in equilibrium. The different signs in the correlations ensure the existence of a pair of $\hat{A}_r$ and κ_r that defines the general equilibrium as shown by figure 6. Red dashed line summarizes the positive relation in the macroeconomic equilibrium and blue solid line the negative relation in the financial market equilibrium. The general equilibrium is found at the intersection of these two curves, represented by the black cross. This intersection defines the unique pair ($\hat{A}_r$,κ_r) where both the financial market and the macroeconomic systems are simultaneously in equilibrium.

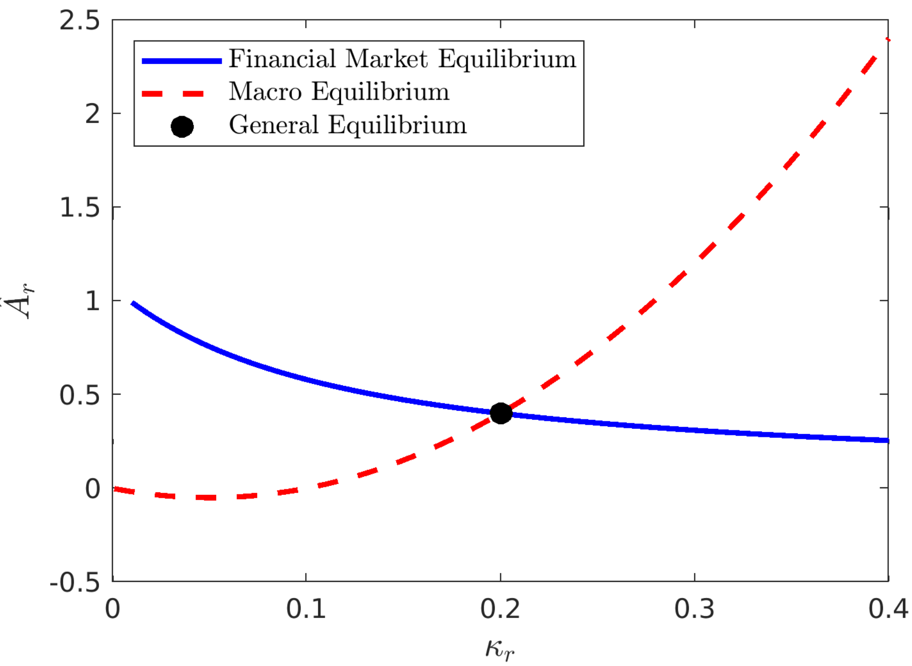


Figure 6: General Equilibrium in the Simple Case.

Note: Parameters are set as follows: $\varsigma = 1$, $\psi_r = 0.1$, $\phi_x = 0.5$.

4.3.1 Extension 1 - Balance Sheet Cost

In this extension of the simple model, I introduce a balance sheet cost associated with holding assets, which modifies the equilibrium dynamics. The main assumptions for this extension are:

1. Constant price elasticity of habitat demand: $\alpha(\tau) = 1$.

2. Balance sheet cost: $\lambda > 0$.

The inclusion of a balance sheet cost introduces positive non-pecuniary cost ($\lambda > 0$) that affects the excess return on Treasury investments. This means that even without risk aversion or demand risk, bond prices deviate from the expectations hypothesis purely due to the balance sheet cost channel, which operates through the elastic habitat demand. To maintain analytical tractability, I assume the bond price elasticity $\alpha(\tau)$ is consistent across all maturities.

LEMMA 4. *(Affine coefficient, simple case extension 1).* *In extension 1, given κ_r, the* *affine coefficient $A_r(\tau)$ is*

$$A_r(\tau) = \frac{1 - e^{-(\lambda+\kappa_r)\tau}}{(\lambda + \kappa_r)}.$$

The linearized macroeconomic system is same as in the simple case:

$$\Upsilon = \begin{bmatrix} \psi_r & -\psi_r \phi_x \\ -\varsigma^{-1} \hat{A}_r & 0 \end{bmatrix},$$

where

$$\hat{A}_r = \frac{1}{T} \int_0^T \frac{1 - e^{-(\lambda+\kappa_r)\tau}}{(\lambda + \kappa_r)\tau} d\tau. \tag{36}$$

Equation 36 shows that the relationship between $\hat{A}_r$ and κ_r remains negative even with the introduction of the balance sheet cost. However, the presence of λ lowers the intermediary efficiency $\hat{A}_r$ for any maturity compared to the basic simple case. This reduction occurs because the balance sheet cost makes carry trades more expensive, diminishing the responsiveness of bond prices to short rate changes. Despite the introduction of λ, the macroeconomic equilibrium continues to be summarized by the same equations as in the basic simple case. The negative correlation between $\hat{A}_r$ and κ_r in the financial market equilibrium and the positive correlation in the macroeconomic equilibrium still ensure the existence of a general equilibrium. The general equilibrium in this extended model will be characterized by a lower κ_r compared to the basic case. The reasoning is that the balance sheet cost reduces intermediary efficiency, making the short rate less effective in stimulating output. Consequently, the central bank needs to maintain policy changes for a longer period, leading to a smaller general equilibrium value of κ_r. Figure 7 illustrates the impact of the balance sheet cost on the general equilibrium. The graph shows the modified relationships between $\hat{A}_r$ and κ_r under the new assumptions, highlighting how the balance sheet cost lowers the equilibrium value of $\hat{A}_r$ and κ_r.

4.3.2 Extension 2 - Dealers' Market Power

In this extension, I explore a scenario where dealers benefit from economies of scale by expanding their holdings of Treasuries, in contrast to the balance sheet cost introduced in Extension 1. The key assumptions for this extension are:

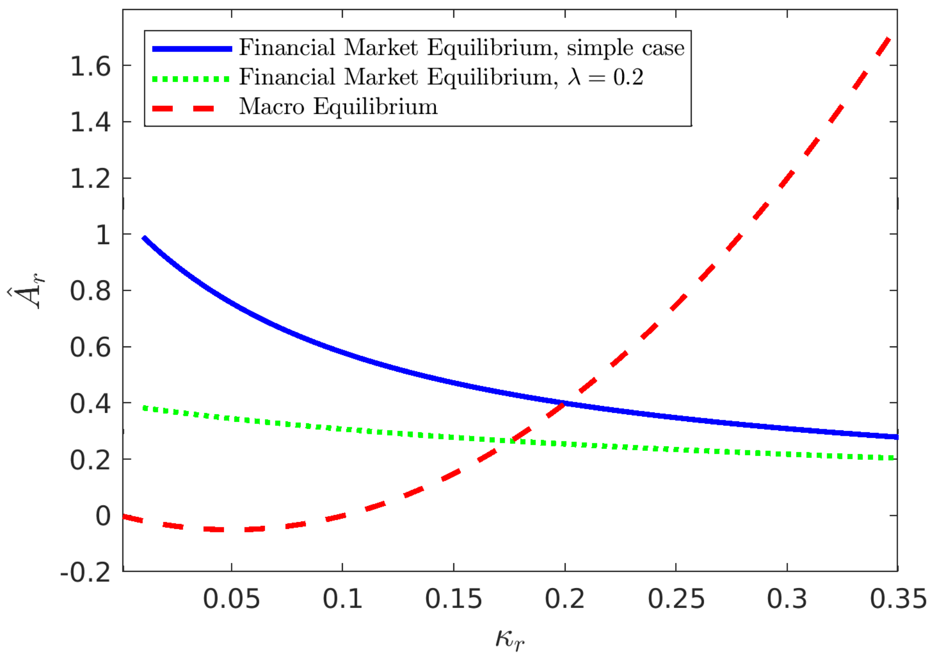


Figure 7: General Equilibrium in Extension 1.

Note: Parameters are set as follows: $\varsigma = 1$, $\psi_r = 0.1$, $\phi_x = 0.5$.

1. Constant price elasticity of habitat demand: $\alpha(\tau) = 1$.

2. Dealers achieve economies of scale by expanding holdings: $\lambda < 0$.

In this setup, the non-pecuniary value of holding Treasuries ($\lambda < 0$)effectively reduces the required excess return for investors. Holding Treasuries brings market power benefit to investors. This reflects the market power benefit that dealers gain by holding Treasuries, which is in addition to any pecuniary profit. The assumption of constant price elasticity is retained to keep the derivation of the affine coefficient tractable.

LEMMA 5. *(Affine coefficient, simple case extension 2). In extension 2, given κ_r, the affine coefficient $A_r(\tau)$ is the same as in Lemma 3. What's more, the intermediary efficiency $\hat{A}$ is also the same as in equation (36).*

Lemma 5 reveals that the solution for the affine coefficient $A_r(\tau)$ remains unchanged from Extension 1. The macroeconomic equilibrium is also identical to that in Extension 1. Consequently, the general equilibrium solution can be directly applied from the previous extension. However, this similarity in solutions does not imply that the equilibrium variables are the same as in Extension 1. In this extension, λtakes a negative value, which influences

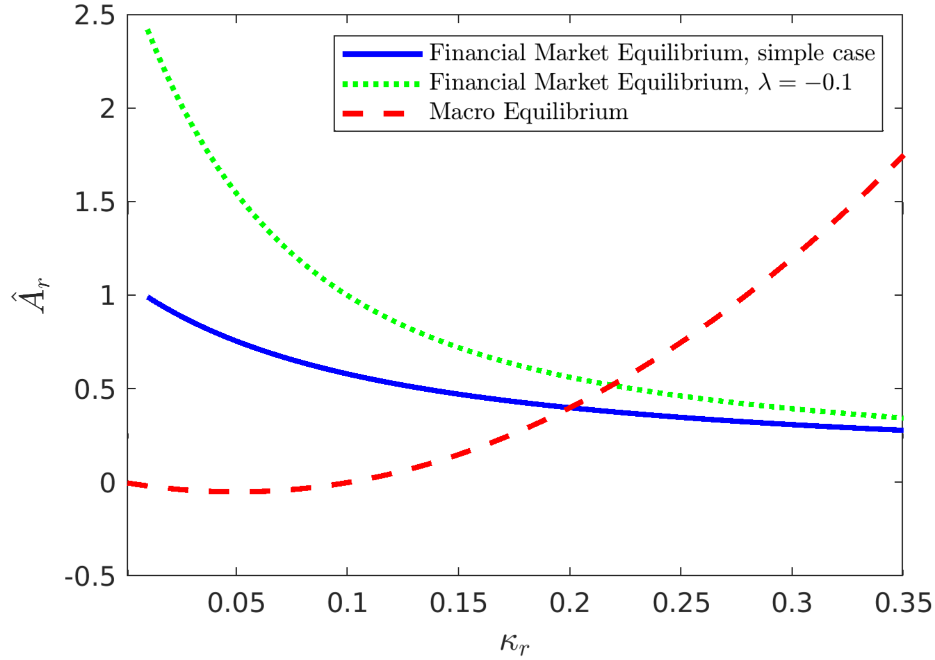


Figure 8: General Equilibrium in Extension 2.

Note: Parameters are set as follows: $\varsigma = 1$, $\psi_r = 0.1$, $\phi_x = 0.5$.

the model in a specific way. Since $A_r(\tau)$ is inversely related to λ, a negative λ increases $A_r(\tau)$ across all maturities. This is because arbitrageurs, who now gain market power benefits in addition to pecuniary profits from carry trades, are more effective in transmitting changes in the short rate to bond prices. As a result, when the financial intermediary is more efficient, a given short rate drop has a larger impact on boosting output. Consequently, the policy shock does not need to be sustained for as long, allowing it to revert back to its mean level more quickly. Figure 8 compares the general equilibrium in the basic simple case with the case where $\lambda = -0.1$. The figure illustrates how the enhanced intermediary efficiency (larger $\hat{A}_r$) due to dealers' market power shifts the general equilibrium, leading to a faster reversion of the policy rate (larger κ_r).

4.3.3 Extension 3 - Repo Importance

In this final exercise, I relax the assumption regarding the composition of the aggregate nominal rate $\tilde{r}_t$:

1. The aggregate nominal interest rate $\tilde{r}_t$ now has both yield and repo components. The relative weight is governed by the parameter η, i.e., $\eta^i(\tau) = \frac{1-\eta}{T}$ and $\eta^R(\tau) = \frac{\eta}{T}$.

The weights of bond yields and repo rates sum to 1. Which component is with more influence on the aggregate nominal rate is determined by the parameter η. When $\eta < 0.5$, bond yields are more important than repo rates in determining the aggregate nominal rate. Oppositely, when $\eta > 0.5$, repo rates become more important than bond yields. Since bond yields and repo rates have different compositions of excess returns, they also respond differently to short rate shocks. Therefore, the efficiency of financial intermediaries, represented by $\hat{A}_t$, depends on the relative importance of these two rates.

LEMMA 6. *(Affine coefficient, simple case extension 3). In extension 3, given κ_r, the affine coefficient $A_r(\tau)$ is the same as in Lemma 3. However, the intermediary efficiency $\hat{A}_t$ is now given by:*

$$\hat{A}_r = \frac{1-\eta}{T}\int_0^T \frac{1-e^{-\kappa_r\tau}}{\kappa_r\tau}d\tau + \eta. \tag{37}$$

The macroeconomic equilibrium in this extension is the same as in extension 1. In equation (37), the term $\frac{1-e^{-\kappa_r\tau}}{\kappa_r\tau}$ is always less than 1 for any τ, leading to the inequality $\int_0^T \frac{1-e^{-\kappa_r\tau}}{\kappa_r\tau}d\tau < T$. Therefore, $\int_0^T \frac{1-e^{-\kappa_r\tau}}{\kappa_r\tau}d\tau/T < 1$, implying that $\hat{A}_r$ is increasing in η. Letting $eta = 0.5$, figure 9 illustrates how both $\hat{A}_r$ and κ_r increase relative to the basic simple case. This outcome can be understood as follows. Repo rates are typically more sensitive to changes in the short rate compared to bond yields. Thus, when a higher portion of lending contracts is indexed by repo rates (η is larger), the short rate achieves better passthrough to the aggregate nominal rate. With improved passthrough, the short rate policy is more effective at boosting output. Consequently, the policy can revert to its mean level more quickly, resulting in a higher equilibrium value of κ_r.

5 Explanations for GFC and Covid-19 Observations

In this section, I illustrate how the model can explain the differentiated observations seen during the Global Financial Crisis (GFC) and the COVID-19 pandemic. During the GFC, the financial system experienced a significant net demand shock for short-term Treasuries. In contrast, the COVID-19 pandemic brought about a net supply shock for long-term Treasuries. The model captures the distinct dynamics in Treasury markets during these two crises, driven by different types of shocks.

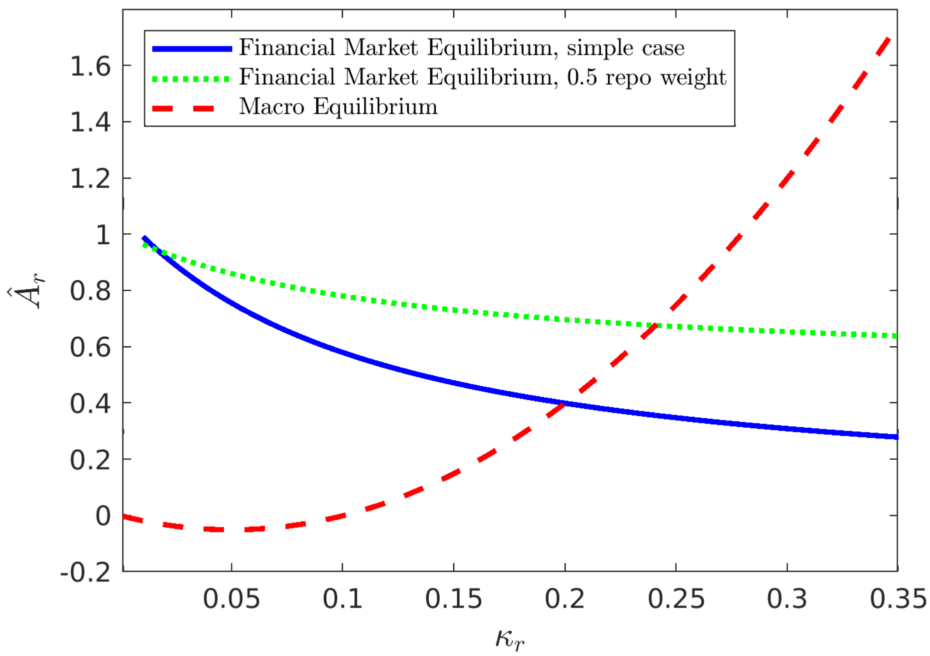


Figure 9: General Equilibrium in Extension 3.

Note: Parameters are set as follows: $\varsigma = 1$, $\psi_r = 0.1$, $\phi_x = 0.5$.

5.1 Flight-to-Liquidity in GFC

I posit that there was a pronounced flight-to-liquidity, particularly evident in the surge in demand for Treasury bonds, especially those with shorter maturities [21]. This phenomenon can be explained by the behavior of habitat investors in response to the changing risk environment during the crisis. The collapse of major financial institutions, such as Bear Stearns and Lehman Brothers, amplified investors' concerns about asset safety and liquidity. In response, there was a strong preference for assets that were not only safe but also highly liquid, like short-term Treasuries.

The increased demand for these safe assets is represented by a negative value of the private habitat demand shifter ($\beta_t < 0$). In the context of the model, this increased demand for Treasuries created an endogenous non-pecuniary value. For dealers, holding Treasuries during this period provided a convenience yield, as it allowed them to avoid the potential costs associated with short selling, such as reputation loss and regulatory penalties. This convenience benefit made it attractive for dealers to hold Treasuries despite the low pecuniary returns.

[21] He, Nagel and Song (2022) documents "a large increase (by about $350 billion) in foreign investors' holdings of long- term Treasuries since July 31, 2007".

Figure 10 plots model-implied IRFs to an exogenous demand shock for short-term Treasuries. In this exercise, I replace the Taylor rule with a simple Ornstein-Uhlenbeck process for the short rate process with the mean-reversion speed κ_r and volatility σ_r. This process is more suitable for the short observation windows, as shown in figure 3, during which the central bank did not have time to adjust the short rate. Most of the term-structure parameter values are taken from Vayanos and Vila (2021). The parameter λ is set to 0.1, indicating a 10% marginal non-pecuniary benefit for holding Treasuries, which reflects the market conditions during the crisis where holding Treasuries provided additional safety and liquidity. The demand shock is assumed to be 40% of the Treasuries held by dealers, aligning with empirical observations during the GFC, documented in figure 3.

Panel A plots the instantaneous response in the Treasury yield curve. The repo wedge, defined as the spread between the repo rate and the short rate, is depicted by panel B. In this model, dealers lend cash to hedge funds at the repo rate and finance lending at the short rate. Therefore, the repo-short rate spread measures the (in)convenience yield in the Treasury repo market. Panel C shows the Treasury-OIS spread, calculated as the difference between Treasury yields and OIS rates, reflecting the additional premium investors were willing to pay for holding Treasuries over risk-free contracts. To derive the Treasury-OIS spread, I first construct the OIS rates which are the returns of investing in contracts with the same cash flows as for Treasuries but do not involve the ownership of the securities. Specifically, the shadow price of the OIS contract is defined as the equilibrium bond price with $\lambda = 0$, i.e.,

$$P_t^{OIS}(\tau) = exp(A_r^{OIS}(\tau)r_t + A_\beta^{OIS}(\tau)\beta_t + C^{OIS}(\tau)),$$

where $A_r^{OIS}(\tau)$, $A_\beta^{OIS}(\tau)$, and $C^{OIS}(\tau)$ are affine coefficients with $\lambda = 0$. It follows that the OIS rate is equal to

$$i_t^{OIS}(\tau) = -\frac{1}{\tau} log P_t^{OIS}(\tau).$$

Panel D shows dealers' aggregate balance sheet size, the sum of the direct and repo holding of Treasuries. Finally, panel E and F depict responses in output gap and inflation rate, respectively.

The model predicts that an increase in demand leads to a rise in their prices, which in turn causes a decline in their yields. Since the demand shock is concentrated at the short end of the maturity spectrum, the yields for short-term Treasuries drop significantly.

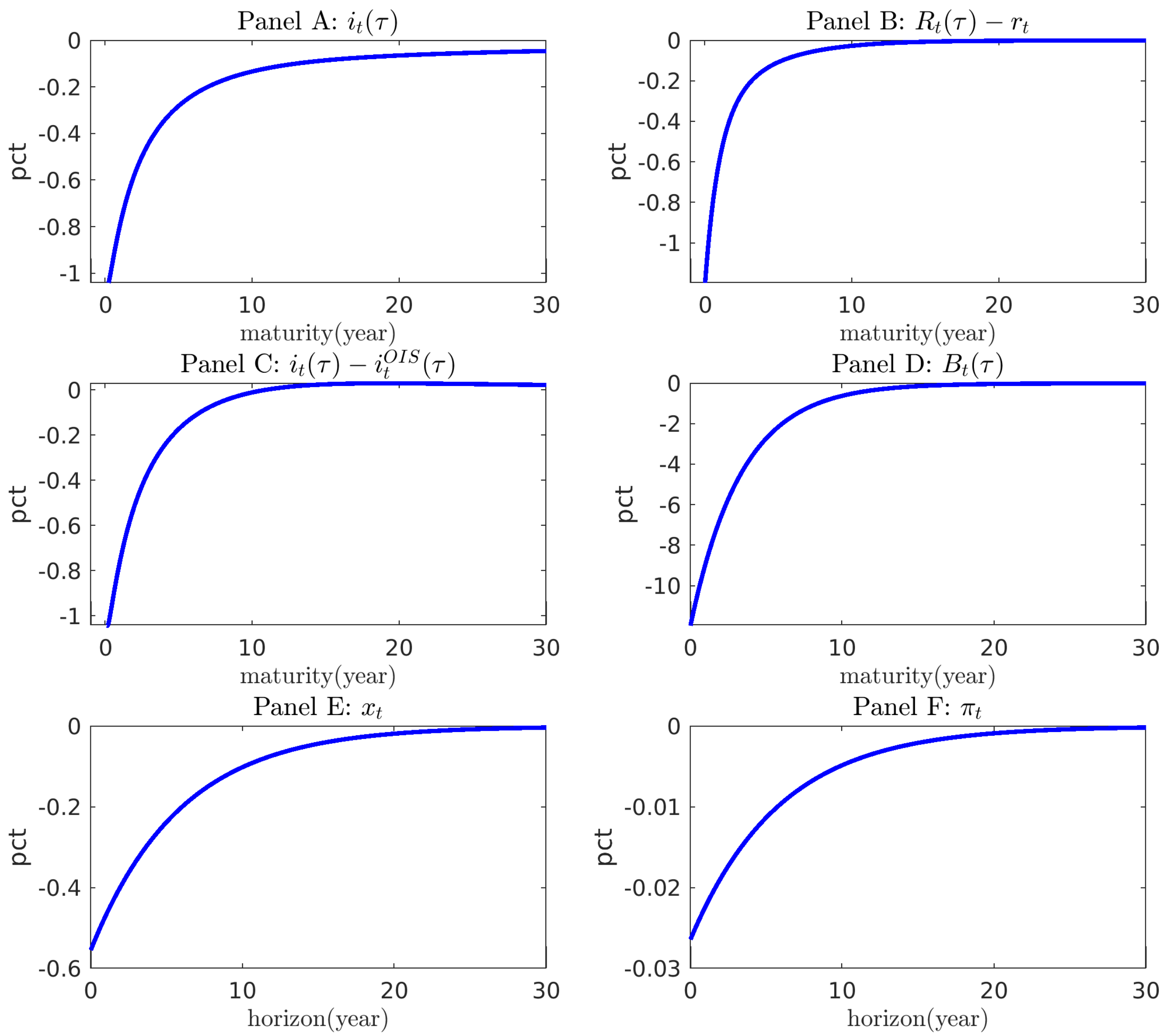


Figure 10: Model-implied responses to a demand shock concentrated at short maturity.

Note: The demand shock is 40% of Treasuries held by dealers. Parameters are set as follows: $\kappa_{r\beta} = 0$, $\kappa_r = 0.125$; $\sigma_r = 0.0146$; $\kappa_\beta = 0.17$; $\sigma_\beta = 0.0146$; $\alpha(\tau) = 5.21 * exp(-0.297 * \tau)$, $\theta(\tau) = 15 * exp(-0.297 * \tau)$, $\rho = 7$, $\lambda = 0.1$,$\eta^i(\tau) = 1/30$, $\eta^R(\tau) = 0$, $\chi = 0.04$, $\delta = 0.01$, and $\varpi = -0.3$.

he long end of the yield curve remains mostly unaffected because the demand shock does not target these maturities. The difference between long-term and short-term yields (term spread) increases, reflecting a steeper yield curve shown in panel A. With higher demand for short-term Treasuries, these securities become more valuable as collateral in the repo market. The repo rates, particularly for short-term Treasuries, decrease because the increased demand lowers the cost of borrowing against these securities. This is captured by the larger responses in repo rates with short-term underlying Treasuries in panel B.

Dealers respond to the demand shock by increasing the supply of Treasuries through direct sales or by lending them via repo agreements. As dealers provide more Treasuries, the amount of naked short selling increases, raising the marginal cost of such activities. The heightened scarcity of Treasuries and the increased scale of short selling amplify the non-pecuniary value of holding Treasuries, especially due to their liquidity and safety benefits. The convenience yield, represented by the negative Treasury-OIS spread [22], becomes more pronounced. This can be seen from panel C, with the spread reflecting the non-pecuniary value of Treasuries, making their yields lower compared to OIS rates.

Based on the previous analysis, the demand shock causes a decrease in the Treasury yields and repo rates. The drop in Treasury yields and repo rates leads to a lower aggregate nominal rate, which typically discourages saving and encourages spending. In a standard model, lower interest rates would boost output and inflation by stimulating consumption. However, the model incorporates a second channel - the quality preference - that increases households' marginal utility of holding savings products. If the quality preference channel is strong enough, it can lead to a decline in macroeconomic variables (output and inflation), even in the face of lower borrowing costs, as in panel E and F.

5.2 Flight-From-Safety in Covid-19

Unlike the GFC, where a flight-to-liquidity led to lower yields and a negative Treasury-OIS spread, the COVID-19 period saw higher Treasury yields, higher repo wedges, and positive Treasury-OIS spreads, signaling a shift in market dynamics and investor preferences. I posit that there was a large supply shock to long-term Treasuries resulted in rising long-term yields.

