A New Keynesian Preferred Habitat Model with Repo

Qian Wu*

August 1, 2024

Abstract

This paper documents different observations in the Treasury cash and repo markets during the Global Financial Crisis (GFC) and the Covid-19 pandemic. To account for these observations, I develop a New Keynesian Preferred Habitat model with repo assets, featuring market segmentation, financial frictions, and liquidity/safety preference. I show that there was a flight-to-liquidity demand for short-term Treasuries during the GFC and a flight-from-safety supply for long-term Treasuries 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

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, providing central banks with a reference for analyzing fluctuations in macroeconomic variables and their relationships to monetary policies. 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. 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 et al. (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 non-pecuniary holding benefits/costs. There are four key departures from the benchmark New Keynesian model. First, the market segmentation introduced by habitat investors allows QE 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 utility from holding savings due to safety and liquidity 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 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 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 investors' attitudes towards Treasuries, along with shifts in financial intermediation conditions, together explain the discrepancies in key financial market variables. Specifically, there was a net demand shock for short-term Treasuries during the GFC, leading to a flight-to-liquidity shock that decreased short-term yields and steepened the term structure of the Treasury yield curve. Conversely, during Covid-19, there was selling pressure for long-term Treasuries, resulting in a flight-from-safety shock that increased long-term yields and similarly steepened the term structure. During both recessions, the intermediation conditions for financial institutions changed considerably. Updated capital regulations under Basel III eliminated the convenience value of holding Treasuries by increasing the balance sheet cost (Klingler and Sundaresan, 2023). This explains the different signs of the Treasury-

OIS spread, a measure of the inconvenience yield in the Treasury cash market. The repo market experienced similar shifts. During the GFC, there was a net convenience value in holding Treasuries. When demand soared, securities holders benefited from the convenience premium through lower financing costs in the repo market. However, during Covid-19, the convenience value shrank or even turned negative, causing soaring supply to force securities holders to accept higher financing costs in the repo market. On the macroeconomic side, when there is a flight-to-liquidity motivation, households invest in savings products not only for their pecuniary return but also for their safety premium. In this scenario, households do not increase consumption even with lower interest rates, leading to drops in output and inflation alongside declining interest rates.

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 high balance sheet expansion costs, 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 balance sheet costs 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 balance sheet costs, 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. In other words, conventional expansion is more effective in the repo regime. The impact of QE on boosting the output gap is similar in both regimes; however, in the repo regime, this boost is accompanied by smaller inflation, making QE more effective in this context. 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.

Why important

- By now QE is a norm along with short rate policy. But what do we know about the efficacy of these policies in the post-GFC era? From empirics, the two most recessions witness different financial market observations, which are hard to be accounted for by the current New Keynesian literature. What's more, recent institutional reforms such as the Supplementary Leverage Ratio (SLR) requirement and wholesale market indexing benchmark reform ask for an updated revision of the standard New Keynesian framework to better understand the monetary policies perform in the new era.

- The important ongoing benchmark interest rate transition from LIBOR to SOFR 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. Unfortunately, the current literature does not have a great answer to this issue.

- Finally, by introducing new features into the New Keynesian Preferred Habitat framework, I also contribute to the literature from a methodological perspective. Incorporating repo assets introduces complexity to the general equilibrium, and I propose a generic numerical method to solve the model.

Related Literature

- **Preferred habitat models.** Vayanos & Vila (2021); Ray (2019); He et al. (2022).
- **Empirical evidence: Treasury cash and repo market frictions.**
  - Imperfect competition: Huber (2023); Eisenschmidt, Ma, and Zhang (2023)
  - Collateral scarcity: D'Amico, Fan, and Kitsul (2018); Arrata et al. (2020); Corradin and Maddaloni (2020)
  - Balance sheet cost: Duffie and Krishnamurthy (2016); Infante et al. (2023)
- **Benchmark rate reform.** Jermann (2019); Duffie et al. (2023)

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[2]. The US Treasury market features a deep and broad secondary cash market, with daily trading volumes exceeding $884 billion as of June 2024 [3]. 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 et al., 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 [4]. Additionally, broker-dealers are key participants in the Treasury secondary cash market. Figure 1 summarizes the main components of the Treasury secondary

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

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

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

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.