The increased supply is represented by a positive value of the private habitat demand

[22] The non-pecuniary value pushes Treasury price to be higher than the OIS shadow price and yield to be lower than the OIS rate. That being said, the convenience yield is represented by a negative Treasury-OIS spread.

shifter ($\beta_t > 0$). In the context of the model, this increased demand for Treasuries created an endogenous non-pecuniary cost. For dealers, holding Treasuries during this period involved an inconvenience yield, as the higher supply of long-term Treasuries forced dealers to expand their balance sheets, which incurred a higher balance sheet cost. Thus, higher excess returns are required to compensate the extra balance sheet cost.

I use the model to general IRFs to a supply shock as shown in figure 11. The increase in supply predominantly targets long-term Treasuries. The presented variables in figure 11 are constructed in the same way as for the GFC. Most of the parameter values remain unchanged, with the exception of the demand shock location function $\theta(\tau)$ shifting from short to long end concentrated. The non-pecuniary cost parameter λ still takes the value of 0.1, but the interpretation is different. Specifically, $\lambda = 0.1$ implies a 10% marginal balance sheet cost when dealers expands the Treasury holdings to absorb the extra supply in the market. The size of the supply shock is about 75% of total Treasuries held by dealers to match the dynamics of primary dealers' repo holdings documented in figure 4.

Similar to the GFC, the yield curve steepens, but this time, the steepening is due to a different mechanism. The increase in supply predominantly targets long-term Treasuries. The influx of long-term Treasuries drives down their prices, leading to a significant increase in their yields. Conversely, the short end of the yield curve remains relatively unchanged due to stable short-term Treasury supply. The rise in long-term yields relative to short-term yields increases the term spread, causing the yield curve to steepen. Panel A depicts this result. The increased supply of long-term Treasuries reduces their scarcity, which diminishes the privilege of using these securities as collateral in the repo market. Consequently, the borrowing cost (repo rate) for long-term Treasuries rises. Although the supply shock mainly targets long maturities, short maturity bond prices also drop (see panel A). lower short-term Treasury prices increase habitat demand, which reduces the available collateral in the repo market, thereby lowering the repo rate. That explains the different signs of the responses in the repo rates, depicted by panel B.

Panel C replicates the dynamics in Treasury-OIS spread during Covid-19. The higher supply of long-term Treasuries forces dealers to expand their balance sheets, which incurs a higher non-pecuniary cost (interpreted as a balance sheet cost). This makes holding Treasuries less convenient, leading to a more positive Treasury-OIS spread (indicating an inconvenience yield. For short-term Treasuries, the lower prices increase habitat demand, which reduces the dealers' holdings and alleviates the balance sheet pressure. This causes the Treasury-OIS spread to become more negative for short maturities. Such mechanism is

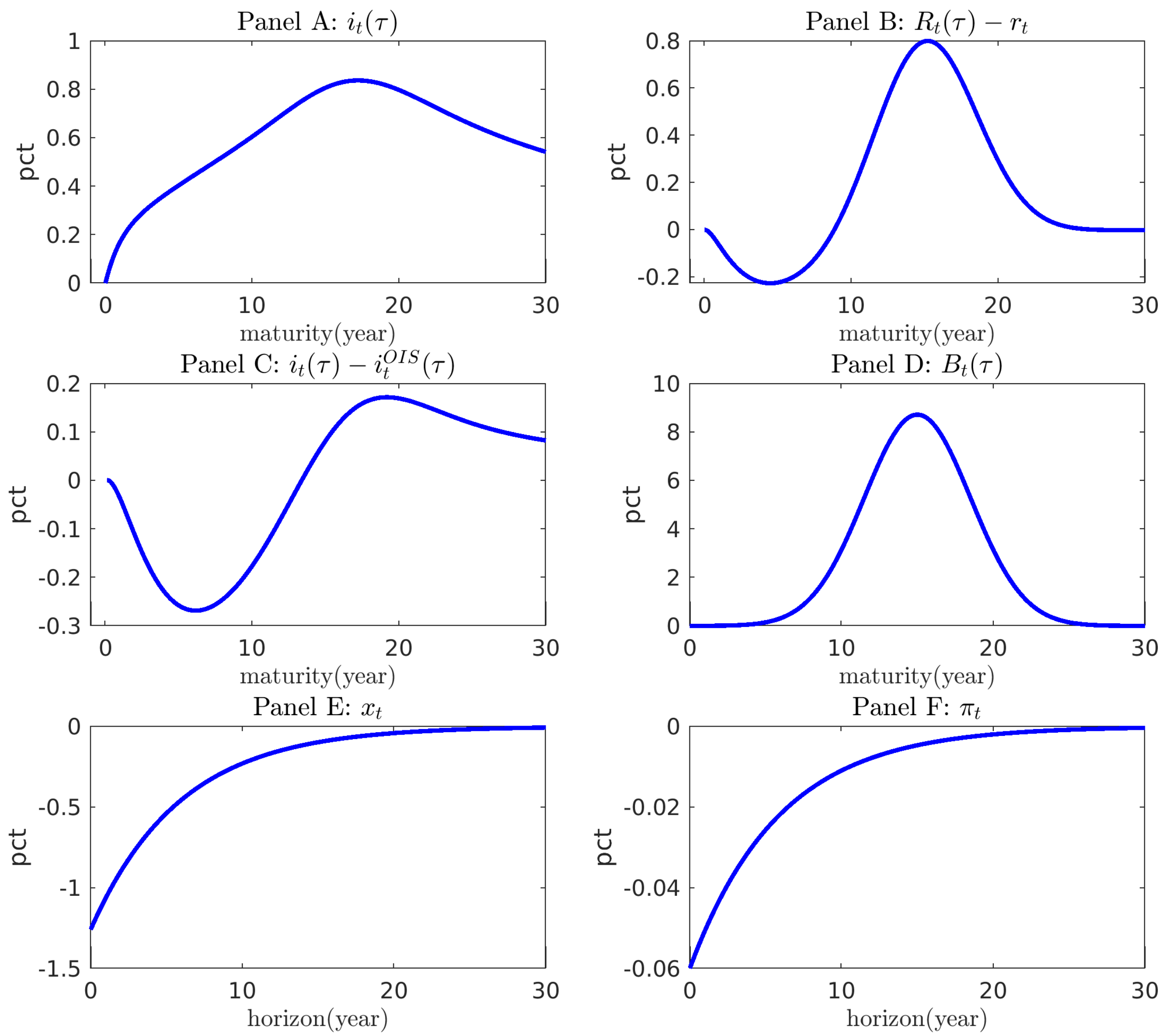


Figure 11: Model-implied instantaneous responses to a drop in habitat demand concentrated at long maturity.

Note: The demand shock is 75% of Treasuries held by dealers. Parameters are set as follows: $\kappa_{r\beta} = 0$, $\kappa_r = 0.125$; $\sigma_r = 0.0146$; $\kappa_\beta = 0.17$; $\sigma_\beta = 0.0146$; $\alpha(\tau) = 5.21 * exp(-0.297 * \tau)$, $\theta(\tau) = 51 * normpdf(15, 3.5)$, $\rho = 7$, $\lambda = 0.1$,$\eta^i(\tau) = 1/30$, $\eta^R(\tau) = 0$, $\chi = 0.04$, $\delta = 0.01$, and $\varpi = -0.3$.

strong for short maturities and weak for long maturities [23], therefore only the short end of the Treasury-OIS spread exhibits a negative response.

The last two panels E and F provide the responses in output gap and inflation rate. The increase in bond yields and repo rates translates to higher overall borrowing costs for households, which typically reduces consumption. On the other hand, the extra supply of Treasuries diminishes the quality preference for saving products, thus reducing the marginal utility of saving and potentially encouraging consumption. In this scenario, the dominant effect is the first channel - higher borrowing costs - which leads to a decrease in output and inflation. This reflects typical recessionary dynamics where economic activity contracts in response to increased financial frictions.

6 Calibration

In this section, I explain the methods used to calibrate the model's parameters, ensuring that the model accurately reflects observed economic behavior. The parameters are grouped into three categories: those affecting the aggregate borrowing rate, macroeconomic dynamics, and term structures. Table 1 summarizes the calibration results, where each parameter's value is chosen to match specific empirical moments or targets. The parameters $\eta^i(\tau)$ and $\eta^R(\tau)$ determine the relative importance of bond yields and repo rates in the overall borrowing cost. The macroeconomic dynamics parameters consist of: the preference parameters, namely the discount factor and the intertemporal substitution elasticity (χ and ς^{-1}); price rigidity δ; Taylor rule inertia (ψ_r)[24]; and the short rate volatility (σ_r). The term structure parameters consist of: the demand factor location function ($\theta(\tau)$); the habitat demand price elasticity ($\alpha(\tau)$); arbitrageurs aggregate risk aversion ($\frac{1}{\rho_h+\rho_d}$); demand shock volatility (σ_β); demand shock inertia (κ_β); the marginal non-pecuniary cost (λ), and the safety preference ϖ.

6.1 Yield and Repo Weights

The weighting functions $\eta^i(\tau)$ and $\eta^R(\tau)$ are critical in determining the relative significance of bond yields and repo rates in calculating the overall borrowing cost. To estimate $\eta^i(\tau)$, I align it with the average maturity distribution of Treasury securities held by the public between

[23] This can be seen from the functional format of $\alpha(\tau)$. It is a decreasing function in τ.

[24] In fact, Taylor rule also involves the output and inflation weight parameters ϕ_x and ϕ_π. Later I discuss the under-identification of these two parameters and that they do not affect key equilibrium objectives in this model.

Table 1: Calibration Results.

Parameter	Value	Description	Target
Aggregate Borrowing Rate			
$\eta^i(\tau)$	figure 4	Bond yield weight	Treasury maturity distribution
$\eta^R(\tau)$	figure 4	Repo rate weight	MMF repo collateral maturity dist
Macroeconomic Dynamics			
χ	0.04	Discount factor	Long-run interest rate
ς^{-1}	1	Intertemporal subs elasticity	Balanced growth
δ	0.1354	Price rigidity	CEE(2005) IRF of π_t to r_t
ψ_r	0.2236	Short rate policy inertia	CEE (2005) IRF of x_t to r_t
σ_r	0.0218	Short rate risk volatility	Var(r_t)
Term Structures			
$\theta(\tau)$	figure 5	Demand factor location	LSAP1 targets
$\alpha(\tau)$	$5.21e^{-0.297\tau}$	Habitat demand elasticity	Vayanos & Vila (2021)
$\rho = \frac{1}{\rho_h+\rho_d}$	6	Risk aversion	FFR 2-year yield inst. response
σ_β	0.027	Demand shock volatility	LSAP1 10-year yield inst. response
κ_β	0.22	Demand shock inertia	LSAP1 10-year yield response half-life
λ	0.55	Marginal non-pecuniary cost	LSAP1 avg. repo rate inst. response
ϖ	-0.36	Quality preference	$Var(x_t)/Var(r_t)$

Note: Parameters in the macro dynamics and term structures groups are calibrated jointly within each group. The parameters are listed alongside the target which is most sensitive to the given parameter.

1985 and 2007. This period was characterized by low financial risk and minimal market friction, making it ideal for calibrating macroeconomic parameters under such assumptions. The data is sourced from the CRSP U.S. Treasury database, which provides comprehensive details on bond characteristics, including issue dates, maturity dates, face values, market prices, and coupon rates. Using monthly data, I apply Gaussian kernel density estimation to derive the maturity distribution of Treasuries. The resulting distribution, shown in Figure 12, highlights that the maturity structure of public-held Treasuries is skewed toward shorter maturities, with the highest density around 2-year bonds. A smaller peak is observed around 8-9 years, indicating some preference for mid- to long-term bonds, though long maturities occupy a smaller portion of the market.

Similarly, I estimate $\eta^R(\tau)$ by matching the average maturity structure of Treasury securities used as collateral in repos by Money Market Funds (MMFs) [25]. Data for estimating

[25] "The role of MMFs as cash investors in repos has increased over the last 20 years. As of September 30, 2020, the Financial Accounts of the United States show that MMFs accounted for close to 22% of the total repo assets." See Baklanova, Kuznits and Tatumetal (2021) for a comprehensive analysis of MMFs participation on repo market.

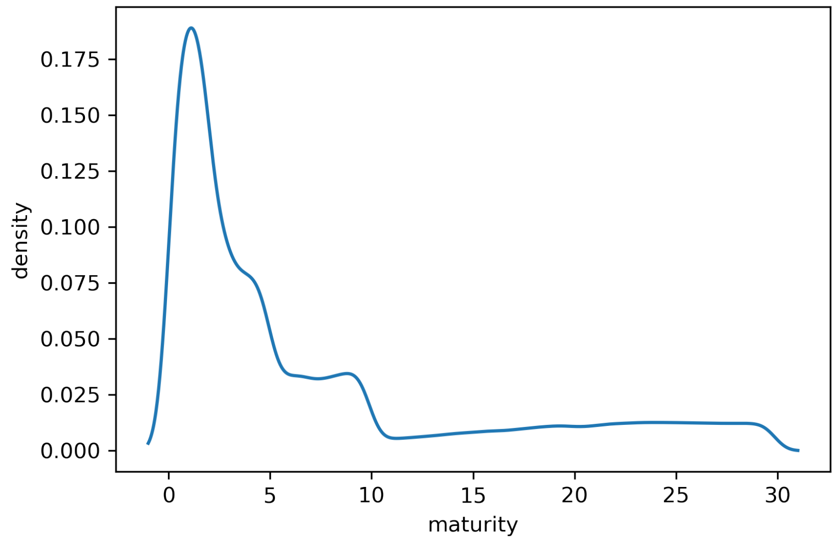


Figure 12: Estimation of bond yield weights $\eta^i(\tau)$.

Note: The curve represents the average density of the remaining maturity for all treasury securities outstanding during the period of 1985-2006.

$\eta^R(\tau)$ comes from the N-MFP database[26], which is publicly available and provides detailed information on repo collateral, including asset types, face values, market prices, and maturity details. I use quarterly data from 2011 to 2018 to exclude distortions from the pandemic period. The Gaussian density estimation, as shown in Figure 13, reveals a more dispersed maturity distribution for repos than for bond yields. The density peaks are around 5-year and 26-year maturities, indicating a broader range of maturities used as collateral in repo transactions by MMFs.

6.2 Macroeconomic Parameters

To streamline the calibration of macroeconomic variables, I temporarily disable the term structure dynamics. This simplification involves assuming no risk aversion ($\rho = \frac{1}{\rho_h+\rho_d} = 0$)[27] and setting the non-pecuniary value to zero ($\lambda = 0$), with the short rate as the sole risk factor ($S_t = r_t$). Additionally, considering that the repo weight was calibrated using post-GFC data, I assume that the aggregate nominal rate depends solely on bond yields ($\eta^R(\tau) =$

[26] Another non-public source: Federal Reserve Liquidity Monitoring Report (FR 2052) has daily data about BHCs' repo positions by collateral class and maturity.

[27] I replace $\frac{1}{\rho_h+\rho_d}$ with ρ hereafter.

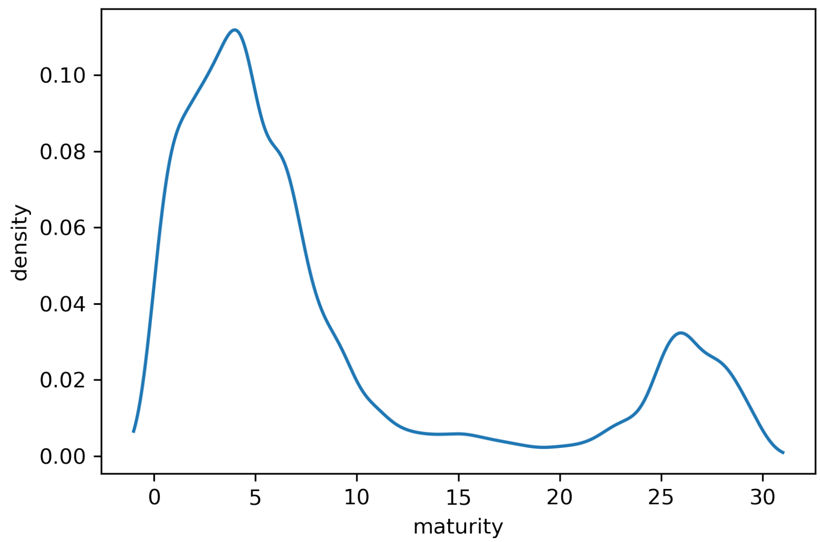


Figure 13: Estimation of repo rate weights $\eta^R(\tau)$.

Note: The curve represents the average density of the remaining maturity for treasury securities used as collateral in a repo transaction by MMFs during the period of 2011-2018.

0). These assumptions remove the complexities associated with term structure parameters, thus facilitating a more straightforward calibration of the macroeconomic parameters. The discount factor χ is set to 0.04 to match the long-run interest rate. The intertemporal substitution elasticity ς^{-1} is set to 1 to maintain a balanced growth path, reflecting standard economic assumptions.

The taylor rule paramters for inflation (ϕ_π) and output (ϕ_x) enter the Taylor rule as products $\psi_r\phi_\pi$ and $\psi_r\phi_x$. This can be easily seen by rewriting the Taylor rule as follows:

$$\mathrm{d}r_t = -(\psi_r r_t + \psi_r\phi_\pi\pi_t + \psi_r\phi_x x_t + \psi_r r^*)\mathrm{d}t + \sigma_r\ \mathrm{d}B_{r,t}.$$

In this simplified model with a single risk factor, macroeconomic variables become linear functions of the short rate r_t: $x_t = \omega_{xr}(r_t - r^{ss})$ and $\pi_t = \omega_{\pi r}(r_t - r^{ss})$. Here, ω_{xr} and $\omega_{\pi r}$ represent contemporaneous responses of output and inflation to the short rate shock. The mean-reverting speed of the short rate, accounting for feedback from macroeconomic performance, is given by:

$$\kappa_r = \psi_r(1 + \phi_\pi\omega_{\pi r} + \phi_x\omega_{xr}).$$

Given this, only one of the parameters ψ_r, ϕ_x, or ϕ_π can be uniquely identified, as any values leading to the same κ_r will be indistinguishable from the output observables. For calibration, I choose $\pi_\pi = 3$ and $\pi_x = 2$ to satisfy the condition for a unique solution.

To calibrate the remaining macroeconomic parameters, I use a joint moments-matching approach, relying on IRFs from Christiano, Eichenbaum and Evans (2005). Specifically, I consider the instantaneous responses in the output gap and inflation rate to a conventional monetary policy shock, where a 1% decrease in the policy rate is associated with approximately a 0.8% reduction in the output gap and a 0.15% decline in the inflation rate. Additionally, I use the variance of the 3-month Treasury yield, calculated as 3.5013, as a target moment in the calibration process. The detailed steps for this calibration, along with sensitivity analyses, are provided in Appendix C1.

6.3 Term Structure Parameters

The final group of parameters consists of five scalar values and two functions that define the term structure. I begin by adopting the functional form for the habitat demand price elasticity, $\alpha(\tau)$, as proposed by Vayanos and Vila (2021). This specification suggests that habitat demand elasticity decreases with bond maturity. For the demand factor location, $\theta(\tau)$, I assume it follows the density function of a truncated normal distribution. The mean and standard deviation of this distribution are estimated to align with the maturity distribution of Treasury purchases under the first Large-Scale Asset Purchase program (LSAP1) between March 18 and October 31, 2009. Data from D'Amico and King (2013) inform this calibration. The resulting estimated demand location function, $\theta(\tau)$, is depicted in Figure 14, which shows a concentration of purchases targeting maturities of 7-8 years.

Given the weighting functions, macroeconomic parameters, habitat demand elasticity, and the demand factor location, I proceed to calibrate the remaining term structure parameters, except for the safety preference ϖ. In the baseline specification, I continue to assume that the aggregate nominal rate depends solely on bond yields ($\eta^R(\tau) = 0$). The four term structure parameters - ρ, σ_β, κ_β, and λ - are chosen simultaneously to match the observed responses of yields and repo rates to conventional monetary policy and LSAP1 shocks. Since LSAP1 is a government purchasing program, I assume it does not alter household valuations of saving products, so $\varpi = 0$. The optimal parameter values are obtained by minimizing the sum of squared errors between model-generated and actual data moments. Appendix C2 provides a sensitivity analysis for these four term structure parameters. Below, I outline the key calibration targets used in this exercise.

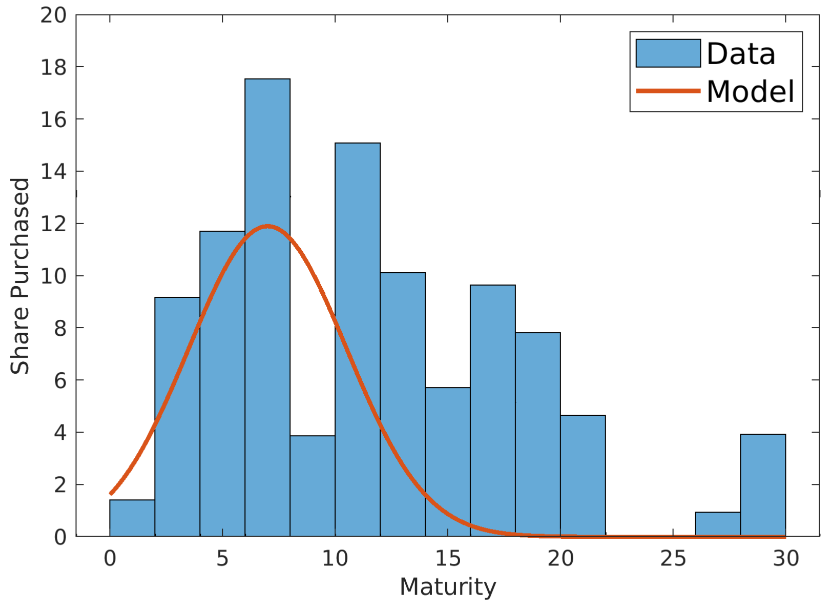


Figure 14: Estimation of demand factor location.

Note: Truncated normal distribution fitting of the share of maturity purchased during the LSAP1 between March 18 and October 31, 2009.

First, Swanson (2021) estimates the yield curve's response to FFR shocks using data from 1991-2019. When FFR drops by 8.46 bps, the 6-month yield drops by 4.4 bps, 2-year yield drops by 3.88 bps, 5-year yield drops by 2.26 bps, and 10-year yield drops by 1.11 bps. Second, using the the Gürkaynak, Sack and Wright (2007) data, I document that the LSAP1 shock causes a 50 bps instantaneous drop in the 10-year yield, consistent with the moment used in Ray (2019) Third, Ihrig et al. (2012) estimate that during LSAP1, the half-life of the 10-year yield response is approximately 2 years after the shock.

The last target is the response in average overnight repo rate to the LSAP1 shock. D'Amico and King (2013) report that during LSAP1, around $300 billion in Treasury securities were purchased, with 71% being off-the-run securities. d'Amico, Fan and Kitsul (2018) further detail that during 2009/3-2012/12, the ratio of Fed purchasing to selling was 46:1 for on-the-run and 1.82:1 for off-the-run securities. Moreover, Fed purchases of $0.29 billion in on-the-run securities decreased the repo rate by 0.224 bps, while purchases of $0.24 billion in off-the-run securities decreased the repo rate by 0.085 bps. Based on these data, I calculate that during 2009/3-2012/12, Fed net purchased $85 billion in on-the-run securities and $96 billion in off-the-run securities, resulting in a 65 bps drop in the repo rate for on-the-run securities and a 34 bps drop for off-the-run securities. The average response in the repo rate to the LSAP1 shock is thus approximately 48 bps.

The final parameter to be calibrated is the marginal safety preference, ϖ, which measures the sensitivity of consumption growth to private demand shocks through changes in households' preferences for safe saving products. I calibrate ϖ to match the ratio of output gap variance to short rate variance using data from 1985-2007. During this period, the variance of the 3-month Treasury yield was approximately 3.5, and the variance of the output gap was about 1.5, resulting in a ratio of 0.43. The calibration of ϖ concludes the parameterization of the model. In the next section, I will conduct several exercises to evaluate the model's performance in replicating key results from the literature.

7 Model Performance

This section evaluates the model's performance by replicating key results from the literature. First, I demonstrate that the model can replicate the positive relationship between risk premia and the slope of the term structure, a well-documented empirical finding originating with Fama and Bliss (1987). Next, I examine the model's implications for the repo market by studying repo specialness, as investigated by Jordan and Jordan (1997). Finally, I evaluate the macroeconomic implications by comparing the model's performance with the four-equation New Keynesian model developed by Sims, Wu and Zhang (2023).

7.1 Excess Return Predictability

Fama and Bliss (1987) examine the predictability of risk premia by running the following regression:

$$\frac{1}{\Delta\tau}\log\left(\frac{P_{t+\Delta\tau}(\tau-\Delta\tau)}{P_t(\tau)}\right)-y_t(\Delta\tau)=a_{\mathrm{FB}}(\tau)+b_{\mathrm{FB}}(\tau)\left(f_t(\tau-\Delta\tau,\tau)-y_t(\Delta\tau)\right)+e_{t+\Delta\tau}(\tau). \tag{38}$$

In equation (38), the dependent variable is the excess return from holding a bond of maturity τ over a period of $\Delta\tau$. This is defined as the average return during the holding period minus the spot rate for the same length of time. The independent variable is the slope of the term structure, defined as the forward rate between maturities $\tau-\Delta\tau$ and τ minus the spot rate for maturity $\Delta\tau$. Fama and Bliss (1987) finds that the coefficient $b_{FB}(\tau)$ is greater than 1 for most maturities, and its magnitude increases with τ.

To replicate the Fama-Bliss regression, I follow these steps: First, using the calibrated state variable transition matrix Γ and volatility matrix Σ, I simulate short rate and demand factor according to equation (33). Second, with the simulated short rate and demand factor, I simulate bond prices and yields using equations (15) and (25). Third, I compute forward rates for various maturities using the formula:

$$f_t(\tau - \Delta\tau, \tau) \quad = \frac{1}{\Delta\tau}\Big(A_r(\tau)r_t + A_\beta(\tau)\beta_t + C(\tau) - [A_r(\tau - \Delta\tau)r_t + A_\beta(\tau - \Delta\tau)\beta_t + C(\tau - \Delta\tau)]\Big).$$

Forth, I calculate the excess holding returns for various maturities using:

$$\begin{aligned}&\frac{1}{\Delta\tau}\log\left(\frac{P_{t+\Delta\tau}(\tau - \Delta\tau)}{P_t(\tau)}\right)\\ =&\frac{1}{\Delta\tau}\Big(A_r(\tau)r_t + A_\beta(\tau)\beta_t + C(\tau) - [A_r(\tau - \Delta\tau)r_{t+\Delta\tau} + A_\beta(\tau - \Delta\tau)\beta_{t+\Delta\tau} + C(\tau - \Delta\tau)]\Big).\end{aligned}$$

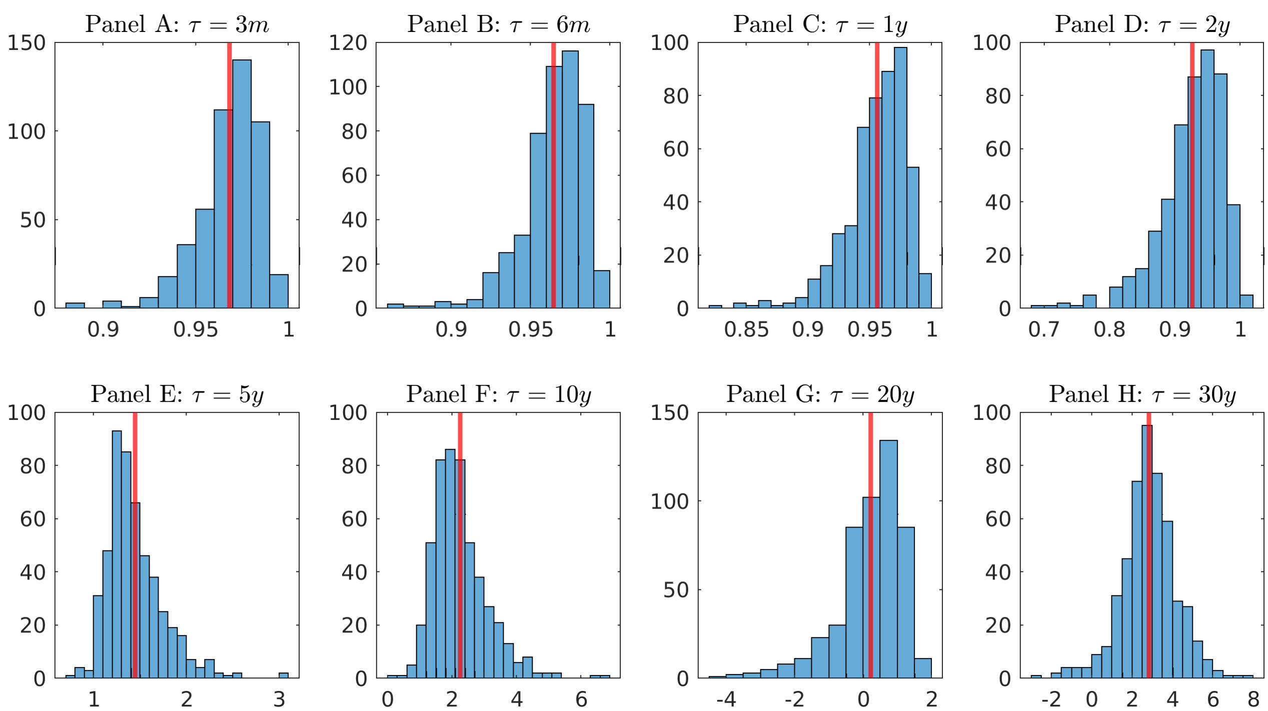


Figure 15: Estimated Fama-Bliss coefficient.

Note: Results are based on 500 simulations. The blue rectangles are histograms of the estimated 500 coefficients and the red line represents the mean.

With the simulated dependent and independent variables, I estimate the Fama-Bliss regression using equation (38). The sequences are simulated 500 times, and the estimated coefficients are reported based on these 500 simulations. The minimum maturity $\Delta\tau$ is set to 11 days and the time step is 1 day.

The estimation results are presented in Figure 15. Each panel represents a selected maturity, with the blue rectangles indicating the histograms of the estimated 500 coefficients and the red line showing the mean. The model performs well in generating a positive premia-slope relationship for most maturities. Generally, the estimated coefficient increases with maturity, except for the 20-year maturity. The magnitude of the coefficient is slightly below 1 for shorter maturities and greater than 1 for most mid- to long-term maturities.

7.2 Repo Specialness

Jordan and Jordan (1997) provide empirical evidence that repo specialness spills over into the cash market. They estimate the relationship using the following regression:

$$\log\left(\frac{P_t(\tau)}{P_t^{ref}(\tau)}\right) = a_{JJ}(\tau) + b_{JJ}(\tau)Special_t(\tau) + c_{JJ}(\tau)OTR_t(\tau) + \epsilon_t(\tau),$$

where the dependent variable is the specialness premium in the cash market, defined as the log difference between the observed bond price $P_t(\tau)$ and the potential reference price $P_t^{ref}(\tau)$, which would prevail without specialness. On the right hand side, $Special_t(\tau) = 1$ indicates that bond τ is special (i.e., its special repo rate is below the general collateral rate), and $OTR_t(\tau) = 1$ indicates that bond τ is on-the-run. Jordan and Jordan (1997) find that $b_{JJ} > 0$, meaning that repo market specialness results in a higher specialness premium in the cash market. In my model, I do not distinguish between on-the-run and off-the-run bonds, so constructing $OTR_t(\tau)$ is not feasible. What's more, Additionally, instead of using a dummy variable for specialness, I replace $Special_t(\tau)$ with the repo premium $r_t - R_t(\tau)$, forming the following updated regression:

$$\log\left(\frac{P_t(\tau)}{P_t^{ref}(\tau)}\right) = a_{JJ}(\tau) + b_{JJ}(\tau)(r_t - R_t(\tau)) + \epsilon_t(\tau). \tag{39}$$

I replicate the Jordan-Jordan regression in the following steps. First, I simulate short rate and demand factor using equation (33). Second, with the simulated short rate and demand factor, I simulate bond prices and repo rates using equations (15) and (26). Third, the reference prices are calculated under the assumption $\lambda = 0$, reflecting a frictionless repo market where the repo rate equals the short rate:

$$\log(P_r^{re}(\tau)) = -(A_r(\tau)^{re} r_t + A_\beta(\tau)^{re}\beta_t + C(\tau)^{ref}),$$

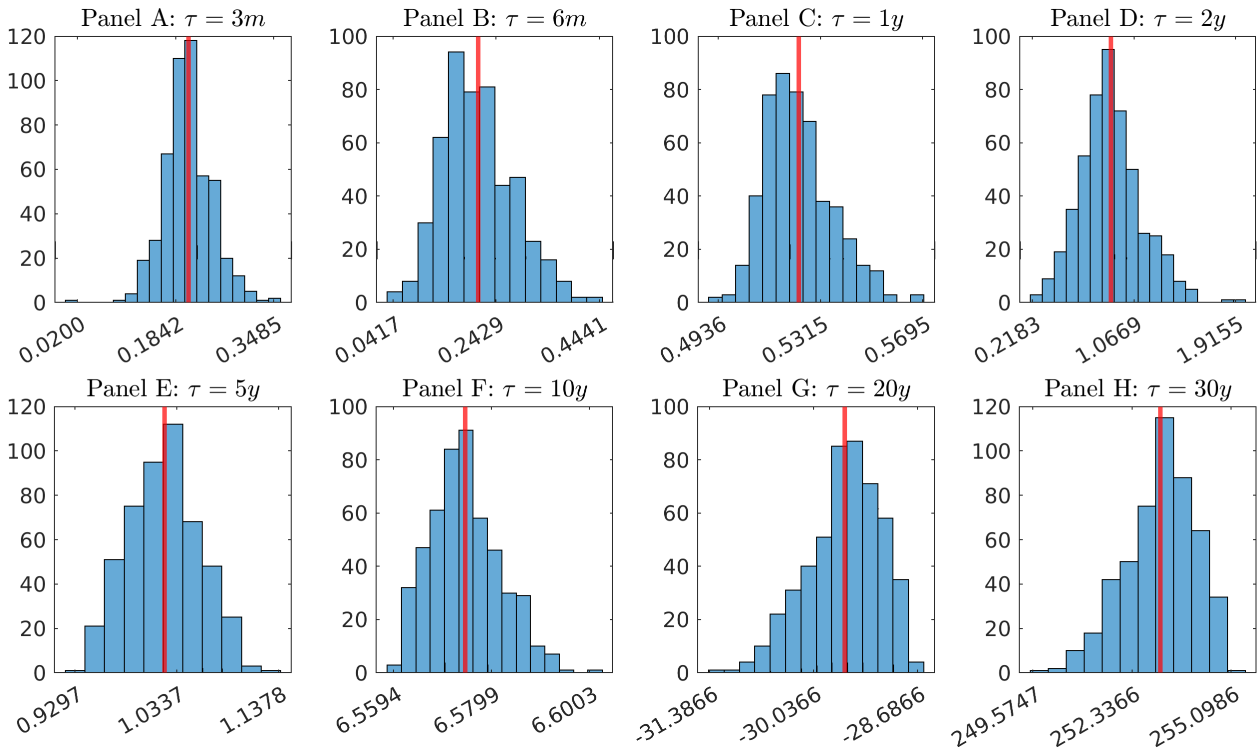


Figure 16: Estimated Jordan-Jordan coefficient.

Note: Results are based on 500 simulations. The blue rectangles are histograms of the estimated 500 coefficients and the red line represents the mean.

where $A_r(\tau)^{re}$, $A_\beta(\tau)^{re}$, and $C(\tau)^{ref}$ are reference affine coefficients with $\lambda = 0$. With all necessary variables simulated, I regress the cash market price premium on the repo premium across different maturities using equation (39). The minimum maturity and time step are consistent with those used in the Fama-Bliss replication. I run 500 simulations to estimate the coefficients.

The estimation results are presented in Figure 16. Each panel shows the histogram of the 500 estimated coefficients $b_{JJ}(\tau)$ across different maturities, with the blue rectangles representing the distribution and the red line indicating the mean value. The results indicate that the model generally produces a positive relationship between repo specialness and cash market premiums for most maturities, consistent with the empirical findings of Jordan and Jordan (1997). However, the exception is the 20-year maturity, where the coefficient is negative, suggesting a weaker or reversed relationship at this longer horizon.

7.3 Comparison with Four-Equation New Keynesian Model

Sims, Wu and Zhang (2023) introduce a four-equation New Keynesian model to analyze unconventional monetary policy, particularly focusing on the impact of QE. Their model incorporates a leverage constraint for financial intermediaries, which affects their long-term bond holdings and, consequently, the term structure of interest rates. A positive QE shock relaxes this leverage constraint, leading to a decrease in long-term rates relative to short-term rates.

To compare their model with mine, I first rewrite the Sims, Wu and Zhang (2023) model in continuous time in Appendix D1. The IS and Phillips curves from their model are represented as follows:

$$\mathrm{d}x_t = a\left(r_t - b * qe_t + c * x_t - d * \pi_t - r^*\right)\mathrm{d}t - b * d(qe_t), \tag{40}$$

$$d\pi_t = (e * qe_t - c * x_t - (1-d)\pi_t)dt. \tag{41}$$

In this specification, qe_t represents the QE shock, which affects the long-term interest rate by relaxing the leverage constraint on financial intermediaries. This relaxation has direct implications for the marginal benefit of saving, which is evident in the IS curve (40). The Phillips curve (41) includes an additional wage channel, where the wealth transfer due to QE affects the real economy by altering labor market dynamics and wages [28].

In contrast, my model features an IS curve with an additional term related to the term premia of bond yields and repo rates. The effect of a QE shock in my model can be evaluated by letting β_t denote the demand shock from the government. Since this demand shift does not stem from private investors, it does not affect household quality preferences. Hence, it is reasonable to assume that $\varpi = 0$ when using β_t as a QE shock. In my model, QE effects are captured through market segmentation and reduced excess returns due to targeted bond purchases.

Without the QE shock, Sims, Wu and Zhang (2023) model reduces to a standard three-equation New Keynesian framework. This is not the case in my model. In my model, the short rate risk is priced endogenously based on financial intermediary's portfolio optimization subject to various frictions. Thus, even without QE shocks, my model deviates from the

[28] Specifically, there are two types of households: Parents save in short-term bonds and transfer wealth to children; Children consume and issue long-term bonds. When a QE shock happens, the long term rate decreases relative to short term rate, implying a wealth transfer from parents to children. Since parents supply labor to firms, the wealth effect says that parents consume less, which puts a downward pressure on wage.

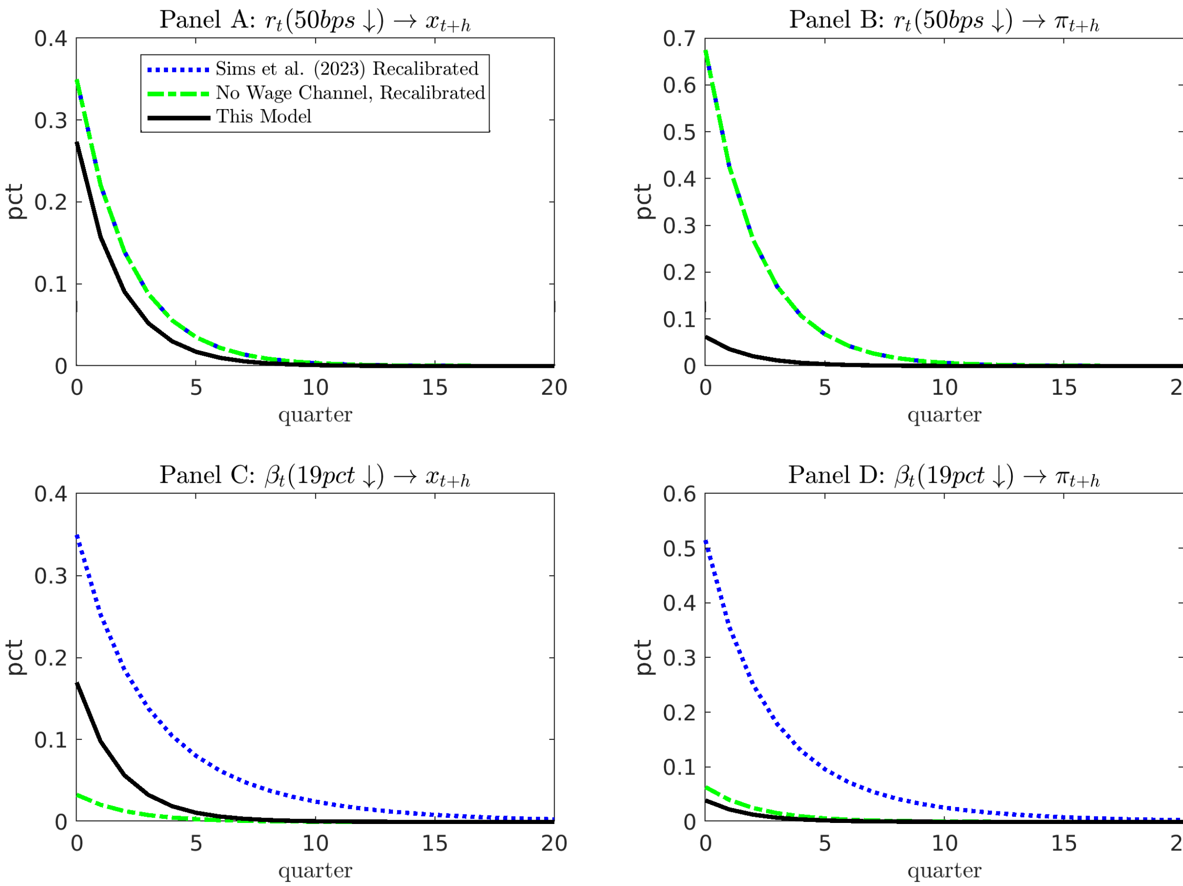


Figure 17: IRFs of recalibrated versions of Sims et al. (2023).

standard three-equation specification. The standard model is recovered in my setup when $\rho = 0$, $\lambda = 0$, and $\kappa_r = 0$, in which scenario both bond yields and repo rates are equivalent to short rate.

To analyze the comparative performance, I recalibrate the Sims, Wu and Zhang (2023) model with and without the wage channel. I then compare the IRFs of my model to these recalibrated versions. The detailed recalibration steps are provided in Appendices D2 and D3, respectively. Figure 17 shows the IRFs for both the Sims, Wu and Zhang (2023) model and my model. All models produce similar responses in the output gap to a short rate shock. A 0.5% decrease in the short rate leads to an approximate 0.3% increase in the output gap. My model generates smaller responses in the inflation rate compared to Sims, Wu and Zhang (2023). The effectiveness of QE in stimulating the macroeconomy is notably different across models. the wage channel significantly amplifies the effectiveness of QE, leading to more substantial increases in the output gap and inflation. The QE shock in Sims, Wu and Zhang (2023) without the wage channel is comparable to the QE effect in my model. A 19%

purchase of outstanding Treasuries results in an increase in the output gap of approximately 0.1% and inflation by about 0.07% in both models.

8 Baseline Results

8.1 Transmission of Short Rate Shock

In a fully segmented economy, where all bond supply is absorbed by habitat investors ($H_t(\tau) = 0$), Vayanos and Vila (2021) demonstrate that bond yields $y_t(\tau)$ become disconnected from the short rate. Applying zero habitat demand to equation (26) results in a zero Repo spread, and repo rates are equal to the short rate. This extreme case of segmentation equilibrium leads to zero short rate passthrough to bond yields and perfect passthrough to repo rates.

When arbitrageurs are present, however, the short rate is transmitted into bond yields through carry trades. To analyze the transmission mechanisms more thoroughly, I use the concept of instantaneous forward rates:

$$f_t(\tau) \equiv -\frac{\partial log P_t(\tau)}{\partial \tau} = A_r'(\tau) r_t + A_\beta'(\tau)\beta_t + C'(\tau). \tag{42}$$

When the short rate drops, arbitrageurs shift wealth from short rates to bonds, raising bond prices and lowering forward rates. Similar to Vayanos and Vila (2021) and Ray (2019), the carry trades in this model involve costs that hinder the passthrough of short rate shocks. These costs include risk cost and non-pecuniary cost. The risk cost is influenced by factors such as risk aversion ρ, short rate volatility σ_r, and habitat demand price elasticity $\alpha(\tau)$. This paper also introduces a non-pecuniary cost, potentially reflecting balance sheet cost or short selling cost. The magnitude of this non-pecuniary cost is an increasing function of the parameter λ.

The transmission of the short rate to overnight repo rates involves only the non-pecuniary cost, with no associated risk cost. In this model, the two risk factors influence the system by affecting bond prices. When hedge funds borrow cash in the repo market using bonds as collateral, they do not face binding borrowing constraints, and all repo contracts are assumed to be free of default risk [29]. Thus, price fluctuations in bond prices do not directly

[29] If hedge funds face binding borrowing constraint, then a lower short rate increase bond price, which relaxed the borrowing constraint and strengthens the passthrough to repo rate. Likewise, if repo contracts

affect arbitrageurs' first-order conditions for repo choices. Repos in this model are risk-free, with the only friction involved being the costly holding due to balance sheet constraints.

Figure 18 illustrates the frictions involved in short rate transmission to instantaneous forward rates $f_t(\tau)$, bond yields $i_t(\tau)$, and repo rates $R_t(\tau)$. The transmission is perfect when there is zero risk aversion ($\rho = 0$) or inelastic habitat demand of price for all maturities ($\alpha(\tau) = 0$), along with zero non-pecuniary cost ($\lambda = 0$). For the instantaneous forward rate and bond yields, the Expectations Hypothesis (EH) holds, as they are equal to the expected future short rate (Panel A and B, blue dashed lines). The reasons for the perfect passthrough differ in these two cases. When risk aversion is zero, arbitrageurs are indifferent to the risks associated with carry trades, making the transmission frictionless. In the second case, when habitat demand is completely price inelastic, arbitrageurs can impact bond prices without engaging in carry trades, corresponding to infinitely price-sensitive arbitrageurs.

The introduction of risk aversion and habitat demand price elasticity breaks the EH. As shown by the red dashed line in Panel A of Figure 18, the instantaneous forward rate underreacts to the short rate compared to the expected future short rate due to the risk costs incurred by risk-averse arbitrageurs conducting carry trades. Unlike forward rates or bond yields, the transmission to repo rates remains perfect in this scenario (Panel C, red dashed line). Dealers do not bear the risk of price fluctuations in the underlying collateral of repo contracts, so the short rate still maps one-to-one with repo rates.

Finally, the non-pecuniary cost further impedes transmission to forward rates (Panels A and B, green dotted lines). Adjusting the balance sheet now incurs additional costs for arbitrageurs, reducing the profitability of carry trades. Repo rates now respond less than one-to-one to the short rate, as indicated by the green dotted line in Panel C of Figure 18. Interestingly, in my model, because carry trades expose arbitrageurs to both risk and non-pecuniary costs, short rate transmission can be less than one-to-one even without risk aversion. When $\alpha(\tau) > 0$, arbitrageurs increase bond holdings to benefit from the yield spread, expanding the balance sheet. If financial intermediaries face additional non-pecuniary costs ($\lambda > 0$), the net expected excess return diminishes, discouraging carry trades and hindering transmission to forward rates. The non-pecuniary cost also spills over into the Treasury repo market, requiring repo rates to compensate dealers for the higher lending costs.

Figure 19 shows the IRFs for the aggregate nominal rate (Panel A), the output gap

have default possibility, a higher bond price helps increase the amount of asset recovered, which also improves the passthrough.

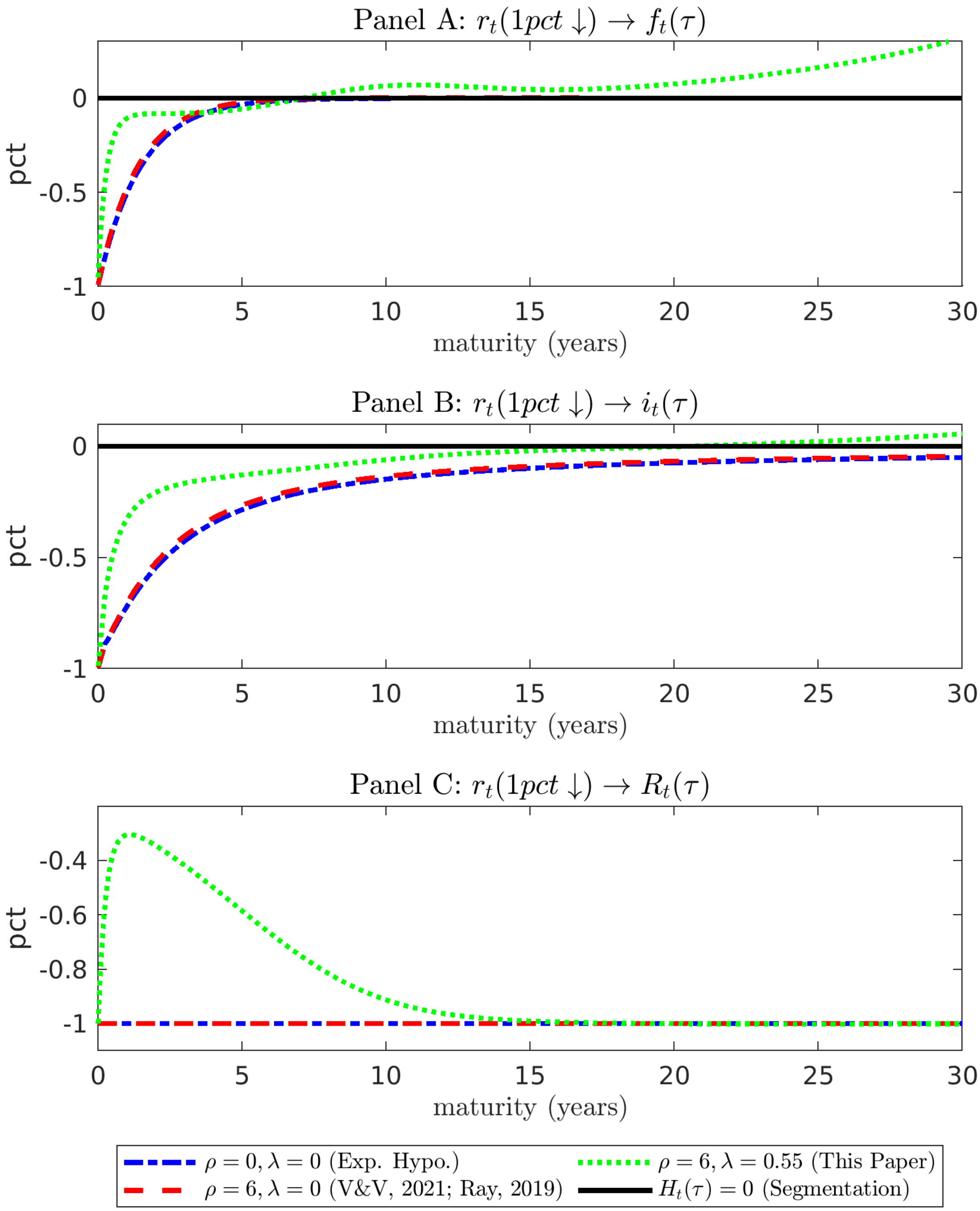


Figure 18: Transmission of short rate shock to asset prices.

Note: Parameters are set as calibrated. The short rate shock is a 1 percentage drop. Panel A shows the change in the instantaneous forward rate. Panel B shows the change in the bond yields. Panel C shows the change in the repo rates.

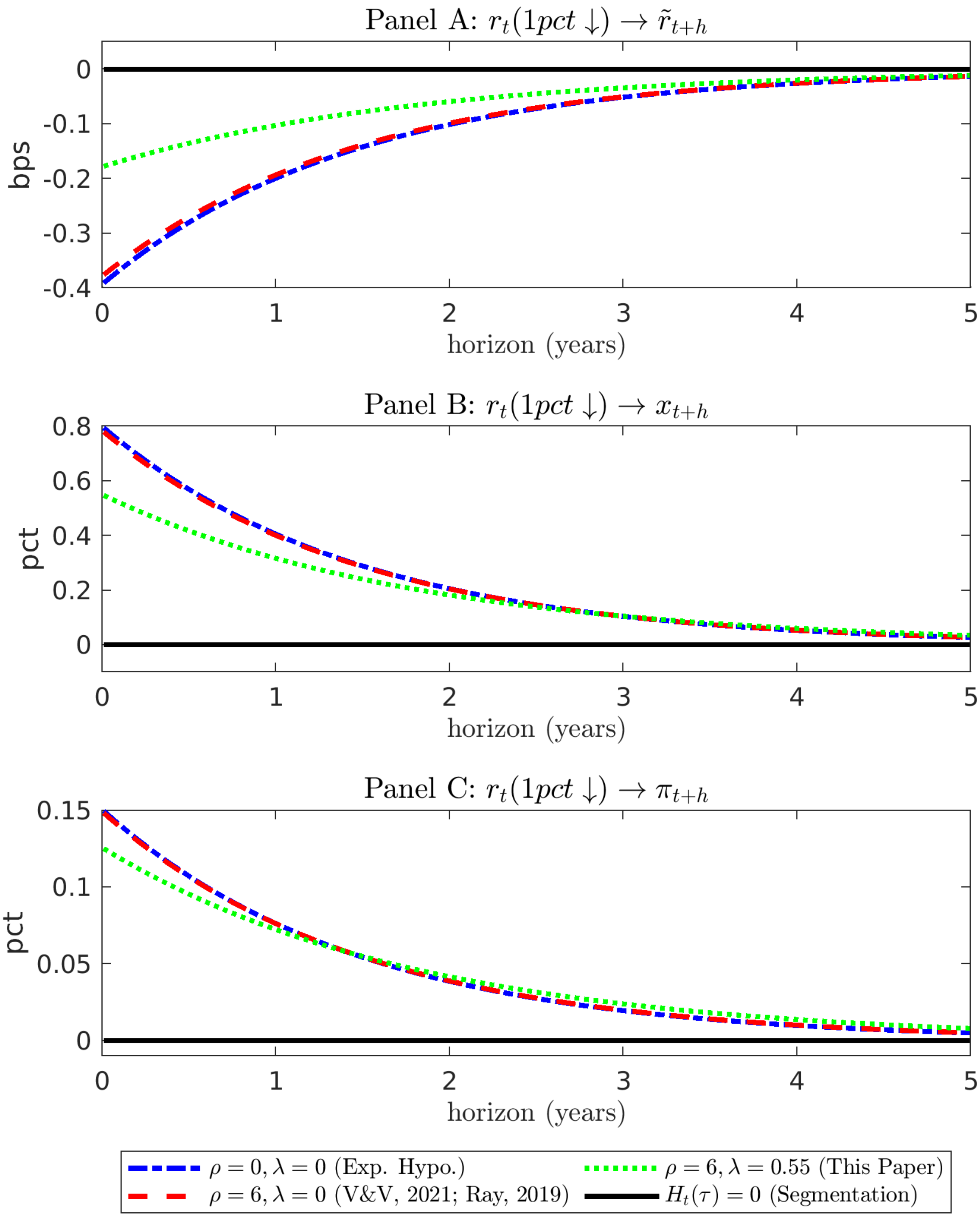


Figure 19: Transmission of short rate shock to macro variables.

Note: Parameters are set as calibrated. The short rate shock is a 1 percentage drop. Panel A shows the change in the aggregate nominal rate. Panel B shows the change in the output gap. Panel C shows the change in the inflation rate.

(Panel B), and the inflation rate (Panel C) in response to the same 1% drop in the short rate. Quantitatively, the calibrated level of risk aversion ($\rho = 6$) does not significantly alter the instantaneous responses of the aggregate nominal rate, output gap, or inflation rate to a short rate shock. A 1% drop in the short rate increases the output gap by 0.4% and the inflation rate by 0.15%. The calibrated marginal non-pecuniary cost $\lambda = 0.55$ significantly hinders the transmission of the short rate to macroeconomic variables. With the non-pecuniary cost in place, the same 1% drop in short rate now boosts the output gap by 0.2% and the inflation rate by 0.125%.

8.2 Transmission of Private Demand Shock

While financial frictions hinder the transmission of short rate shocks to asset prices and macroeconomic variables, they enhance the transmission of demand shocks. When the Expectations Hypothesis holds, forward rates are disconnected from the demand factor and depend solely on expected future short rates. Consequently, demand shocks have no effect in the absence of financial frictions, as shown by the blue dashed line in figure 20.

Financial frictions allow the demand factor to influence asset prices. In the case of a positive shock, the net demand of habitat investors increases. The increased habitat demand reduces arbitrageurs' bond holdings, lowering the marginal cost of investing in the portfolio. Arbitrageurs now require less compensation to hold bonds. When portfolio adjustments induce only risk costs, the demand shock affects asset prices purely by reducing the risk prices associated with the short rate and the demand factor. Since the risk price is consistent across all maturities, the demand shock has a broad effect, causing forward rates to drop across different maturities.

The introduction of non-pecuniary costs adds a local dimension to the demand shock's effects. The demand shock can target specific maturities, and the most targeted maturity benefits the most from reductions in non-pecuniary costs as the demand shock reduces arbitrageurs' holdings of those maturities. This is evident from the shape of the forward rate curve (Panel A, green dotted line). In the baseline calibration, the demand shock is concentrated around the 7.5-year maturity. The largest savings in non-pecuniary costs occur for mid-maturities, explaining the hump in the forward rate curve response. Additionally, the very long end of the forward rate curve also experiences a significant drop, but for different reasons than the mid-maturities. Specifically, the bond price sensitivity to the demand factor relative to the short rate increases with maturity. Although long-term bonds do not benefit from savings in non-pecuniary costs, they do benefit from reduced risk costs, explaining the

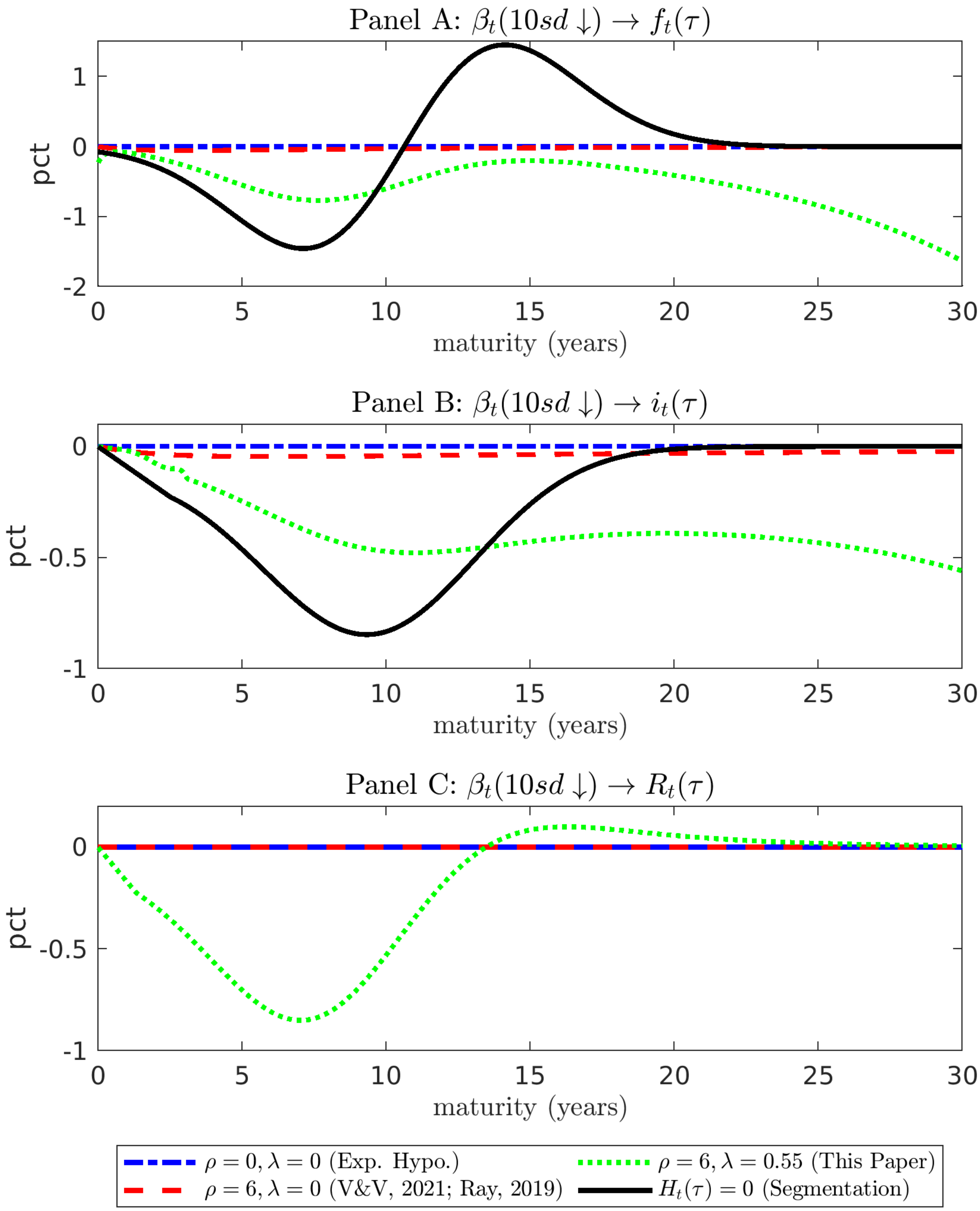


Figure 20: Transmission of demand shock to asset prices.

Note: Parameters are set as calibrated. The demand shock is a 10sd drop (125% central bank balances sheet expansion). Panel A shows the change in the instantaneous forward rate. Panel B shows the change in the bond yields. Panel C shows the change in the repo rates.

substantial drop at the long end of the forward rate curve.

The demand shock also affects repo rates when the non-pecuniary cost is accounted for. Since repo investors are not directly exposed to bond price risks, the repo spread reflects only the non-pecuniary cost. The repo curve's response is shaped by the maturity distribution of the calibrated baseline demand shock, with the largest response occurring at the mid maturities.

It is important to distinguish the different roles played by risk aversion (ρ) and the marginal non-pecuniary cost (λ). In equation (24), ρ is multiplied by the risk factor volatility and the aggregate porfolio's sensitivity to risk factors, neither of which depends on maturity. On the other hand, λ is multiplied by the habitat demand function, which varies by maturities. Intuitively, ρ governs the slope, while λ governs the curvature of the yield curve response. This distinction is particularly relevant in the case of a demand shock, as it shifts aggregate habitat demand $H_t(\tau)$ both directly through the location function $\theta(\tau)$ and indirectly through changes in bond price. In contrast, short rate shocks affect aggregate demand only indirectly via bond prices. Panel B of Figure 20 confirms this analysis, with the red line acting as a "drag" on the blue line and the green line adding a "tilt."

The best transmission occurs when financial markets are completely segmented, meaning there are no arbitrageurs trading across products. This is consistent with Vayanos and Vila (2021), who show that in a segmented economy, the price for any maturity is disconnected from the short rate and depends only on the demand shifter for that specific maturity. In the absence of arbitrageurs, a demand shock would cause bond prices to rise significantly, discouraging habitat investors from shifting their holdings. The extent to which bond prices rise depends on habitat investors' price elasticity for that specific maturity $\alpha(\tau)$.

The findings extend to macroeconomic variables, where the private demand shock generates quantitatively different responses. In the baseline calibration, the aggregate nominal rate depends only on bond yields. The same demand shock causes larger drops in yields when financial frictions are present, leading to more substantial growth in the output gap and inflation. Figure 21 presents the IRFs for the aggregate nominal rate (Panel A), the output gap (Panel B), and the inflation rate (Panel C) in response to a 10 standard deviation decrease in the demand factor. The positive demand shock simultaneously encourages consumption by reducing market rates and discourages consumption by increasing safety preferences. Financial frictions enhance the stimulus, with non-pecuniary cost alone accounting for a 0.05% increase in the output gap and a 0.013% increase in the inflation rate, based on the calibration. Risk aversion introduces an additional 0.05% increase in the output

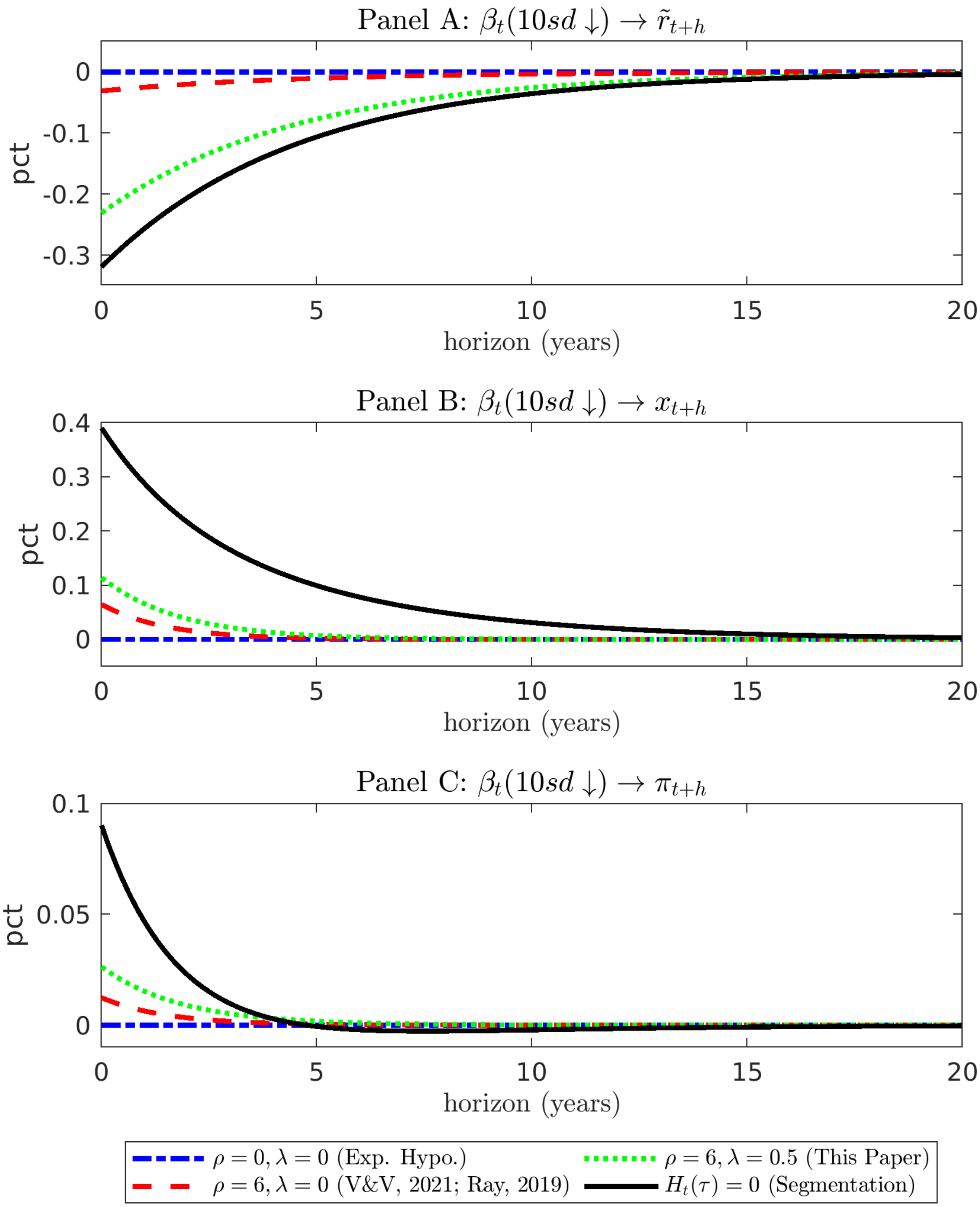


Figure 21: Transmission of demand shock to macro variables.

Note: Parameters are set as calibrated. The demand shock is a 10sd drop (125% central bank balances sheet expansion). Panel A shows the change in the aggregate nominal rate. Panel B shows the change in the output gap. Panel C shows the change in the inflation rate.

gap and a 0.012% increase in the inflation rate.

9 Policy Analysis

In the baseline model, the Taylor rule serves as an endogenous conventional monetary policy. From the previous section, I find that the efficacy of conventional monetary policy decreases as financial frictions increase. Moreover, the same change in the short-term interest rate can have different effects on bond yields and repo rates. In this section, I modify the baseline model to study two alternative monetary policies: forward guidance and QE. Additionally, I compare how these monetary policies behave before and after the benchmark rate reform.

9.1 Forward Guidance

Consider an unanticipated drop in the long-term mean of the short rate r^{ss} at time zero, which reverts deterministicaly to zeros at the rate $\kappa_{\bar{r}}$. To compute the effect, I denote a third state variable $d\bar{r}_t = -\kappa_{\bar{r}}\bar{r}_t$. The habitat demand now becomes

$$H_t(\tau) = -\alpha(\tau)logP_t(\tau) + \theta(\tau)\beta_t,$$

where $logP_t(\tau) = -[A_r(\tau)r_t + A_\beta(\tau)\beta_t + A_{\bar{r}}(\tau)\bar{r}_t]$. Collecting the three state variables into a vector:

$$d\begin{bmatrix} r_t \\ \beta_t \\ \bar{r}_t \end{bmatrix} = -\begin{bmatrix} \kappa_r & \kappa_{r\beta} & \kappa_{r\bar{r}} \\ 0 & \kappa_\beta & 0 \\ 0 & 0 & \kappa_{\bar{r}} \end{bmatrix} \left(\begin{bmatrix} r_t \\ \beta_t \\ \bar{r}_t \end{bmatrix} - \begin{bmatrix} r^{ss} - \bar{r}_t \\ \beta^{ss} \\ 0 \end{bmatrix} \right) dt + \begin{bmatrix} \sigma_r & 0 & 0 \\ 0 & \sigma_\beta & 0 \\ 0 & 0 & 0 \end{bmatrix} d\begin{bmatrix} B_{r,t} \\ B_{\beta,t} \\ B_{\bar{r},t} \end{bmatrix},$$

which simplifies to:

$$ds_t = -\Gamma(s_t - s^{ss}) + \Sigma dB_t.$$

To solve for affine coefficients, first denote $\vartheta(\tau) = \begin{bmatrix} 0 & \theta(\tau) & 0 \end{bmatrix}'$ as a 3x1 vector, $\epsilon = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}'$ as a 3x1 vector, and $E = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ as a 3x3 matrix . Then dealer's or hedge fund's FOC can be rewritten as:

$$A'(\tau)'s_t + C'(\tau) + A(\tau)'\Gamma\left(s_t - \underbrace{\bar{r}_t\epsilon}_{Es_t}\right) + \frac{1}{2}A(\tau)'\Sigma\Sigma'A(\tau) - \underbrace{r_t}_{\epsilon' s_t} - \lambda\Big[\underbrace{\theta(\tau)\beta_t}_{\vartheta(\tau)'s_t} - \alpha(\tau)\Big(A(\tau)'s_t + C(\tau)\Big)\Big]$$

$$= \frac{1}{\rho_d + \rho_h}A(\tau)'\Sigma\Sigma'\left[\int_0^T \Big[\underbrace{\theta(\tau)\beta_t}_{\vartheta(\tau)'s_t} - \alpha(\tau)\Big(A(\tau)'s_t + C(\tau)\Big)\Big]A(\tau)d\tau\right].$$

Collecting all terms relevant to s_t, I have:

$$A'(\tau) + \Gamma'A(\tau) - E' + \lambda\Big[\alpha(\tau)A(\tau) - \vartheta(\tau)\Big] = \frac{1}{\rho_d + \rho_h}\left[\int_0^T \Big[\vartheta(\tau) - \alpha(\tau)A(\tau)\Big]A(\tau)'d\tau\right]\Sigma\Sigma'A(\tau).$$

The system describing the conditions for $A_r(\tau)$ and $A_\beta(\tau)$ remains the same as in the baseline model, without the third state variable $\bar{r}(\tau)$. The affine coefficient $A_{\bar{r}}(\tau)$ satisfies:

$$\begin{aligned}
&A'_{\bar{r}}(\tau) + \kappa_{\bar{r}}A_{\bar{r}}(\tau) - \kappa_r A_r(\tau) + \lambda\alpha(\tau)A_{\bar{r}}(\tau) \\
=&\frac{1}{\rho_d + \rho_h}\left[A_r(\tau)\sigma_r^2\left(\int_0^T -\alpha(\tau)A_{\bar{r}}(\tau)A_r(\tau)d\tau\right)\right. \\
&\left.+A_\beta(\tau)\sigma_\beta^2\left(\int_0^T -\alpha(\tau)A_{\bar{r}}(\tau)A_\beta(\tau)d\tau\right)\right].
\end{aligned} \tag{43}$$

Since the bond price is $P_t(\tau) = exp(-A_r(\tau)r_t - A_\beta(\tau)\beta_t - A_{\bar{r}}(\tau)\bar{r}_t - C(\tau))$ in the three-state variable case, the marginal effect of this shock on bond yield is $A_{\bar{r}}(\tau)/\tau$, where $A_{\bar{r}}$ is solved from equation (43). The general equilibrium linear system becomes:

$$d\begin{bmatrix} r_t \\ \beta_t \\ \bar{r}_t \\ x_t \\ \pi_t \end{bmatrix} = -\underbrace{\begin{bmatrix} \psi_r & 0 & 0 & -\psi_r\phi_x & -\psi_r\phi_\pi \\ 0 & \kappa_\beta & 0 & 0 & 0 \\ 0 & 0 & \kappa_{\bar{r}} & 0 & 0 \\ -\varsigma^{-1}\hat{A}_r & -\varsigma^{-1}(\hat{A}_\beta + \varpi) & -\varsigma^{-1}\hat{A}_{\bar{r}} & 0 & \varsigma^{-1} \\ 0 & 0 & 0 & \delta & -\chi \end{bmatrix}}_{\Upsilon}\left(\begin{bmatrix} r_t \\ \beta_t \\ \bar{r}_t \\ x_t \\ \pi_t \end{bmatrix} - \begin{bmatrix} r^{ss} \\ \beta^{ss} \\ \bar{r}^{ss} \\ x^{ss} \\ \pi_{ss} \end{bmatrix}\right)dt + \underbrace{\begin{bmatrix} \sigma_r & 0 \\ 0 & \sigma_\beta \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}}_{\begin{bmatrix} S \\ 0 * I_2 \end{bmatrix}} d\begin{bmatrix} B_{r,t} \\ B_{\beta,t} \end{bmatrix}.$$

where $\hat{A}_{\bar{r}} = \int_0^T \frac{1}{\tau}\eta^i(\tau)A_{\bar{r}}(\tau)d\tau + \int_0^T \eta^R(\tau)\Big[-\lambda\alpha(\tau)A_{\bar{r}}(\tau)\Big)\Big]d\tau$. The general equilibrium solution is

$$d\begin{bmatrix} r_t \\ \beta_t \\ \bar{r}_t \end{bmatrix} = -\Gamma\left(\begin{bmatrix} r_t \\ \beta_t \\ \bar{r}_t \end{bmatrix} - \begin{bmatrix} r^{ss} \\ \beta^{ss} \\ \bar{r}^{ss} \end{bmatrix}\right)dt + SdB_t,$$

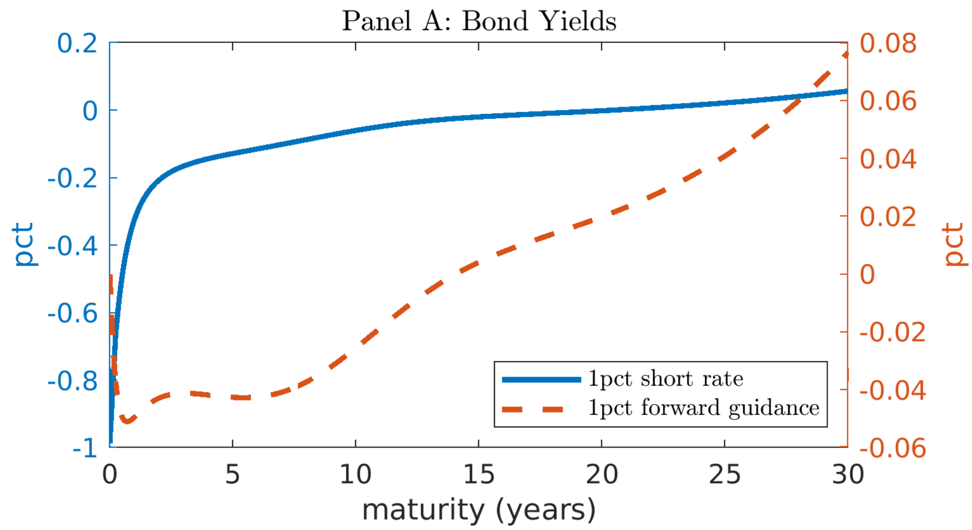
$$\begin{bmatrix} x_t \\ \pi_t \end{bmatrix} = \Omega\left(\begin{bmatrix} r_t \\ \beta_t \\ \bar{r}_t \end{bmatrix} - \begin{bmatrix} r^{ss} \\ \beta^{ss} \\ \bar{r}^{ss} \end{bmatrix}\right),$$

where the transition matrix Γ now is 3x3 and the output matrix Ω now is 2x3, both of which involve elements as functions of the eigenvalues and eigenvectors of Υ. The solution algorithm is the same as in the baseline case.

Figure 22 compares the impact of a conventional short rate expansion with that of forward guidance on bond yields and the repo rate. In the case of forward guidance, the future path of the short rate is altered without immediate changes to the current short rate. When there is an expansionary shock, $\bar{r}_t > 0$ and the long run short rate $r^{ss} - \bar{r}_t$ drops. Arbitrageurs' expected short rate in the future decreases, so they engage in carry trades of selling short rate and buying bonds, pushing bond prices to higher levels and bond yields to lower levels. This reduction in the expected future short rate leads arbitrageurs to engage in carry trades - selling short rates and purchasing bonds. These actions drive bond prices higher and yields lower. However, due to the risks and non-pecuniary costs associated with carry trades, the transmission is not perfect. In this sense, forward guidance achieves a similar outcome to a short rate shock.

However, unlike short rate, which has the most substantial impact on short maturities, forward guidance produces a hump-shaped response across the yield curve. After an expansionary forward guidance shock, the anticipated future short rate declines, while the current short rate remains unchanged. Therefore, at the very short end of the yield curve, the net benefit of holding bonds is unaffected by forward guidance. Quantitatively, the effect of forward guidance is significantly smaller than that of a standard short rate drop, with the largest response in the yield curve being below 0.1%. At the long end of the yield curve, both policies exhibit limited influence on long-term bond yields, but forward guidance is particularly weak, even causing a slight upward shift in long-term yields.

Surprisingly, forward guidance leads to an increase in repo rates. This occurs because forward guidance does not alter the current short rate, so its transmission to repo rates operates solely through the bond price channel. Higher bond prices reduce the habitat holdings,



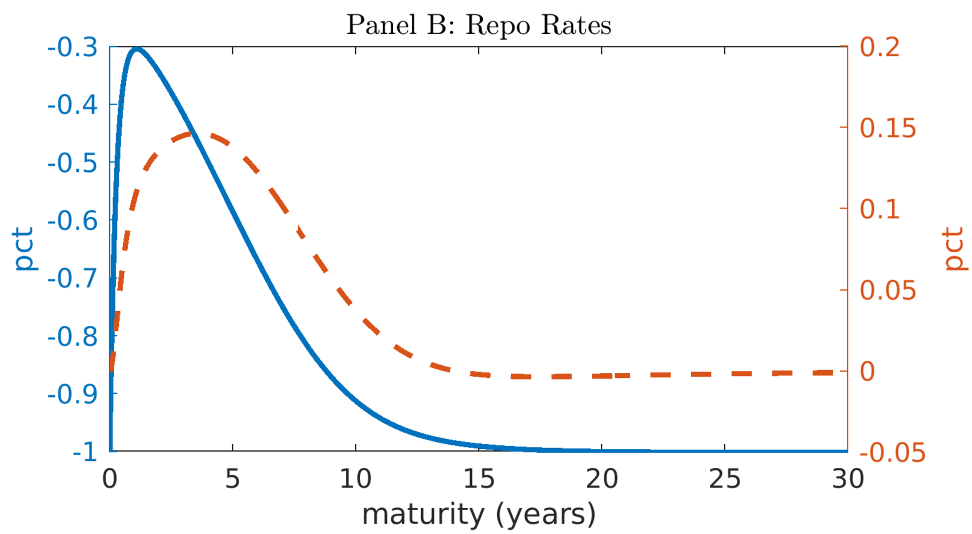


Figure 22: Transmission of short rate and forward guidance to asset prices.

Note: Parameters are set as calibrated. Both short rate shock and forward guidance shock refer to a 1 percentage decrease and are subject to the same speed of mean-reverting.

forcing arbitrageurs to absorb more supply, thereby raising their non-pecuniary costs. Consequently, an expansionary forward guidance shock can actually result in higher repo rates. Specifically, a 1% forward guidance shock raises the average repo rate by approximately 0.05%.

Figure 23 illustrates that the impact of a forward guidance shock is primarily governed by the mean-reversion speed parameter $\kappa_{\bar{r}}$. A faster mean-reversion implies a shorter-lived shock, leading to a smaller effect on the expected future short rate and, consequently, a

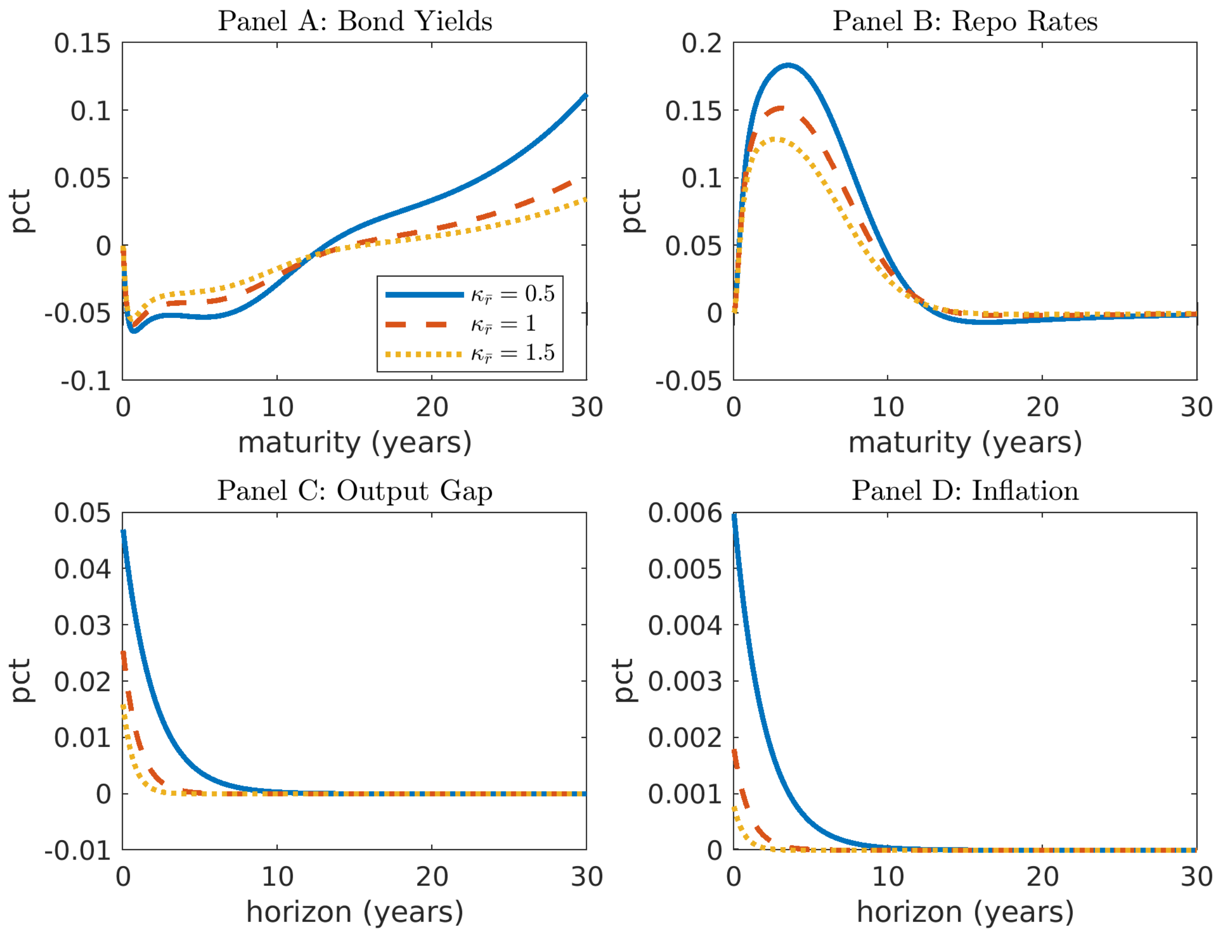


Figure 23: Impact of mean-reverting speed on forward guidance effectiveness.

Note: Parameters are set as calibrated.

reduced profitability from carry trades. Therefore, a larger $\kappa_{\bar{r}}$ diminishes the impact on interest rates and macroeconomic variables. Quantitatively, increasing the half-life of the forward guidance shock from 0.46 year ($\kappa_{\bar{r}} = 1.5$) to 1.49 year ($\kappa_{\bar{r}} = 0.5$) enhances the immediate effect on the output gap by 0.03% and the inflation rate by 0.005%. Interestingly, forward guidance moves bond yields and repo rates in opposite directions. In terms of macroeconomic impact, an expansionary forward guidance shock, similar to a conventional short rate drop, boosts the output gap and inflation in the baseline yield economy. However, in a repo-driven economy, this same forward guidance shock can induce a recession. Therefore, in an economy where borrowing is predominantly at the overnight repo rate, forward guidance may be counterproductive, as it fails to lower overall interest rates and instead exacerbates financial frictions[30].

9.2 QE on Various Maturities

Consider an unanticipated change $\Delta\theta_0(\tau)$ in the intercept of preferred-habitat demand at time zero that reverts deterministically to zero at the rate κ_θ. To compute the effect, I denote a third state variable $d\theta_t = -\kappa_\theta\theta_t$. The habitat demand now becomes:

$$H_t(\tau) = -\alpha(\tau)logP_t(\tau) + \Delta\theta_0(\tau)\theta_t + \theta(\tau)\beta_t,$$

where $logP_t(\tau) = -[A_r(\tau)r_t + A_\beta(\tau)\beta_t + A_\theta(\tau)\theta_t]$. Collecting the three state variables into a vector:

$$d\begin{bmatrix} r_t \\ \beta_t \\ \theta_t \end{bmatrix} = -\begin{bmatrix} \kappa_r & \kappa_{r\beta} & \kappa_{r\theta} \\ 0 & \kappa_\beta & 0 \\ 0 & 0 & \kappa_\theta \end{bmatrix}\left(\begin{bmatrix} r_t \\ \beta_t \\ \theta_t \end{bmatrix} - \begin{bmatrix} r^{ss} \\ \beta^{ss} \\ \theta^{ss} \end{bmatrix}\right)dt + \begin{bmatrix} \sigma_r & 0 & 0 \\ 0 & \sigma_\beta & 0 \\ 0 & 0 & 0 \end{bmatrix} d\begin{bmatrix} B_{r,t} \\ B_{\beta,t} \\ B_{\theta,t} \end{bmatrix},$$

which simplifies to:

$$s_t = -\Gamma(s_t - s^{ss}) + \Sigma dB_t.$$

To solve for affine coefficients, first denote $\vartheta(\tau) = \begin{bmatrix} 0 & \theta(\tau) & \Delta\theta_0(\tau) \end{bmatrix}'$ as a 3x1 vector, $\epsilon = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}'$ as a 3x1 vector selecting the first element. Then I can rewrite dealer's or hedge fund's FOC as:

$$A'(\tau)'s_t + C'(\tau) + A(\tau)'\Gamma\left(s_t - r^{ss}\epsilon\right) + \frac{1}{2}A(\tau)'\Sigma\Sigma'A(\tau) - \underbrace{r_t}_{\epsilon' s_t} \tag{44}$$

$$-\lambda\left[\underbrace{\Delta_0(\tau)\theta_t + \theta(\tau)\beta_t}_{\vartheta(\tau)'s_t} - \alpha(\tau)\Big(A(\tau)'s_t + C(\tau)\Big)\right]$$

$$= \frac{1}{\rho_d + \rho_h}A(\tau)'\Sigma\Sigma'\left[\int_0^T \left[\underbrace{\Delta_0(\tau)\theta_t + \theta(\tau)\beta_t}_{\vartheta(\tau)'s_t} - \alpha(\tau)\Big(A(\tau)'s_t + C(\tau)\Big)\right]A(\tau)d\tau\right]. \tag{45}$$

[30] This result is not plotted in figure 23 but can easily be speculated from figure 22.

Collecting all terms relevant to s_t gives:

$$A'(\tau) + \Gamma' A(\tau) - \epsilon + \lambda\Big[\alpha(\tau)A(\tau) - \vartheta(\tau)\Big] = \frac{1}{\rho_d + \rho_h}\left[\int_0^T \Big[\vartheta(\tau) - \alpha(\tau)A(\tau)\Big]A(\tau)'d\tau\right]\Sigma\Sigma' A(\tau). \tag{46}$$

Note that the system still describes the same conditions for $A_r(\tau)$ and $A_\beta(\tau)$ as in the case without the third state variable $\theta(\tau)$. The affine coefficient $A_\theta(\tau)$ satisfies:

$$\begin{aligned}
&A'_\theta(\tau) + \kappa_\theta A_\theta(\tau) + \lambda\Big(\alpha(\tau)A_\theta(\tau) - \Delta\theta_0(\tau)\Big) \\
=&\frac{1}{\rho_d + \rho_h}\left[A_r(\tau)\sigma_r^2\left(\int_0^T \Big(\Delta\theta_0(\tau) - \alpha(\tau)A_\theta(\tau)\Big)A_r(\tau)d\tau\right)\right. \\
&\left.+A_\beta(\tau)\sigma_\beta^2\left(\int_0^T \Big(\Delta\theta_0(\tau) - \alpha(\tau)A_\theta(\tau)\Big)A_\beta(\tau)d\tau\right)\right].
\end{aligned} \tag{47}$$

Given that the bond price is $P_t(\tau) = exp(-A_r(\tau)r_t - A_\beta(\tau)\beta_t - A_\theta(\tau)\theta_t - C(\tau)$ in the three-state variable case, the marginal effect of this shock on bond yield is $A_\theta(\tau)/\tau$, where $A_\theta(\tau)$ is solved from equation (47). The QE shock θ_t is a demand shock similar to the private habitat demand shock β_t, but it refers to the central bank's balance sheet policy. Consequently, QE shock θ_t does not affect households' safety preferences. The general equilibrium linear system in this case becomes:

$$d\begin{bmatrix} r_t \\ \beta_t \\ \theta_t \\ x_t \\ \pi_t \end{bmatrix} = -\underbrace{\begin{bmatrix} \psi_r & 0 & 0 & -\psi_r\phi_x & -\psi_r\phi_\pi \\ 0 & \kappa_\beta & 0 & 0 & 0 \\ 0 & 0 & \kappa_\theta & 0 & 0 \\ -\varsigma^{-1}\hat{A}_r & -\varsigma^{-1}(\hat{A}_\beta + \varpi) & -\varsigma^{-1}\hat{A}_\theta & 0 & \varsigma^{-1} \\ 0 & 0 & 0 & \delta & -\chi \end{bmatrix}}_{\Upsilon}\left(\begin{bmatrix} r_t \\ \beta_t \\ \theta_t \\ x_t \\ \pi_t \end{bmatrix} - \begin{bmatrix} r^{ss} \\ \beta^{ss} \\ \theta^{ss} \\ x^{ss} \\ \pi_{ss} \end{bmatrix}\right)dt + \underbrace{\begin{bmatrix} \sigma_r & 0 \\ 0 & \sigma_\beta \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}}_{\begin{bmatrix} S \\ 0 * I_2 \end{bmatrix}} d\begin{bmatrix} B_{r,t} \\ B_{\beta,t} \end{bmatrix}$$

where $\hat{A}_\theta = \int_0^T \frac{1}{\tau}\eta^i(\tau)A_\theta(\tau)d\tau + \int_0^T \eta^R(\tau)\Big[\lambda\big(\Delta\theta_0(\tau) - \alpha(\tau)A_\theta(\tau)\big)\Big]d\tau$. The general equilibrium solution is:

$$d\begin{bmatrix} r_t \\ \beta_t \\ \theta_t \end{bmatrix} = -\Gamma\left(\begin{bmatrix} r_t \\ \beta_t \\ \theta_t \end{bmatrix} - \begin{bmatrix} r^{ss} \\ \beta^{ss} \\ \theta^{ss} \end{bmatrix}\right)dt + SdB_t,$$

$$\begin{bmatrix} x_t \\ \pi_t \end{bmatrix} = \Omega\left(\begin{bmatrix} r_t \\ \beta_t \\ \theta_t \end{bmatrix} - \begin{bmatrix} r^{ss} \\ \beta^{ss} \\ \theta^{ss} \end{bmatrix}\right),$$

where the transition matrix Γ now is 3x3 and the output matrix Ω now is 2x3, both involving elements as functions of the eigenvalues and eigenvectors of Υ. The solution algorithm is the same as in the baseline case.

QE operates by reducing the marginal cost of arbitrageurs' portfolio investments. Since different bond maturities (tenors) have varying sensitivities to demand risk, QE shocks of the same magnitude can generate different effects depending on which maturities are targeted. Moreover, the inclusion of non-pecuniary costs adds another layer of specificity to QE effects, as maturities that are more directly targeted by QE enjoy greater reductions in these costs.

Figure 24 shows the effects of QE targeting different maturities. The short-end QE concentrates on 3-year maturities, mid-end QE on 15-year maturities, and long-end QE on 27-year maturities. In the baseline model, the importance of demand risk relative to short rate risk increases with maturity. As a result, targeting longer maturities achieves more significant effects on bond yields. Compared to bond yields, the responses in repo rates are more localized. Repo rates do not include risk premia, so their responses to QE are shaped solely by changes in non-pecuniary costs. More targeted maturities see larger drops in repo rates, whereas less targeted maturities may experience smaller declines or even reverse responses due to indirect effects. This reversal in repo rate responses at shorter maturities when QE targets the mid or long end is particularly notable. Although these shorter maturities benefit weakly from the shock, their high elasticity in habitat demand means that the indirect effects (i.e., increased bond prices leading to reduced habitat holdings) can become strong enough to reverse the sign of repo rate responses. In terms of macroeconomic outcomes, because long-end QE produces larger drops in bond yields, it has a more substantial impact on the output gap and inflation. Quantitatively, long-end QE improves the output gap by 1% and the inflation rate by 0.25% compared to short-end QE.

Next, I examine QE effects when the central bank operates under a fixed budget constraint. When conducting asset purchases, the central bank has a specific amount of re-

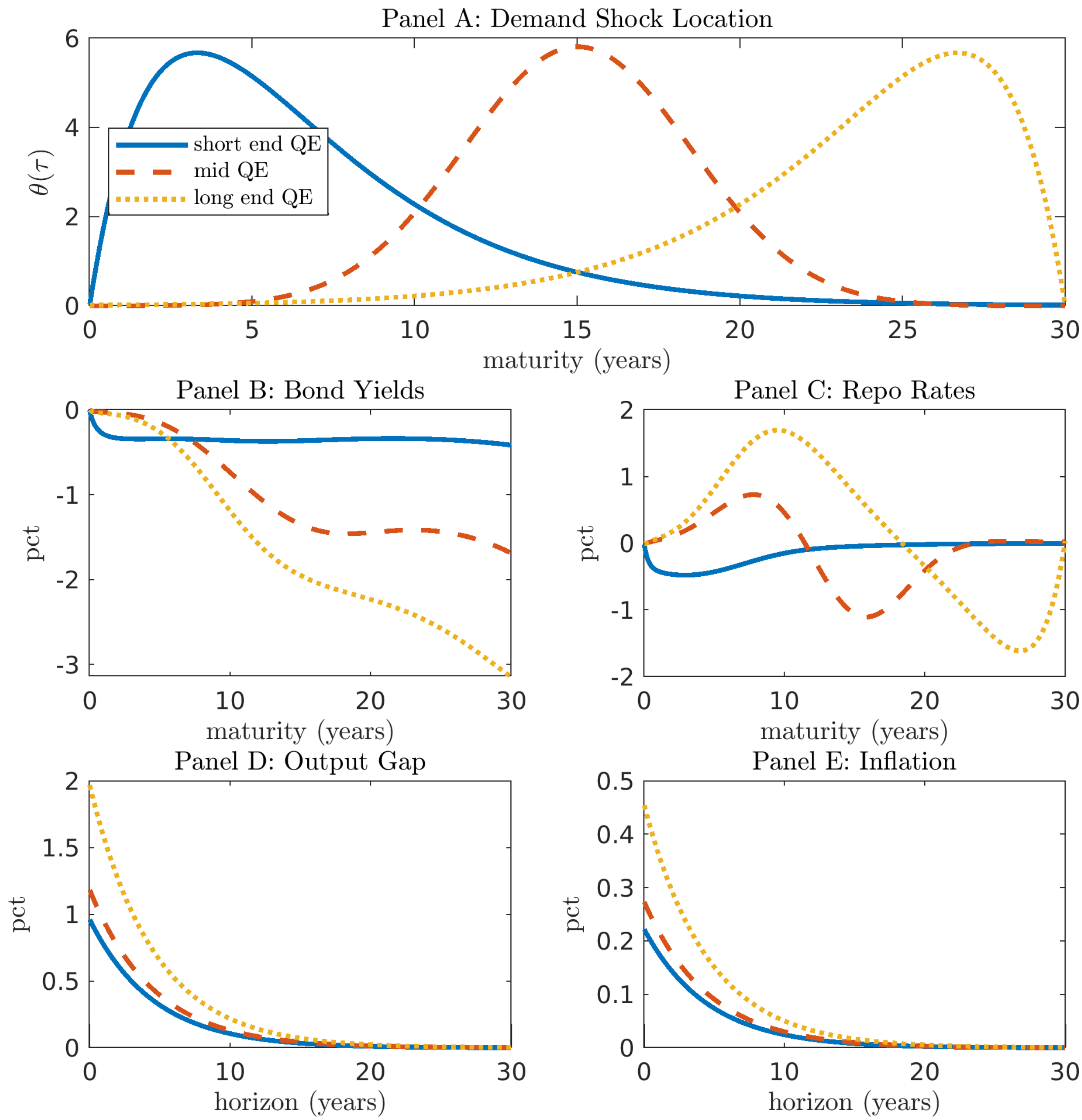


Figure 24: Impact of targeting maturity on QE effectiveness.

Note: Parameters are set as calibrated. Size of the QE shock is 10 standard deviations of the demand factor (27%) in all cases.

sources to allocate, and the price of Treasuries also influences the aggregate effect of QE.

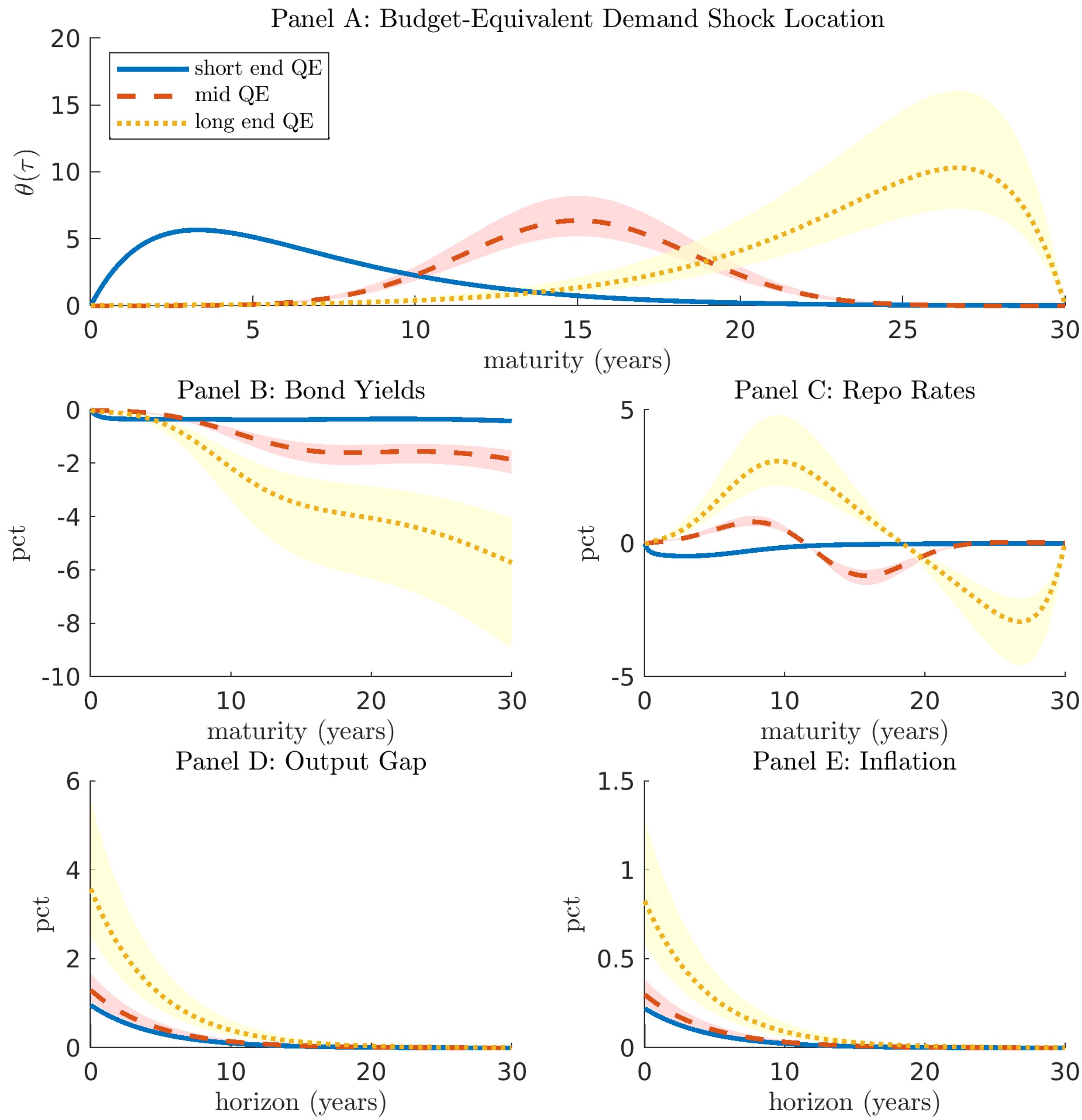


Figure 25: Impact of targeting maturity on budget-equivalent QE effectiveness.

Note: Parameters are set as calibrated. Size of the QE shock is chosen to have the same cost as for a 10 standard deviation (27%) purchase of short end. Bands are 90% confidence intervals.

Typically, long-term bonds are cheaper due to their higher risk, meaning that a fixed budget can purchase a larger quantity of long-term bonds compared to short-term ones.

Figure 25 replicates the analysis under the condition that all purchasing schemes are budget-equivalent, specifically at a cost equivalent to a 27% purchase of the short-end scheme. However, because the two state variables, the short rate r_t and the private demand factor β_t, are stochastic, the budget-equivalent purchasing scales are uncertain. Simulating the state variables 100 times, I calculate the budget-equivalent purchasing scales and aggregate effects for each scenario. Then the bands in figure 25 represent the 90% confidence interval based on these simulations. When constrained by the same budget, long-end QE outperforms even more due to the lower price of long-term Treasuries. Specifically, budget-equivalent long-end QE results in a 3% greater response in the output gap and a 0.55% larger increase in the inflation rate compared to short-end QE.

9.3 Benchmark Rate Reform

In this section, I explore the performance of monetary policies in two distinct scenarios: the "yield economy" and the "repo economy." The yield economy is characterized by the aggregate nominal interest rate depending solely on bond yields, i.e., $\eta^R(\tau) = 0$, while the repo economy is defined by the aggregate nominal interest rate depending solely on repo rates, i.e., $\eta^i(\tau) = 0$.

These scenarios are intended to reflect different indexing schemes before and after the U.S. benchmark rate reform from LIBOR to SOFR. Although the Treasury yield and LIBOR are derived from different mechanisms, they closely track each other historically, as shown in Figure 26. Therefore, using the Treasury yield as a proxy for LIBOR is reasonable, particularly when focusing on dynamics rather than absolute levels.

9.3.1 Monetary Policy Transmission

To understand how monetary policies transmit under these two regimes, I first examine the IRFs of macroeconomic variables following a conventional expansionary shock, as depicted in figure 27. The blue solid line represents the yield economy, and the green dashed line represents the repo economy. The short rate shock affects bond yields through both the risk cost and non-pecuniary cost, while its transmission to repo rates is hindered only by the non-pecuniary cost. Thus, repo rates are expected to respond more sharply to changes in the short rate, as reflected in the different magnitudes of the average response in the yield curve (panel B) and repo rates (panel C).

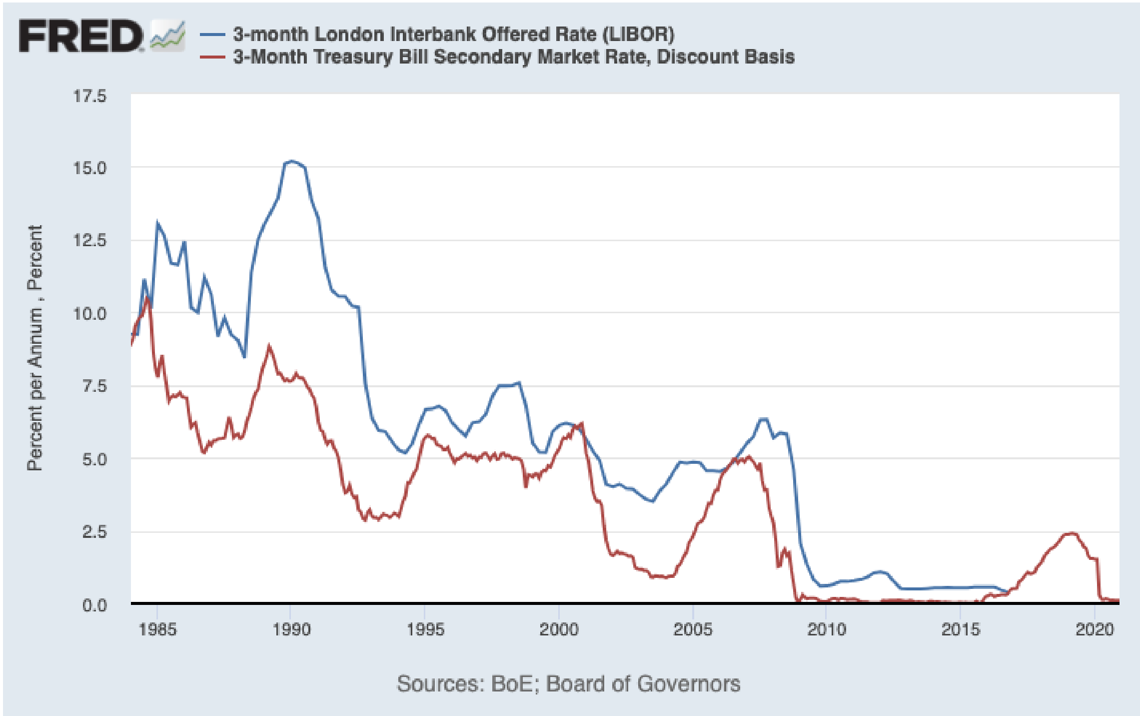


Figure 26: 3-month LIBOR and Treasury yield.

In the repo regime, the conventional expansionary shock has a more significant impact on the aggregate nominal rate due to the direct indexing of savings products by repo rates. This lower interest rate discourages saving, thereby boosting consumption and output, which, in turn, drives up inflation. Comparing the effects of a 1% conventional expansionary shock across the two regimes, the repo regime exhibits 0.4% higher consumption growth and 0.08% higher inflation than the yield regime.

Next, I explore the differential effects of a forward guidance shock across the two regimes, as shown in Figure 28. The results are surprising: while an expansionary forward guidance shock (a decrease in long run short rate r^{ss}) stimulates the yield economy, it causes a recession in the repo economy. This counterintuitive result arises from the indirect bond price channel. The forward guidance shock lowers bond yields across most maturities but raises repo rates. The increase in repo rates is due to the higher non-pecuniary cost as arbitrageurs take on more holdings due to the discouragement of habitat investors from higher bond prices. As calibrated, a 1% drop in the long-run short rate increases the output gap by 0.035% and the inflation rate by 0.0045% in the yield economy. In contrast, in the repo economy, the same shock reduces the output gap by 0.075% and the inflation rate by 0.005%.

Finally, I compare the effects of a QE shock across the two regimes, illustrated in Figure

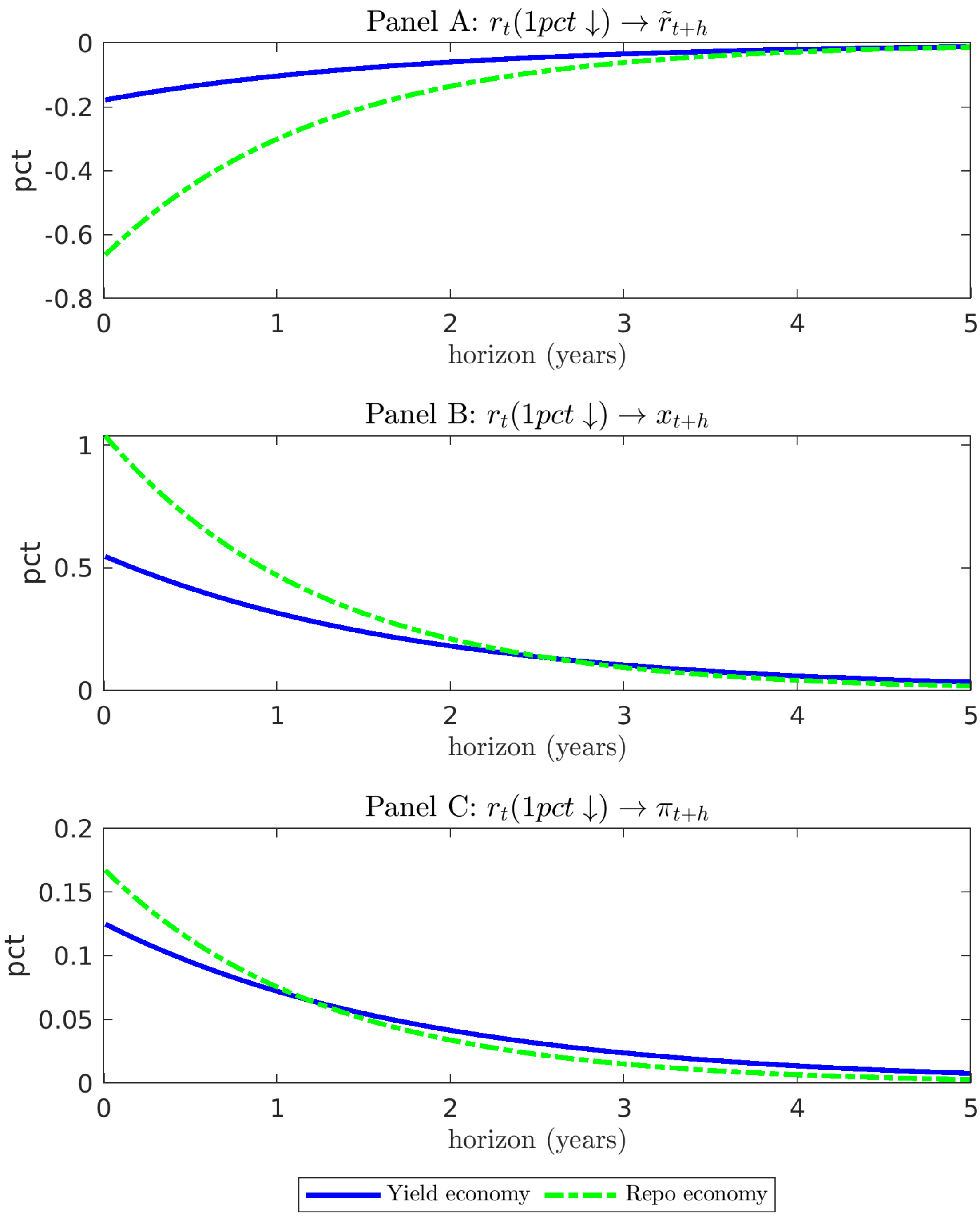


Figure 27: Transmission of short rate shock to macro variables.

Note: Parameters are set as calibrated. The short rate shock is a 1 percentage drop. Panel A shows the change in the aggregate nominal rate. Panel B shows the change in the output gap. Panel C shows the change in the inflation rate.

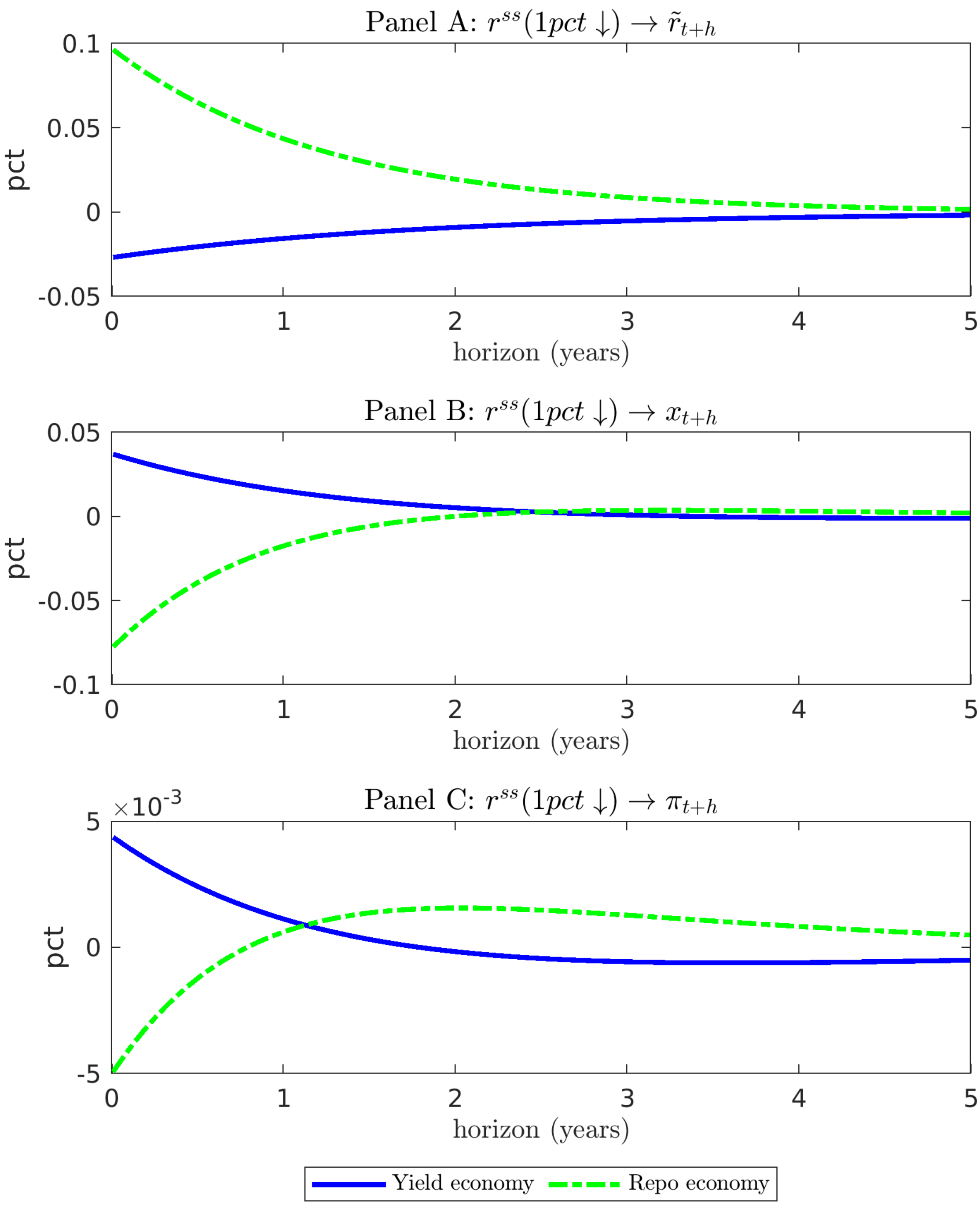


Figure 28: Transmission of forward guidance shock to macro variables.

Note: Parameters are set as calibrated. The forward guidance shock is a 1 percentage drop in the long run short rate. Panel A shows the change in the aggregate nominal rate. Panel B shows the change in the output gap. Panel C shows the change in the inflation rate. The mean-reverting speed of the forward guidance shock is the same as for the short rate shock.

29. When repo rates index households' savings products, the QE's effectiveness in stimulating the macroeconomy is diminished. From Figure 24, it is evident that QE has a more "global" impact on bond yields, producing larger yield drops across longer maturities, regardless of the QE's originating tenor. However, the same QE shock has more "localized" effects on repo rates, with limited or even reversed impacts on non-targeted tenors. As a result, QE is more effective in influencing overall borrowing costs in the yield economy, especially when the shock originates from mid- to long-end maturities. The precise magnitude of QE's effects depends on the targeted maturities. In the baseline calibration, where the QE shock is set to match the maturity targets during LSAP1, the impact of a QE purchase sized at 27% of outstanding Treasuries is 0.14% smaller on the output gap and 0.09% smaller on inflation in the repo economy compared to the yield economy.

9.3.2 Optimal Long-Run Policy Rate Target

In the standard New Keynesian model, the aggregate nominal rate $\tilde{r}_t$ is equivalent to the policy rate r_t. The optimal long-run policy rate should be set as equal to the natural interest rate $r^* = \bar{r}$. To see this is the case, first, consider the steady-state version of the IS curve:

$$\mathrm{d}x^{ss} = \varsigma^{-1} \left(\tilde{r}^{ss} - \pi^{ss} - \bar{r}\right) \mathrm{d}t,$$

the steady-state aggregate nominal rate $\tilde{r}^{ss}$ must be equal to the natural rate $\tilde{r}$, i.e., $\tilde{r}^{ss} = \tilde{r}$. Next, consider the steady-state version of the Taylor rule:

$$\mathrm{d}r^{ss} = -\psi_r \left(r^{ss} - \phi_\pi \pi^{ss} - \phi_x x^{ss} - r^*\right) \mathrm{d}t,$$

the steady-state policy rate r^{ss} must be equal to the policy target r^*, i.e., $r^{ss} = r^*$. In a natural economy, without financial frictions, bond yields and repo rates for all maturities are identical to the short rate. Thus, the steady-state aggregate nominal rate $\tilde{r}^{ss}$ is equal to the steady-state short rate r^{ss}, i.e., $\tilde{r}^{ss} = r^{ss}$. Consequently, the following equivalence holds: $\tilde{r}^{ss} = r^{ss} = r^* = \bar{r}$.

However, financial frictions deviate the aggregate nominal rate from the short rate. Denoting the yield and repo economies by subscripts y and r, respectively, the IS curve implies that the steady-state aggregate nominal rate is still equal to the natural rate. However, the steady-state short rate now differs from the steady-state aggregate nominal rate. To find the

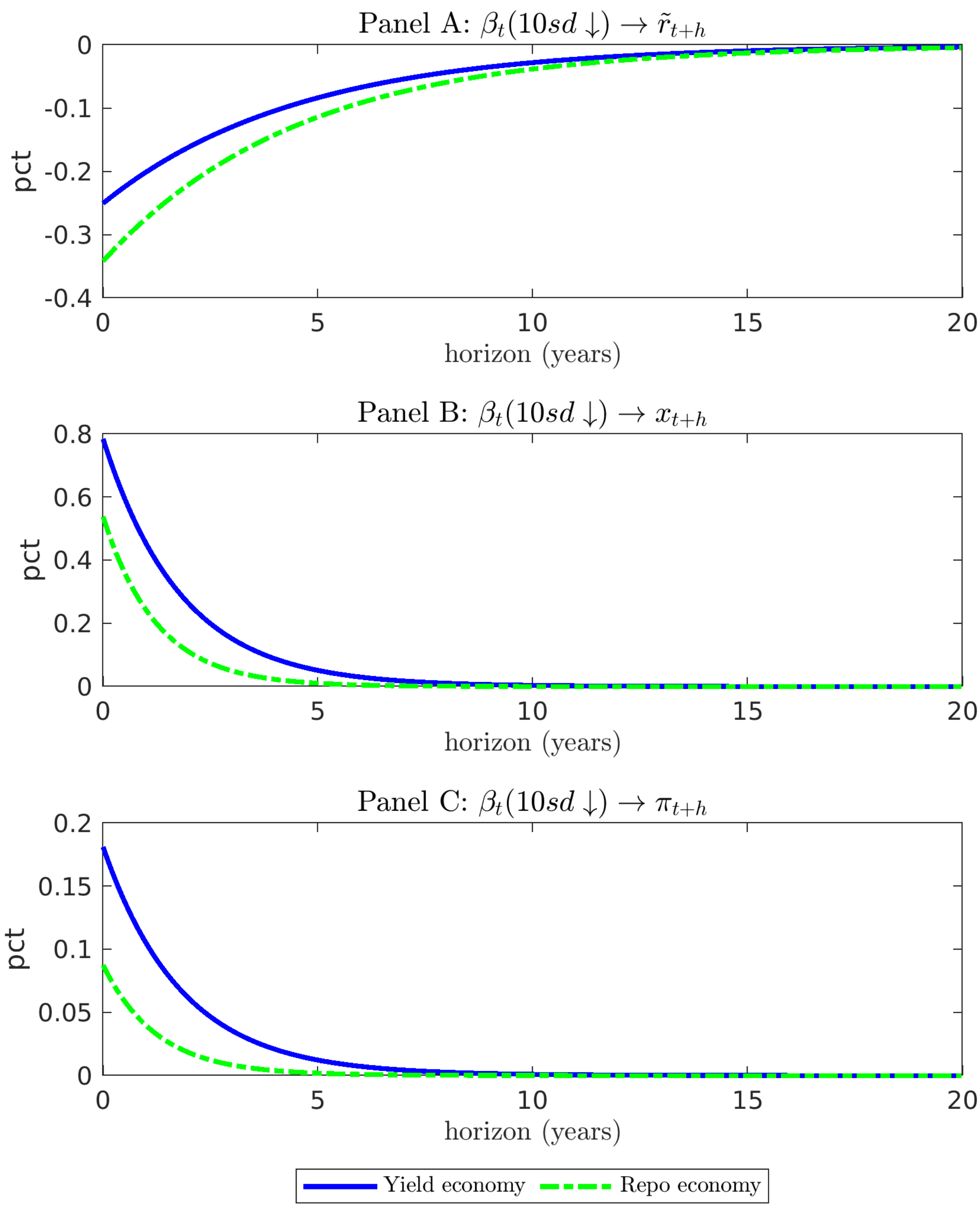


Figure 29: Transmission of QE shock to macro variables.

Note: Parameters are set as calibrated. The QE shock is 10 standard deviations of the demand factor (27%). Panel A shows the change in the aggregate nominal rate. Panel B shows the change in the output gap. Panel C shows the change in the inflation rate.

optimal long-term policy target r^*, I first derive:

$$r^{ss,y} = r^{*,y},$$
$$r^{ss,r} = r^{*,r}.$$

I can further express the aggregate rate $\tilde{r}^{ss}$ in the yield and repo economy, respectively, as:

$$\tilde{r}^{ss,y} = \hat{A}_r^y r^{ss,y} + \hat{C}^y,$$
$$\tilde{r}^{ss,r} = \hat{A}_r^r r^{ss,r} + \hat{C}^r,$$

where

$$\hat{A}_r^y \equiv \int_0^T \frac{1}{\tau}\eta^i(\tau)A_r^y(\tau)d\tau,$$
$$\hat{A}_r^r \equiv \int_0^T \eta^R(\tau)\Big(1 - \lambda\alpha(\tau)A_r^r(\tau)\Big)d\tau,$$
$$\hat{C}^y \equiv \int_0^T \frac{1}{\tau}\eta^i(\tau)C^y(\tau)d\tau,$$
$$\hat{C}^r \equiv -\lambda\int_0^T \eta^R(\tau)\alpha(\tau)C^r(\tau)d\tau.$$

Since the demand factor is assumed to be zero in the steady states, the optimal long-run policy target that ensures a zero output gap is independent of the affine coefficient of the demand factor A_β. Given that the steady-state aggregate nominal rates in both regimes, $\tilde{r}^{ss,y}$ and $\tilde{r}^{ss,r}$, are equal to the natural rate $\bar{r}$, the optimal long-run policy target in the two regimes can be derived as:

$$r^{*,y} = (\bar{r} - \hat{C}^y)/\hat{A}_r^r,$$
$$r^{*,r} = (\bar{r} - \hat{C}^r)/\hat{A}_r^r.$$

I compute the numerical values of the optimal long-run policy target in the yield economy $r^{*,y}$ and in the repo economy $r^{*,r}$ in figure 30. The natural interest rate, which corresponds to the prevailing rate $\bar{r}$ when the potential GDP equals the true GDP, is set to 1%. Panel A plots the optimal rates as functions of the risk aversion ρ. In the yield economy, because risk aversion hinders the transmission of the short rate, the central bank needs to set a more aggressive policy target to achieve a zero output gap in the long run. This explains

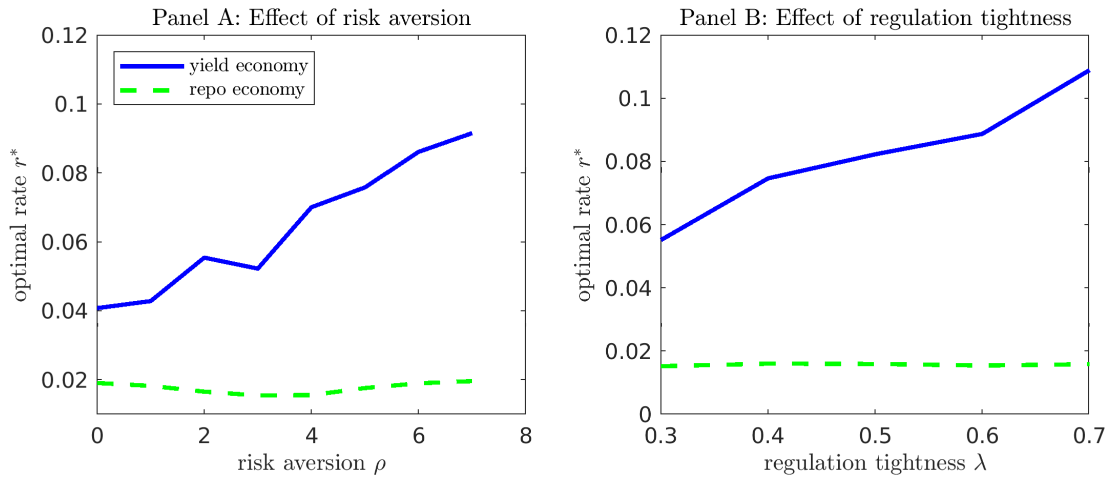


Figure 30: Optimal long-run policy rate target in yield and repo economy.

Note: Parameters are set as calibrated. The natural interest rate $\bar{r}$ is 1%.

the upward slope of the blue solid line. In contrast, since excess returns on repo rates do not include a risk component, the efficiency of conventional policy in the repo economy is less sensitive to risk aversion. Consequently, the green dashed line remains nearly flat with respect to ρ. Panel B depicts the optimal rates as functions of the marginal non-pecuniary cost λ. Similarly, non-pecuniary costs impede the transmission of the short rate to bond yields, necessitating a higher policy target when the marginal non-pecuniary cost is high. Notably, the required optimal rate in the yield economy is generally higher than in the repo economy. This is because the short rate's impact on macroeconomic variables is weaker in the yield economy, prompting the central bank to adopt a more ambitious policy target.

10 Conclusion

References

Acharya, Viral V, and Toomas Laarits. 2023. "When Do Treasuries Earn the Convenience Yield?: A Hedging Perspective."

Adrian, Tobias, Richard K Crump, and Erik Vogt. 2019. "Nonlinearity and flight-to-safety in the risk-return trade-off for stocks and bonds." *The Journal of Finance*, 74(4): 1931–1973.

Angeletos, George-Marios, and Chen Lian. 2018. "Forward guidance without common knowledge." *American Economic Review*, 108(9): 2477–2512.

Arrata, William, Benoît Nguyen, Imène Rahmouni-Rousseau, and Miklos Vari. 2020. "The scarcity effect of QE on repo rates: Evidence from the euro area." *Journal of Financial Economics*, 137(3): 837–856.

Augustin, Patrick, Mikhail Chernov, Lukas Schmid, and Dongho Song. 2021. "Benchmark interest rates when the government is risky." *Journal of Financial Economics*, 140(1): 74–100.

Baker, Scott R, Nicholas Bloom, Steven J Davis, Kyle Kost, Marco Sammon, and Tasaneeya Viratyosin. 2020. "The unprecedented stock market reaction to COVID-19." *The review of asset pricing studies*, 10(4): 742–758.

Baklanova, Viktoria, Adam M Copeland, and Rebecca McCaughrin. 2015. "Reference guide to US repo and securities lending markets." *FRB of New York Staff Report*, , (740).

Baklanova, Viktoria, Cecilia Caglio, Marco Cipriani, and Adam Copeland. 2019. "The use of collateral in bilateral repurchase and securities lending agreements." *Review of Economic Dynamics*, 33: 228–249.

Baklanova, Viktoria, Isaac Kuznits, and Trevor Tatumetal. 2021. "Primer: Money Market Funds and the Repo Market." *Division of Investment Management's Analytics Office of the US Securities and Exchange Commission, February*, 18.

Banegas, Ayelen, Phillip J Monin, and Lubomir Petrasek. 2021. "Sizing hedge funds' Treasury market activities and holdings."

Banerjee, Snehal, and Jeremy J Graveline. 2013. "The cost of short-selling liquid securities." *The Journal of Finance*, 68(2): 637–664.

Bansal, Ravi, Wilbur John Coleman, and Christian T Lundblad. 2010. "Endogenous liquidity supply."

Beber, Alessandro, Michael W Brandt, and Kenneth A Kavajecz. 2009. "Flight-to-quality or flight-to-liquidity? Evidence from the euro-area bond market." *The Review of Financial Studies*, 22(3): 925–957.

Bodenstein, Martin, Pablo Cuba-Borda, Nils Gornemann, Ignacio Presno, Andrea Prestipino, Albert Queralto, and Andrea Raffo. 2023. "Global Flight to Safety, Business Cycles, and the Dollar." *International Finance Discussion Paper*, , (1381).

Christiano, Lawrence J, Martin Eichenbaum, and Charles L Evans. 2005. "Nominal rigidities and the dynamic effects of a shock to monetary policy." *Journal of political Economy*, 113(1): 1–45.

Christiano, Lawrence J, Roberto Motto, and Massimo Rostagno. 2014. "Risk shocks." *American Economic Review*, 104(1): 27–65.

Cochran, Paul, Sebastian Infante, Lubomir Petrasek, Zack Saravay, and Mary Tian. 2023. "Dealers' Treasury Market Intermediation and the Supplementary Leverage Ratio."

Cooperman, Harry R, Darrell Duffie, Stephan Luck, Zachry Wang, Yilin Yang, et al. 2023. "Bank funding risk, reference rates, and credit supply."

Copeland, Adam, Isaac Davis, Eric LeSueur, and Antoine Martin. 2014. "Lifting the veil on the us bilateral repo market."

Corradin, Stefano, and Angela Maddaloni. 2020. "The importance of being special: repo markets during the crisis." *Journal of Financial Economics*, 137(2): 392–429.

Duffie, Darrell. 1996. "Special repo rates." *The Journal of Finance*, 51(2): 493–526.

Duffie, Darrell. 2018. "Financial regulatory reform after the crisis: An assessment." *Management Science*, 64(10): 4835–4857.

Du, Wenxin, Joanne Im, and Jesse Schreger. 2018. "The us treasury premium." *Journal of International Economics*, 112: 167–181.

D'Amico, Stefania, and Thomas B. King. 2013. "Flow and stock effects of large-scale treasury purchases: Evidence on the importance of local supply." *Journal of Financial Economics*, 108(2): 425–448.

d'Amico, Stefania, Roger Fan, and Yuriy Kitsul. 2018. "The scarcity value of treasury collateral: Repo-market effects of security-specific supply and demand factors." *Journal of Financial and Quantitative Analysis*, 53(5): 2103–2129.

Eisenbach, Thomas M, and Gregory Phelan. 2023. "Fragility of safe asset markets."

Eisenschmidt, Jens, Yiming Ma, and Anthony Lee Zhang. 2024. "Monetary policy transmission in segmented markets." *Journal of Financial Economics*, 151: 103738.

Fama, Eugene F, and Robert R Bliss. 1987. "The information in long-maturity forward rates." *The American Economic Review*, 680–692.

Gallmeyer, Michael, Burton Hollifield, Francisco Palomino, and Stanley E Zin. 2007. "Arbitrage-free bond pricing with dynamic macroeconomic models."

Gallmeyer, Michael F, Burton Hollifield, and Stanley E Zin. 2005. "Taylor rules, McCallum rules and the term structure of interest rates." *Journal of Monetary Economics*, 52(5): 921–950.

Goldreich, David, Bernd Hanke, and Purnendu Nath. 2005. "The price of future liquidity: Time-varying liquidity in the US Treasury market." *Review of Finance*, 9(1): 1–32.

Greenwood, Robin, and Dimitri Vayanos. 2014. "Bond supply and excess bond returns." *The Review of Financial Studies*, 27(3): 663–713.

Guerrieri, Veronica, and Guido Lorenzoni. 2017. "Credit crises, precautionary savings, and the liquidity trap." *The Quarterly Journal of Economics*, 132(3): 1427–1467.

Gürkaynak, Refet S, Brian Sack, and Jonathan H Wright. 2007. "The US Treasury yield curve: 1961 to the present." *Journal of monetary Economics*, 54(8): 2291–2304.

Harkrader, James C, and Daniel J Weitz. 2020. "How Do Principal Trading Firms and Dealers Trade around FOMC Statement Releases?"

He, Zhiguo, Stefan Nagel, and Zhaogang Song. 2022. "Treasury inconvenience yields during the COVID-19 crisis." *Journal of Financial Economics*, 143(1): 57–79.

Huber, Amy Wang. 2023. "Market power in wholesale funding: A structural perspective from the triparty repo market." *Journal of Financial Economics*, 149(2): 235–259.

Ihrig, Jane E, Elizabeth Klee, Canlin Li, Brett Schulte, and Min Wei. 2012. "Expectations about the Federal Reserve's balance sheet and the term structure of interest rates."

Jermann, Urban J. 2019. "Is SOFR better than LIBOR?" *Available at SSRN 3361942*.

J Jermann, Urban. 2020. "Negative swap spreads and limited arbitrage." *The Review of Financial Studies*, 33(1): 212–238.

Jordan, Bradford D, and Susan D Jordan. 1997. "Special repo rates: An empirical analysis." *The Journal of Finance*, 52(5): 2051–2072.

Kekre, Rohan, and Moritz Lenel. 2024. "The flight to safety and international risk sharing." *American Economic Review*, 114(6): 1650–1691.

Klingler, Sven, and Suresh Sundaresan. 2019. "An explanation of negative swap spreads: Demand for duration from underfunded pension plans." *The journal of finance*, 74(2): 675–710.

Klingler, Sven, and Suresh Sundaresan. 2023. "Diminishing Treasury convenience premiums: Effects of dealers' excess demand and balance sheet constraints." *Journal of Monetary Economics*, 135: 55–69.

Krishnamurthy, Arvind, and Annette Vissing-Jorgensen. 2012. "The aggregate demand for treasury debt." *Journal of Political Economy*, 120(2): 233–267.

Longstaff, Francis A. 2002. "The flight-to-liquidity premium in US Treasury bond prices."

Mazur, Mieszko, Man Dang, and Miguel Vega. 2021. "COVID-19 and the march 2020 stock market crash. Evidence from S&P1500." *Finance research letters*, 38: 101690.

Moskowitz, Tobias J, Chase P Ross, Sharon Y Ross, and Kaushik Vasudevan. 2024. "Quantities and Covered-Interest Parity." *Available at SSRN 4820243*.

Nagel, Stefan. 2016. "The liquidity premium of near-money assets." *The Quarterly Journal of Economics*, 131(4): 1927–1971.

Ray, Walker. 2019. "Monetary policy and the limits to arbitrage: Insights from a new keynesian preferred habitat model." 692.

Ray, Walker, and Rupal Kamdar. 2024. "Optimal Macro-Financial Stabilization in a New Keynesian Preferred Habitat Model."

Sims, Eric, Jing Cynthia Wu, and Ji Zhang. 2023. "The four-equation new keynesian model." *Review of Economics and Statistics*, 105(4): 931–947.

Swanson, Eric T. 2021. "Measuring the effects of federal reserve forward guidance and asset purchases on financial markets." *Journal of Monetary Economics*, 118: 32–53.

Vayanos, Dimitri, and Jean-Luc Vila. 2021. "A preferred-habitat model of the term structure of interest rates." *Econometrica*, 89(1): 77–112.

Werning, Ivan. 2011. "Managing a liquidity trap: Monetary and fiscal policy."

Appendix A: The Simple Case

Proof of LEMMA 1.

The only risk in this simple economy is the short rate:

$$dr_t = -\kappa_r(r_t - r^{ss})dt + \sigma_r dB_{r,t}.$$

It follows that the equilibrium bond prices are

$$P_t(\tau) = exp\Big[- \Big(A_r(\tau) r_t + C(\tau) \Big) \Big].$$

Using Ito's Lemma, the instantaneous return is

$$\frac{dP_t(\tau)}{P_t(\tau)} = \mu_t(\tau)dt - A_r(\tau)\sigma_r dB_{r,t},$$

where

$$\mu_t(\tau) \equiv A_r'(\tau)r_t + C'(\tau) + A_r(\tau)\kappa_r\left(r_t - r^{ss}\right) + \frac{1}{2}A_r(\tau)^2\sigma_r^2.$$

In this simple case, the FOC for hedge funds is

$$\mu_t(\tau) - r_t = 0$$

Collecting terms related to r_t, I have

$$A_r'(\tau) + A_r(\tau)\kappa_r - 1 = 0.$$

The solution is

$$A_r(\tau) = \frac{1 - e^{-\kappa_r \tau}}{\kappa_r}.$$

□

Proof of LEMMA 2.

The equilibrium system is given by

$$d\begin{bmatrix} r_t \\ x_t \end{bmatrix} = -\underbrace{\begin{bmatrix} \psi_r & -\psi_r\phi_x \\ -\varsigma^{-1}\hat{A}_r & 0 \end{bmatrix}}_{\Upsilon}\left(\begin{bmatrix} r_t \\ x_t \end{bmatrix} - \begin{bmatrix} r^{ss} \\ x^{ss} \end{bmatrix}\right)dt + \begin{bmatrix} \sigma_r & 0 \\ 0 & 0 \end{bmatrix} d\begin{bmatrix} B_{r,t} \\ B_{\beta,t} \end{bmatrix}.$$

The general equilibrium solution is

$$r_t = -\kappa_r(r_t - r^{ss})dt + \sigma_r dB_{r,t}$$
$$x_t = x^{ss} + \omega_x(r_t - r^{ss}),$$

where κ_r is a positive eigenvalue of Υ. For this simple matrix Υ, its eigenvalue can be found by solving for κ_r in the following equation:

$$-\kappa_r(\psi_r - \kappa_r) - \psi_r\phi_x\varsigma^{-2}\hat{A}_r = 0$$
$$\Rightarrow \hat{A}_r = \frac{\kappa_r(\kappa_r - \psi_r)}{\psi_r\phi_x\varsigma^{-1}}$$

Denoting $f(c) = c(c - \psi_r) - \psi_r\phi_x\varsigma^{-2}\hat{A}_r$, then I have $f0) = f(\psi_r) = -\psi_r\phi_x\varsigma^{-2}\hat{A}_r$. When $\hat{A}_r > 0$, $f0) = f(\psi_r) < 0$. The function $f(c) = 0$ has two roots with one being positive and the other one being negative. The mean reverting coefficient κ_r takes the value of the only positive eigenvalue.

The eigenvector associated with the positive eigenvalue is $q_1 = \begin{bmatrix} -\frac{\kappa_r}{\varsigma^{-1}\hat{A}_r} \\ 1 \end{bmatrix} = \begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix}$. For the simple case, the dynamics of the output gap x_t is fully characterized by

$$\omega_x = q_{12}/q_{11} = -\frac{\varsigma^{-1}\hat{A}_r}{\kappa_r}$$
$$= -\frac{\varsigma^{-1}}{\kappa_r}\frac{\kappa_r(\kappa_r - \psi_r)}{\psi_r\phi_x\varsigma^{-1}}$$
$$= \frac{\psi_r - \kappa_r}{\psi_r\phi_x}.$$

□

Appendix B: Solution of the Generic Model

B.1 Equilibrium Affine Coefficients

Substituting $\mu_t(\tau)$ using equation (17), $H_t(\tau)$ using equation (7), and $P_t(\tau)$ using equation (15), I can rewrite equation (24) as

$$A'(\tau)'s_t + C'(\tau) + A(\tau)'\Gamma\left(s_t - r^{ss}\epsilon\right) + \frac{1}{2}A(\tau)'\Sigma\Sigma'A(\tau) - \underbrace{r_t}_{\epsilon' s_t} - \lambda\Big[\theta(\tau)\underbrace{\beta_t}_{\gamma' s_t} - \alpha(\tau)\Big(A(\tau)'s_t + C(\tau)\Big)\Big]$$

$$= \frac{1}{\rho_d + \rho_h}A(\tau)'\Sigma\Sigma'\left[\int_0^T \Big[\theta(\tau)\gamma' s_t - \alpha(\tau)\Big(A(\tau)'s_t + C(\tau)\Big)\Big]A(\tau)d\tau\right],$$

where $\epsilon = [1\ 0]'$ and $\gamma = [0\ 1]'$ are vectors selecting r_t and β_t respectively from the risk factor vector s_t. Collecting terms related to s_t, I have

$$A'(\tau) + \Gamma' A(\tau) - \epsilon + \lambda\Big[\alpha(\tau)A(\tau) - \theta(\tau)\gamma\Big] = \frac{1}{\rho_d + \rho_h}\left[\int_0^T \Big[\theta(\tau)\gamma - \alpha(\tau)A(\tau)\Big]A(\tau)'d\tau\right]\Sigma\Sigma'A(\tau). \tag{B.1}$$

Then the remaining terms consist of the following equation

$$C'(\tau) - r^{ss}A(\tau)'\Gamma\epsilon + \frac{1}{2}A(\tau)'\Sigma\Sigma'A(\tau) + \lambda\alpha(\tau)C(\tau) = -\frac{1}{\rho_d + \rho_h}A(\tau)'\Sigma\Sigma'\Big[\int_0^T \alpha(\tau)C(\tau)A(\tau)d\tau\Big]. \tag{B.2}$$

Equation (B.1) is a two-dimensional vector equation that can be used to pin down $A(\tau)$. With $A(\tau)$ solved, I can use equation (B.2) to find $C(\tau)$. Then the term structures for bond yields and repo rates are given as functions of bond prices. Unfortunately, these two equations do not have analytic solutions except in soe simple cases. The general form has to be solved numerically.

B.2 Rational Expectations Linear System

Let $Y_t = [x_t, y_t]'$ be the vector of all variables where x_t is the vector of state variables and y_t is the vector of jump variables. Y_t follows the process

$$dY_t = -\Upsilon(Y_t - Y^{ss})dt + [S, 0]'dB_t. \tag{B.3}$$

The eigen-decomposition of Υ is

$$\Upsilon = Q\Lambda Q^{-1},$$

where Λ is the diagonal matrix with eigenvalues on the diagonal and Q is the eigenvector matrix with eigenvectors on the columns. I can divide the two matrices into blocks:

$$\Lambda = \begin{bmatrix} \Lambda_x & 0 \\ 0 & \Lambda_y \end{bmatrix}, \; Q = \begin{bmatrix} Q_{xx} & Q_{xy} \\ Q_{yx} & Q_{yy}, \end{bmatrix}$$

where the partition corresponds to the state and jump vectors. If Υ is full-ranked and the number of negative eigenvalues is the same as that of the state variables, then the solution to equation (B.3) is equation (32) and (33) with

$$\Gamma = Q_{xx}\Lambda_x Q_{xx}^{-1}, \tag{B.4}$$

$$\Omega = Q_{yx}Q_{xx}^{-1}. \tag{B.5}$$

B.3 Solution Algorithm

The key general equilibrium objective is the transition matrix for the state variables Γ, which is pinned down by the coincidence between the macroeconomic and the term structure sides of the model. The solution algorithm can be summarized as a root-finding problem involving two layers of numerical computation.

1. *Given value of Γ, do the following.*
   - (a) *Numerically solve for the Affine coefficients $A_r(\tau)$ and $A_\beta(\tau)$ according to asset pricing equations (B.1), which is a system of ordinary differential equations involving integrals.*
     - i. *Treat the intergral terms as given, solve for $A_r(\tau)$ and $A_\beta(\tau)$.*
     - ii. *Update the integral terms with the solution of $A_r(\tau)$ and $A_\beta(\tau)$.*
     - iii. *Repeat step i until the solution converges.*
   - (b) *Using the Affine coefficients calculated in step (a), compute the aggregate coefficients $\hat{A}_r$ and $\hat{A}_\beta$ according to the system of equations (28).*
   - (c) *Using the aggregate coefficients calculated in step (b), construct the parameter matrix for the rational expectation system Υ according to equation (31).*

(d) Recalculate the transition matrix from the macroeconomic side of the model according to equation (B.4). Denote this recalculated transition matrix as $\Gamma^*(\Gamma)$*.*

2. *The general equilibrium is defined as a root finding problem of* Γ *such that* $F(\Gamma) = \Gamma - \Gamma^*(\Gamma) = 0$*.*

Appendix C: Calibration

C.1 Macroeconomic Parameters

After simplification and partial parameterization, the coefficient matrix is given by:

$$\Upsilon = \begin{bmatrix} \psi_r & -\psi_r\phi_x & -\psi_r\phi_\pi \\ -\hat{A}_r & 0 & 1 \\ 0 & \delta & -0.04 \end{bmatrix}.$$

The characteristic polynomial of this matrix is

$$|\Upsilon - \lambda I| = (\psi_r - \lambda)[\lambda(0.04 + \lambda) - \delta] + \psi_r\phi_x\hat{A}_r(0.04 + \lambda) + \psi_r\phi_\pi\hat{A}_r\delta = 0. \tag{C.1}$$

There must be an eigenvalue that is equal to κ_r, the short rate reverting speed in the general equilibrium. Therefore, equation (C.1) must satisfy

$$(\psi_r - \kappa_r)[\kappa_r(0.04 + \kappa_r) - \delta] + \psi_r\phi_x\hat{A}_r(0.04 + \kappa_r) + \psi_r\phi_\pi\hat{A}_r\delta = 0. \tag{C.2}$$

The eigenvector associated with that eigenvalue is defined as a three-dimensional vector $v = [v_1, v_2, v_3]'$ such that

$$\begin{bmatrix} \psi_r - \kappa_r & -\psi_r\phi_x & -\psi_r\phi_\pi \\ -\hat{A}_r & -\kappa_r & 1 \\ 0 & \delta & -0.04 - \kappa_r \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = 0$$

$$\Rightarrow \begin{cases} (\psi_r - \kappa_r)v_1 - \psi_r\phi_x v_2 - \psi_r\phi_\pi v_3 = 0 \\ \hat{A}_r v_1 + \kappa_r v_2 - v_3 = 0 \\ \delta v_2 - (0.04 + \kappa_r)v_3 = 0 \end{cases}$$

From the last equation, I can derive that

$$\frac{v_3}{v_2} = \frac{\delta}{0.04 + \kappa_r}. \tag{C.3}$$

From the second equation, I can derive that

$$\frac{v_2}{v_1} = \frac{\hat{A}_r}{\frac{\delta}{0.04+\kappa_r} - \kappa_r}.$$

Recall that the output matrix is a function of elements in v. From equation (B.5), I can map that

$$\omega_{xr} = v_2/v_1 \Rightarrow cov(x,r) = \frac{\hat{A}_r}{\frac{\delta}{0.04+\kappa_r} - \kappa_r},$$
$$\omega_{\pi r} = v_3/v_1 \Rightarrow cov(\pi,r) = (\frac{\delta}{0.04 + \kappa_r})(\frac{\hat{A}_r}{\frac{\delta}{0.04+\kappa_r} - \kappa_r}).$$

I can further derive

$$\frac{cov(\pi,r)}{cov(x,r)} = \frac{\delta}{0.04 + \kappa_r} \tag{C.4}$$

$$\Rightarrow cov(x,r) = \frac{\hat{A}_r}{\frac{cov(\pi,r)}{cov(x,r)} - \kappa_r}. \tag{C.5}$$

Note that $\hat{A}_r$ is a function of κ_r. Only one unknown κ_r is in equation (C.5), from which I can numerically find the value of κ_r. Once κ_r is found, I can use equation (C.4) to find

$$\delta = (0.04 + \kappa_r)\frac{cov(\pi,r)}{cov(x,r)}. \tag{C.6}$$

Replacing δ in equation (C.2) using equation (C.6), what I am left with is

$$\psi_r = \frac{\kappa_r}{1 - \phi_x \omega_{xr} - \phi_\pi \omega_{\pi r}}.$$

Finally, the volatility of short rate σ_r can be identified using $var(r)$:

$$var(r) = \frac{\sigma_r^2}{2 * \kappa_r}$$
$$\Rightarrow \sigma_r = \sqrt{2\kappa_r var(r)}.$$

C.2 Term Structure Parameters

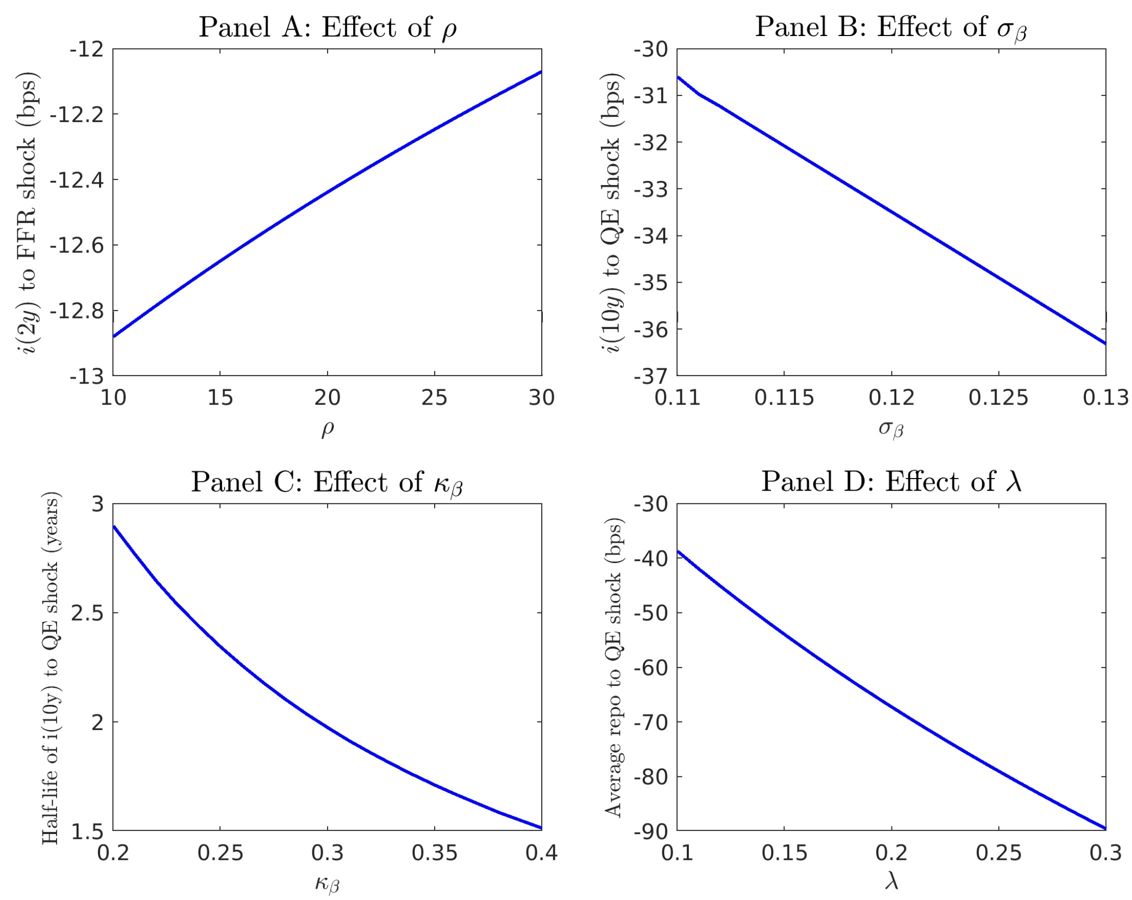


Figure 31: Sensitivity analysis for term-structure parameters.

Note: Parameters are set at optimal values except the one being tested in each panel.

Appendix D: Sims et al. (2023)

D.1 Discrete to Continuous-Time

The discrete-time IS and Phillips curves in Sims et al. (2023) are

$$
\begin{aligned}
x_t =& \mathbb{E}_t x_{t+1} - \frac{1-z}{\sigma}\left(r_t - \mathbb{E}_t \pi_{t+1} - r^*\right) - z\left(\mathbb{E}_t qe_{t+1} - qe_t\right), \\
\pi_t =& \gamma\zeta x_t - \frac{z\gamma\sigma}{1-z} qe_t + \beta \mathbb{E}_t \pi_{t+1}.
\end{aligned}
$$

Let $\frac{x_{t+1}-x_t}{t+1-t} \approx \frac{dx_t}{dt}$, then the IS the Phillips curves become

$$
\begin{aligned}
dx_t &= \frac{1-z}{\sigma}\left(r_t - r^* - \frac{d\pi_t}{dt} - \pi_t\right) dt - z dqe_t, \\
d\pi_t &= \left(\frac{1-\beta}{\beta}\pi_t - \frac{\gamma\varsigma}{\beta} x_t + \frac{z\gamma\sigma}{(1-z)\beta} qe_t\right) dt.
\end{aligned}
$$

Substituting $d\pi_t/dt$ in the IS curve using the Phillips curve, the IS curve becomes

$$
dx_t = \frac{1-z}{\sigma}\left(r_t - r^* - \frac{1}{\beta}\pi_t + \frac{\gamma\varsigma}{\beta} x_t - \frac{z\gamma\sigma}{(1-z)\beta} qe_t\right) dt - z dqe_t.
$$

Adding the same Taylor rule and QE process,

$$
\begin{aligned}
dr_t &= -\psi_r(r_t - \phi_x x_t - \phi_\pi \pi_t) + \sigma_r dB_{r,t}, \\
dqe_t &= -\kappa_\beta qe_t dt + \sigma_\beta dB_{\beta,t},
\end{aligned}
$$

I can further rewrite IS curve as

$$
dx_t = \frac{1-z}{\sigma}\left(r_t - r^* - \frac{1}{\beta}\pi_t + \frac{\gamma\varsigma}{\beta} x_t - \left(\frac{z\gamma\sigma}{(1-z)\beta} - z\kappa_\beta\right) qe_t\right) dt - z\sigma_\beta dB_{\beta,t}.
$$

Collecting terms in matrix form, I have

$$d\begin{bmatrix} r_t \\ \beta_t \\ x_t \\ \pi_t \end{bmatrix} = -\underbrace{\begin{bmatrix} \psi_r & 0 & -\psi_r\phi_x & -\psi_r\phi_\pi \\ 0 & \kappa_\beta & 0 & 0 \\ -\frac{1-z}{\sigma} & \frac{1-z}{\sigma}\left(\frac{z\gamma\sigma}{(1-z)\beta} - z\kappa_\beta\right) & -\frac{1-z}{\sigma}\frac{\gamma\varsigma}{\beta} & \frac{1-z}{\sigma}\frac{1}{\beta} \\ 0 & -\frac{z\gamma\sigma}{(1-z)\beta} & \frac{\gamma\varsigma}{\beta} & 1-\frac{1}{\beta} \end{bmatrix}}_{\Upsilon} \left(\begin{bmatrix} r_t \\ \beta_t \\ x_t \\ \pi_t \end{bmatrix} - \begin{bmatrix} r^{ss} \\ \beta^{ss} \\ x^{ss} \\ \pi_{ss} \end{bmatrix} \right) dt + \begin{bmatrix} \sigma_r & 0 \\ 0 & \sigma_\beta \\ 0 & -b \\ 0 & 0 \end{bmatrix} d\begin{bmatrix} B_{r,t} \\ B_{\beta,t} \end{bmatrix}. \tag{D.1}$$

D.2 Recalibrate the Model

In the following analysis, I present a method to recalibrate the model with and without the wage channel for QE to affect the inflation rate. Let $a = \frac{1-z}{\sigma}$, $b = \frac{z\gamma\sigma}{(1-z)\beta} - z\kappa_\beta$, $c = \frac{\gamma\varsigma}{\beta}$, $d = \frac{1}{\beta}$, $e = \frac{z\gamma\sigma}{(1-z)\beta}$ then the coefficient matrix in equation (D.1) can be simplified as

$$\Upsilon = \begin{bmatrix} \psi_r & 0 & -\psi_r\phi_x & -\psi_r\phi_\pi \\ 0 & \kappa_\beta & 0 & 0 \\ -a & ab & -ac & ad \\ 0 & -e & c & 1-d \end{bmatrix}.$$

Given the coefficient matrix Υ, let $v = \begin{bmatrix} v1 & v2 & v3 & v4 \end{bmatrix}'$ denote the eigenvector for eigenvalue κ_r and let $u = \begin{bmatrix} u1 & u2 & u3 & u4 \end{bmatrix}'$ denote the eigenvector for eigenvalue κ_β. Since there are two state variables and two jump variables, I divide the eigenvalue and egenvector matrices as follows

$$\Lambda = \begin{bmatrix} \kappa_r & & & \\ & \kappa_\beta & & \\ & & \lambda_3 & \\ & & & \lambda_4 \end{bmatrix} = \begin{bmatrix} \Lambda_{xx} & \\ & \Lambda_{yy} \end{bmatrix},$$

$$Q = \begin{bmatrix} q(\kappa_r) & q(\kappa_\beta) & q_3 & q_4 \end{bmatrix} = \begin{bmatrix} v1 & u1 & q_{31} & q_{41} \\ v2 & u2 & q_{32} & q_{42} \\ v3 & u3 & q_{33} & q_{43} \\ v4 & u4 & q_{34} & q_{44} \end{bmatrix} = \begin{bmatrix} Q_{xx} & Q_{xy} \\ Q_{yx} & Q_{yy} \end{bmatrix}.$$

Then the solution of the four-equation system is

$$\Gamma = Q_{xx}\Lambda_x Q_{xx}^{-1},$$
$$\Omega = Q_{xx}^{-1} Q_{yx}.$$

where

$$\Lambda_{xx} = \begin{bmatrix} \kappa_r & \\ & \kappa_\beta \end{bmatrix}, Q_{xx} = \begin{bmatrix} v1 & u1 \\ v2 & u2 \end{bmatrix}, Q_{yx} = \begin{bmatrix} v3 & u3 \\ v4 & u4 \end{bmatrix}.$$

To facilitate a better understanding of the system solution, below I do the computation explicitly.

$$\begin{aligned} Q_{xx}^{-1} &= \frac{1}{v1u2 - v2u1} \begin{bmatrix} u2 & -u1 \\ -v2 & v1 \end{bmatrix} \\ \Rightarrow \Omega = Q_{yx} Q_{xx}^{-1} &= \frac{1}{v1u2 - v2u1} \begin{bmatrix} u2v3 - u3v2 & u3v1 - v3u1 \\ v4u2 - u4v2 & u4v1 - v4u1 \end{bmatrix}, \\ \Rightarrow \Gamma = Q_{xx}\Lambda_x Q_{xx}^{-1} &= \frac{1}{v1u2 - v2u1} \begin{bmatrix} \kappa_r v1 & u1\kappa_\beta \\ v2\kappa_r & u2\kappa_\beta \end{bmatrix} \begin{bmatrix} u2 & -u1 \\ -v2 & v1 \end{bmatrix} \\ &= \frac{1}{v1u2 - v2u1} \begin{bmatrix} \kappa_r v1u2 - u1\kappa_\beta v2 & u1\kappa_\beta v1 - \kappa_r v1u1 \\ v2\kappa_r u2 - u2\kappa_\beta v2 & u2\kappa_\beta v1 - v2\kappa_r u1 \end{bmatrix}. \end{aligned}$$

Since β_t is determined exogenously, $v2 = 0$. The solutions reduce to

$$\begin{aligned} \Omega &= \frac{1}{v1u2} \begin{bmatrix} u2v3 & u3v1 - v3u1 \\ v4u2 & u4v1 - v4u1 \end{bmatrix} = \begin{bmatrix} v3/v1 & \frac{u3v1 - v3u1}{v1u2} \\ v4/v1 & \frac{u4v1 - v4u1}{v1u2} \end{bmatrix}, \\ \Gamma &= \frac{1}{v1u2} \begin{bmatrix} \kappa_r v1u2 & u1\kappa_\beta v1 - \kappa_r v1u1 \\ 0 & u2\kappa_\beta v1 \end{bmatrix} = \begin{bmatrix} \kappa_r & \frac{u1}{u2}(\kappa_\beta - \kappa_r) \\ 0 & \kappa_\beta \end{bmatrix}. \end{aligned}$$

I pick $\kappa_\beta = 0.2$ to match a 0.8 AR(1) coefficient in the discrete model. κ_r is calibrated to be 0.46 to match a 1.5-quarter half life of the conventional monetary policy shock. Here are other moments collected from Sims et al. (2023): $\hat{\omega_{xr}}$=-0.7; $\hat{\omega_{\pi r}}$=-1.35; $\hat{\omega_{x\beta}}/\hat{\omega_{r\beta}}$=5.86; $\hat{\omega_{x\beta}}/\hat{\omega_{\pi\beta}}$=1.82. The following relations hold

$$\hat{\omega_{xr}} = v3/v1, \tag{D.2}$$

$$\hat{\omega_{\pi r}} = v4/v1, \tag{D.3}$$

$$\frac{u3v1 - v3u1}{v1u2} = \hat{\omega_{x\beta}}/\hat{\omega_{r\beta}}\frac{u1}{u2}(\kappa_\beta - \kappa_r), \tag{D.4}$$

$$\hat{\omega_{x\beta}}/\hat{\omega_{\pi\beta}} = \frac{u3v1 - v3u1}{u4v1 - v4u1}, \tag{D.5}$$

$$v1^2 + v3^2 + v4^2 = 1, \tag{D.6}$$

$$u1^2 + u2^2 + u3^2 + u4^2 = 1. \tag{D.7}$$

Taking κ_r and κ_β as given, there are six equations (D.2-D.7) defining seven unknowns $v = \begin{bmatrix} v1 & 0 & v3 & v4 \end{bmatrix}'$, $u = \begin{bmatrix} u1 & u2 & u3 & u4 \end{bmatrix}'$. I can randomly give the value for one unknown then derive the other six unknowns using the equations above. In the exercise, instead I impose another normalization condition $\hat{\omega}_{x\beta} = 0.01$ to achieve the exact identification of the eigenvectors.

Solving the eigenvectors v and u using (D.2-D.7), now I can back out the coefficient values. Because v and u are eigenvectors of Υ, the following relations must hold

$$(\psi_r - \kappa_r)v1 - \psi_r\phi_x v3 - \psi_r\phi_\pi v4 = 0, \tag{D.8}$$

$$(\psi_r - \kappa_\beta)u1 - \psi_r\phi_x u3 - \psi_r\phi_\pi u4 = 0, \tag{D.9}$$

$$-a * v1 + ab * v2 - (ac + \kappa_r) * v3 + ad * v4 = 0, \tag{D.10}$$

$$-a * u1 + ab * u2 - (ac + \kappa_\beta) * u3 + ad * u4 = 0, \tag{D.11}$$

$$-e * v2 + c * v3 + (1 - d - \kappa_r) * v4 = 0, \tag{D.12}$$

$$-e * u2 + c * u3 + (1 - d - \kappa_\beta) * u4 = 0. \tag{D.13}$$

From equations (D.8) and (D.9), the three Taylor rule coefficients ψ_r, ϕ_x, and ϕ_π are under-identified. For simplicity, I assume $\psi_r = 0.2$ then use these two equations to calculate the other two coefficients.

$$\psi_r = \frac{\kappa_\beta * u1 * v4 - \kappa_r * v1 * u4}{u1 * v4 - v1 * u4},$$
$$\psi_\pi = \frac{(\psi_r - \kappa_r) * v1}{\psi_r * v4}.$$

Furthermore, recall that $v2 = 0$ since the QE is exogenous. From equations (D.10) and (D.12),

$$\begin{aligned} ac * v3 &= ad * v4 - a * v1 - \kappa_r * v3, \\ c * v3 &= (\kappa_r + d - 1)v4, \\ \Rightarrow a &= \frac{\kappa_r * v3}{v4 - \kappa_r * v4 - v1}. \end{aligned}$$

Now I am left with three restrictions but four parameters. For simplicity, I assume $d = 1/0.995$, then from equation (D.12),

$$c = (\kappa_r + d - 1)v4/v3.$$

From equation (D.13),

$$e = c * u3/u2 + (1 - d - \kappa_\beta) * u4/u2.$$

From equation (D.11),

$$b = \frac{a * u1 + (ac + \kappa_\beta) * u3 + ad * u4}{a * u2}.$$

D.3 Recalibrate a Version without Wage Channel

To calibrate a version of the model without the wage channel, let $e = 0$. Notice that in this case, there is no direct effect from the QE on inflation, $u3 = u4 = 0$. To compute the eigenvectors satisfying the targeted moments, I impose the condition $u3 = u4 = 0$ on equation (D.4)

$$\frac{-v3}{v1} = \hat{\omega_{x\beta}}/\hat{\omega_{r\beta}}(\kappa_\beta - \kappa_r),$$

which is equal to $-\omega_{xr}$. Therefore, the moment $\hat{\omega_{x\beta}}/\hat{\omega_{r\beta}}$ cannot be matched in this special case. Likewise, imposing the condition $u3 = u4 = 0$ on equation (D.5)

$$\hat{\omega_{x\beta}}/\hat{\omega_{\pi\beta}} = \frac{v3}{v4},$$

which is equal to $\hat{\omega_{xr}}/\hat{\omega_{\pi r}}$. Therefore, the moment $\hat{\omega_{x\beta}}/\hat{\omega_{\pi\beta}}$ cannot be matched in this special case as well. The eigenvectors are computed based on the system (D.2, D.3, D.6, and D.7) along with the normalization $\hat{\omega}_{x\beta} = 0.01$.

To back out the coefficients, imposing $u3 = u4 = 0$ on equation (D.9), I have

$$\psi_r = \kappa_\beta.$$

Assuming $\phi_x = 0$, I can use equation (D.8) to calculate ϕ_π. Likewise, assuming $d = 1/0.995$, I can use equation (D.12) to calculate c. Furthermore, from equations (D.10) and (D.12),

$$a = \frac{\kappa_r * v3}{v4 - \kappa_r * v4 - v1}.$$

Imposing $u3 = u4 = 0$ on equation (D.11), I have

$$b = u1/u2.$$

Appendix E: Micro Foundations for IS Curve

E.1 Overview

I this sector, I lay out a continuous-time New-Keynesian model, in the same spirit of Ray and Kamdar (2024), that results in aggregate nominal interest rate in the format of equation (1).

- Government: supplys government bonds aggregated to 0.
- Central bank: sets policy rate using Taylor Rule.
- Private habitat investors: demand bonds for specific maturities.
- Financial intermediaries (Arbitrageurs):
  - demand bonds as arbitrageurs;
  - take savings as given, repay promised interests, transfer the gains/losses to HHs;
  - dealers provide repo cash in exchange of repo bonds to hedge funds.
- Households:
  - consume final goods, provide labor to intermediate firms;

- save at differentiated interest rates;
- receives profits and price adjustment costs from intermediate firms, and gains/losses from financial intermediaries.

- Firms:
  - Final goods producer aggregates intermediate goods to final goods.
  - Monopolistically competitive intermediate goods producers demand labor and set price with adjustment cost to maximize profits.

E.2 Households

Continuum of HHs differentiated by access to saving markets j. There is a mass $h(j)$ for each type of HH and the aggregate mass $\int_0^J h(j)dj = 1$. HH j chooses consumption, labor, and saving to maximize expected life-time utility subject to budget constraint. Each type of HHs is differentiated by the interest rate to which they have access to $i_t(j)$. HHs own firms and arbitrageurs and receive the equal transfer from the sum of firm profits, price adjustment costs, and arbitrageurs' gains/losses dT_t. I assume that households value the quality of safe saving products, reflected by the $\varpi_t k_t(j)$ term in the utility function, with ϖ_t interpreted as the time-varying marginal quality value.

$$\begin{aligned} &\max_{c_t(j),n_t(j),k_t(j)} \int_0^\infty e^{-\rho t}\left(\log c_t(j) - \frac{n_t(j)^{1+\phi}}{1+\phi} + \varpi_t k_t(j)\right) dt, \\ &s.t. \\ &dk_t(j) = (i_t(j)k_t(j) + W_t n_t(j) - P_t c_t(j))dt + dT_t. \end{aligned} \tag{E.1}$$

Current-value Hamiltonian:

$$H(c_t(j), k_t(j), \lambda_t(j)) = \log c_t(j) - \frac{n_t(j)^{1+\phi}}{1+\phi} + \varpi_t k_t(j) + \lambda_t(j)(i_t(j)k_t(j) + W_t n_t(j) - P_t c_t(j)) + dT_t/dt$$

FOCs:

$$
\begin{aligned}
&\frac{d\lambda_t(j)}{dt} = \rho\lambda_t(j) - \lambda_t(j)i_t(j) - \lambda_t(j)\varpi_t \\
&\Rightarrow \frac{d\lambda_t(j)/dt}{\lambda_t(j)} = \rho - i_t(j) - \varpi_t, \\
&1/c_t(j) = \lambda_t(j)P_t \\
&\Rightarrow \frac{dc_t(j)/dt}{c_t(j)} = -\frac{d\lambda_r(j)/dt}{\lambda_t(j)} - \frac{dP_t/dt}{P_t} = i_t(j) + \varpi_t - \rho - \pi_t, && \text{(E.2)} \\
&n_t(j)^{\phi} = \lambda_t(j)W_t, \\
&\Rightarrow c_t(j)n_t(j)^{\phi} = W_t/P_t. && \text{(E.3)}
\end{aligned}
$$

where $\pi_t = \frac{dP_t/dt}{P_t}$ is the inflation rate.

E.3 Habitat Investors

Private habitat investors demand the following aggregate amount of bonds:

$$H_t(\tau) = -\alpha(\tau)logP_t(\tau) - \theta(\tau)\beta_t.$$

E.4 Arbitrageurs

Arbitrageurs pool households savings and manage the fund through trading on the bond cash and repo markets. Arbitrageurs are risk-averse, they take as given the household savings and make decisions about optimal portfolio. The trading profits are used to payback the promised interests to households. Then the net gains/losses are equally transferred to households.

Hedge fund's problem:

$$
\begin{aligned}
&\max_{Q_t^h(\tau)} E_t\left[dW_t^h\right] - \frac{1}{2\rho_h}\operatorname{Var}_t\left[dW_t^h\right], \\
&s.t. \\
&dW_t^h - W_t^h r_t dt = \int_0^T \underbrace{Q_t^h(\tau)}_{\text{repo demand}} \underbrace{(\frac{dP_t(\tau)}{P_t(\tau)} - R_t(\tau)dt)}_{\text{trading profit}} d\tau. && \text{(E.4)}
\end{aligned}
$$

Dealer's problem:

$$\begin{aligned}
&\max_{X_t(\tau),Q_t^d(\tau)} E_t\left[dW_t^d\right] - \frac{1}{2\rho_d}\operatorname{Var}_t\left[dW_t^d\right],\\
&s.t.\\
&dW_t^d - W_t^d r_t dt = \int_0^T \underbrace{X_t(\tau)}_{\text{direct holdings}} \underbrace{(\frac{dP_t(\tau)}{P_t(\tau)} - r_t dt}_{\text{excess return}} - \underbrace{\Lambda_t(\tau)dt}_{\text{B/S cost}})d\tau\\
&\qquad + \int_0^T \underbrace{Q_t^d(\tau)}_{\text{repo supply}} \underbrace{(R_t(\tau) - r_t}_{\text{repo wedge}} - \underbrace{\Lambda_t(\tau)}_{\text{B/S cost}})dtd\tau. \qquad \text{(E.5)}
\end{aligned}$$

In equilibrium, the cost of direct and repo holdings must be equivalent. Thus, $R_t(\tau) = r_t + \Lambda_t(\tau)$. Dealer and hedge fund face the same problem except for risk-bearing capacity. The optimal portfolio is

$$X_t(\tau) = -\frac{\rho_d}{\rho_d + \rho_h} H_t(\tau), \qquad \text{(E.6)}$$

$$Q_t(\tau) = -\frac{\rho_h}{\rho_d + \rho_h} H_t(\tau). \qquad \text{(E.7)}$$

Embedding equations (E.6) and (E.7) into equations (E.4) and (E.5), I get arbitrageurs' aggregate bedget constraint:

$$d\mathcal{W}_t - \mathcal{W}_t r_t dt = -\int_0^T H_t(\tau)(\frac{dP_t(\tau)}{P_t(\tau)} - R_t(\tau)dt)d\tau,$$

which indicates that the distribution of funds between dealer and hedge fund does not affect the aggregate trading profits. The arbitrageurs make promised interest payment $\int_0^J h(j)i_t(j)k_t(j)dj$ from its wealth growth $d\mathcal{K}_t$ and then transfer the remaining part equally to households.

E.5 Final Goods Producer

A competitive final goods producer aggregates a continuum of intermediate inputs denoted by $k \in [0,1]$ to produce final goods.

$$Y_t = \left(\int_0^1 y_t(k)^{\frac{\varepsilon-1}{\varepsilon}} dk\right)^{\frac{\varepsilon}{\varepsilon-1}}.$$

Taking the final goods Y_t as given, the producer chooses ingredients to minimize cost

$$\min_{y_t(k)} \int_0^1 y_t(k)p_t(k)dk.$$

FOC:

$$y_t(k) = \left(\frac{p_t(k)}{P_t}\right)^{-\varepsilon} Y_t,$$

where

$$P_t = \left(\int_0^1 p_t(k)^{1-\varepsilon}dk\right)^{\frac{1}{1-\varepsilon}}$$

is the aggregate price index which forms the inflation rate.

E.6 Intermediate Goods Producers

A continuum of monopolistically competitive intermediate goods producers $k \in [0,1]$. Assuming quadratic price adjustment cost, the period profit for producer k is

$$p_t(k)y_t(k) - W_t n_t(k) - \frac{\theta}{2}\left(\frac{dp_t(k)/dt}{p_t(k)}\right)^2 P_t Y_t.$$

Assuming linear productivity of labor $y_t(k) = A_t n_t(k)$, the intermediate goods producers' problem becomes

$$\min_{p_t(k)} \int_0^\infty Q_t \left\{ p_t(k)\left(\frac{p_t(k)}{P_t}\right)^{-\varepsilon} Y_t - \frac{W_t}{A_t}\left(\frac{p_t(k)}{P_t}\right)^{-\varepsilon} Y_t - \frac{\theta}{2}\left(\frac{dp_t(k)/dt}{p_t(k)}\right)^2 P_t Y_t \right\} dt, \tag{E.8}$$

where Q_t is the aggregate stochastic discount factor satisfying

$$\begin{aligned}
Q_t &= \Big[exp(\int_0^t -\rho ds)\Big]\Big[exp\Big(\int_0^t \big(\int_1^H \frac{d\lambda_s(j)/ds}{\lambda_s(j)} h(j)dj\big)ds\Big)\Big] \\
&= \Big[exp(\int_0^t -\rho ds)\Big]\Big[exp\Big(\int_0^t \big(\int_1^H (\rho - i_s(j) - \varpi_s)h(j)dj\big)ds\Big)\Big] \\
&= \Big[exp(\int_0^t -\rho ds)\Big]\Big[exp\Big(\int_0^t \big(\rho - \varpi_s - \int_1^H i_s(j)h(j)dj\big)ds\Big)\Big] \\
&= exp\Big(\int_1^H i_t(j)h(j)dj + \varpi_t\Big) \\
&= exp\Big(\tilde{i}_t + \varpi_t\Big)
\end{aligned}$$

Current-value Hamiltonian:

$$\begin{aligned}
&H(dp_t(k)/dt, p_t(k), \lambda_t(k)) \\
=&p_t(k)\left(\frac{p_t(k)}{P_t}\right)^{-\varepsilon} Y_t - \frac{W_t}{A_t}\left(\frac{p_t(k)}{P_t}\right)^{-\varepsilon} Y_t - \frac{\theta}{2}\left(\frac{dp_t(k)/dt}{p_t(k)}\right)^2 P_t Y_t + \lambda_t(k)\left(\frac{dp_t(k)/dt}{p_t(k)}\right)
\end{aligned}$$

FOCs:

$$\lambda_t(k) = \theta \frac{dp_t(k)/dt}{p_t(k)}\frac{P_t}{p_t(k)}Y_t,$$

$$d\lambda_t(k)/dt = (\tilde{i}_t + \varpi_t)\lambda_t(k) - \left[(1-\epsilon)(\frac{p_t(k)}{P_t})^{-\epsilon}Y_t + \epsilon\frac{W_t}{p_t(k)}(\frac{p_t(k)}{P_t})^{-\epsilon}\frac{1}{A_t}Y_t + \theta\left(\frac{dp_t(k)/dt}{p_t(k)}\right)^2\frac{P_t}{p_t(k)}Y\right].$$

Focusing on symmetric equilibrium $p_t(k) = p_t = P_t$, the FOCs become

$$\begin{aligned}
&\lambda_t(k) = \theta\pi_t Y_t \\
\Rightarrow& d\lambda_t(k)/dt = \theta\pi_t dY_t/dt + \theta Y_t d\pi_t/dt, \\
&d\lambda_t(k)/dt = (\tilde{i}_t + \varpi_t)\lambda_t(k) - \Big[(1-\epsilon)Y_t + \epsilon\frac{W_t}{P_t}\frac{1}{A_t}Y_t + \theta\pi_t^2 Y\Big] \\
\Rightarrow& \pi_t\frac{dY_t/dt}{Y_t} + d\pi_t/dt = (\tilde{i}_t + \varpi_t)\pi_t - \Big[\frac{1-\epsilon}{\theta} + \frac{\epsilon}{\theta}\frac{W_t}{P_t}\frac{1}{A_t} + \pi_t^2\Big] \\
\Rightarrow& \pi_t(\tilde{i}_t + \varpi_t - \pi_t - \frac{dY_t/dt}{Y_t}) = \frac{\epsilon - 1}{\theta}\Big(\frac{\epsilon}{\epsilon-1}\frac{W_t}{P_t}\frac{1}{A} - 1\Big) + d\pi_t/dt. \qquad \text{(E.9)}
\end{aligned}$$

E.7 Aggregation

For tractability, I assume a "head of HH" sets transfers such that in equilibrium, wealth is equalized across HHs: $k_t(\tau) = K_t$. Aggregating the individual budget constraint (E.1), I have

$$dK_t = (\tilde{i}_t K_t + W_t N_t - P_t C_t)dt + dT_t,$$

where $\tilde{i}_t = \int_0^1 i_t(j)h(j)dj$. Therefore, the aggregation can be represented by a representative HH who borrows at the rate $\tilde{i}_t$. The representative HH's Euler equation, from equation (E.2), is

$$\frac{dC_t/dt}{C_t} = \tilde{i}_t + \varpi_t - \rho - \pi_t. \tag{E.10}$$

The representative HH's labor-consumption tradeoff is

$$C_t N_t^{\phi} = W_t/P_t. \tag{E.11}$$

Final goods market clearing condition is

$$C_t = Y_t = A_t N_t. \tag{E.12}$$

In the following analysis, I derive the natural economy. The standard NK model assumes the economy is in its natural state when the price can adjust flexibly. In this model, besides the price stickiness, there are three other sources of frictions: the risk of conducting carry trades, the non-pecuniary cost of holding assets, and the quality value of saving. I define the baseline economy as a special case with $\rho = 0$, $\lambda = 0$, $\kappa_r = 0$, and $\theta = 0$. In additional, I remove the quality value by letting $\varpi_t = 0$. In this scenario, there exists flexible price adjustment. The bond yields and repo rates are both the same as the short rate, thus the aggregate nominal rate elapses to short rate.

The flexible price can be derived from minimizing the profit without price adjustment cost, i.e., solving equation (E.9) with $\theta = 0$. In the following analysis, I use the subscript n to denote quantities and prices in the baseline natural economy.

$$\min_{p_t^n(k)} \{p_t^n(k)y_t^n(k) - W_t^n n_t^n(k)\} = \min_{p_t^n(k)} \left\{ p_t^n(k) \left(\frac{p_t^n(k)}{P_t^n}\right)^{-\varepsilon} Y_t^n - W_t^n \left(\frac{p_t^n(k)}{P_t^n}\right)^{-\varepsilon} Y_t^n/A_t \right\}$$
$$\Rightarrow p_t^n(j) = P_t^n = \frac{\epsilon}{\epsilon - 1} W_t^n/A_t.$$

The baseline version of labor-consumption tradeoff (E.11) is $C_t^n (N_t^n)^\phi = W_t^n/P_t^n$. Recall that I assume all profits and costs are transferred to households so there is no real resource loss. The baseline counterparty of the market clearing condition (E.12) is $C_t^n = Y_t^n = A_t N_t^n$. Combining these two equations, I have

$$Y_t^n = C_t^n = A_t \left(\frac{\epsilon - 1}{\epsilon}\right)^{\frac{1}{1+\phi}}. \tag{E.13}$$

The baseline aggregate Euler equation is

$$\frac{dC_t^n/dt}{C_t^n} = \tilde{i}_t^{\,n} - \rho - \pi_t^n. \tag{E.14}$$

From equation (E.13), $\frac{dC_t^n/dt}{C_t^n} = \frac{dA_t/dt}{A_t}$. I assume a constant technology growth, then the natural real interest rate $\tilde{i}_t^{\,n} - \pi_t^n$ is constant. Subtracting the baseline aggregate Euler equation (E.14) from the full aggregate Euler equation (E.10), I have

$$\frac{dY_t/dt}{Y_t} - \frac{dY_t^n/dt}{Y_t^n} = \tilde{i}_t + \varpi_t - \pi_t - (\tilde{i}_t^{\,n} - \pi_t^n)$$
$$\Rightarrow \frac{dX_t/dt}{X_t} = \tilde{i}_t + \varpi_t - \pi_t - r^n,$$

where $X_t = Y_t/Y_t^n$ is the output gap and $r^n = \tilde{i}_t^{\,n} - \pi_t^n$ is the natural real interest rate. Now I will derive the NK Phillips curve. Substitute the Euler equation (E.10) into the inflation dynamic equation (E.9), I have

$$\pi_t \rho = \frac{\epsilon - 1}{\theta} \left(\frac{\epsilon}{\epsilon - 1} \frac{W_t}{P_t} \frac{1}{A_t} - 1 \right) + d\pi_t/dt,$$

where $1/A_t = \frac{\epsilon-1}{\epsilon} P_t^n/W_t^n$ from the flexible price expression. Therefore, the inflation dynamic can be written as

$$\pi_t \rho = \frac{\epsilon - 1}{\theta} \left(\frac{W_t}{P_t} / \frac{W_t^n}{P_t^n} - 1 \right) + d\pi_t/dt,$$

where $\frac{W_t}{P_t} = Y_t^{1+\phi}/A_t^{\phi}$ and $\frac{W_t^n}{P_t^n} = (Y_t^n)^{1+\phi}/A_t^{\phi}$ from the labor-consumption tradeoff and market clearing conditions. The inflation dynamic can be further written as

$$\begin{aligned} \pi_t \rho &= \frac{\epsilon - 1}{\theta} \left(Y_t^{1+\phi}/(Y_t^n)^{1+\phi} - 1 \right) + d\pi_t/dt \\ &= \frac{\epsilon - 1}{\theta} \left(X_t^{1+\phi} - 1 \right) + d\pi_t/dt. \end{aligned}$$

Defining $x_t = logX_t$, I have $X_t^{1+\phi} - 1 = e^(1 + \phi)x_t - 1 \approx (1 + \phi)x_t$. The IS curve and Phillips curve are given by

$$\begin{aligned} dx_t/dt &= \tilde{i}_t + \varpi_t - \pi_t - r^n, \\ d\pi_t/dt &= \rho \pi_t - \frac{\epsilon - 1}{\theta}(1 + \phi)x_t, \end{aligned}$$

the same as equations (1) and (4) in the main text. The model then is closed with the Taylor rule and the demand factor process as in the main text.