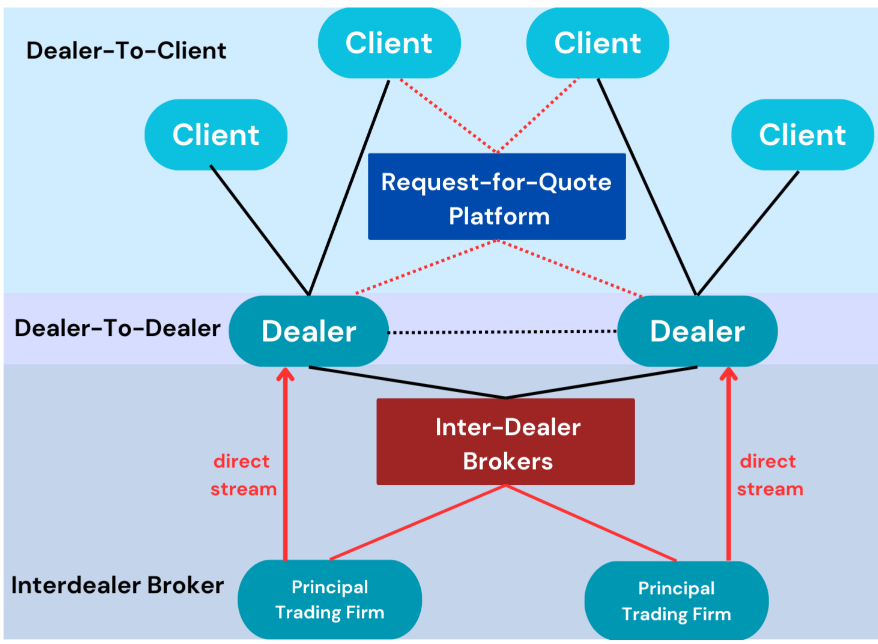


Figure 1: Structure of the Treasury secondary cash market.

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.

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 et al., 2011; Krishnamurthy and Vissing-Jorgensen, 2012; Nagel, 2016; Adrian et al., 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 example, when investors face uninsurable liquidation shocks, which can increase the cost of liquidating assets such as equities and houses, they tend to demand more liquid assets to hedge against these shocks. Since liquidity and safety increase the non-pecuniary benefit of holding Treasuries, I refer to this value as the convenience yield of Treasuries.

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 [5]. I refer to the non-pecuniary cost of holding Treasuries due to the SLR as the inconvenience yield of Treasuries.

What would be the sign of the net convenience yield in the Treasury cash market? The current literature does not provide a clear answer to this question. Given that the SLR was implemented in 2018, a reasonable hypothesis is that the net convenience yield in the Treasury cash market was mostly positive prior to 2018 and likely shrank from 2018 onward. In the next section, I briefly discuss the structure of the Treasury repo market and the convenience/inconvenience yield in that market.

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 ex-

[5] 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.

ample, 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 [6] [7] .

The U.S. repo 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 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 et al., 2021). Broker-dealers are the main securities borrowers in the triparty repo market, sourcing cash from conservative cash investors such as MMFs.

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

[7] 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.

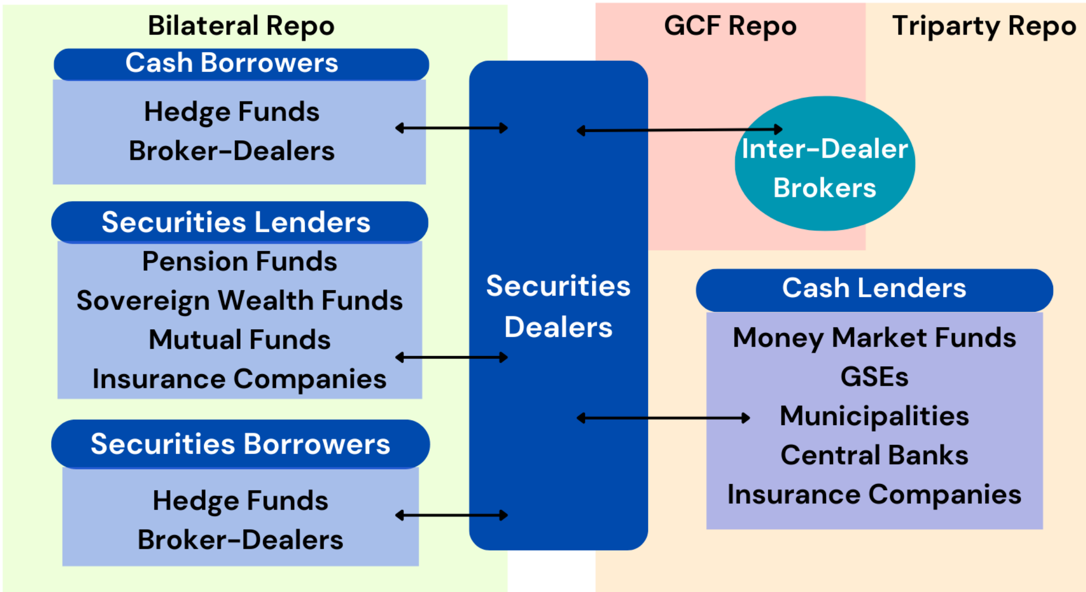


Figure 2: Structure of the Treasury repo market.

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.

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 et al., 2015; He et al., 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 et al., 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 et al., 2018). In this context, holding Treasuries provides investors with an additional non-pecuniary benefit, which I term the convenience yields of Treasury repos.

The SLR affects broker-dealers' intermediation 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. As this balance sheet cost is a non-pecuniary cost associated with holding Treasuries, I refer to it as the inconvenience yield of holding Treasury repos.

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, 2004; Goldreich et al., 2005; Beber et al., 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 1 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 with holding Treasuries [8]. 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.

[8] 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.

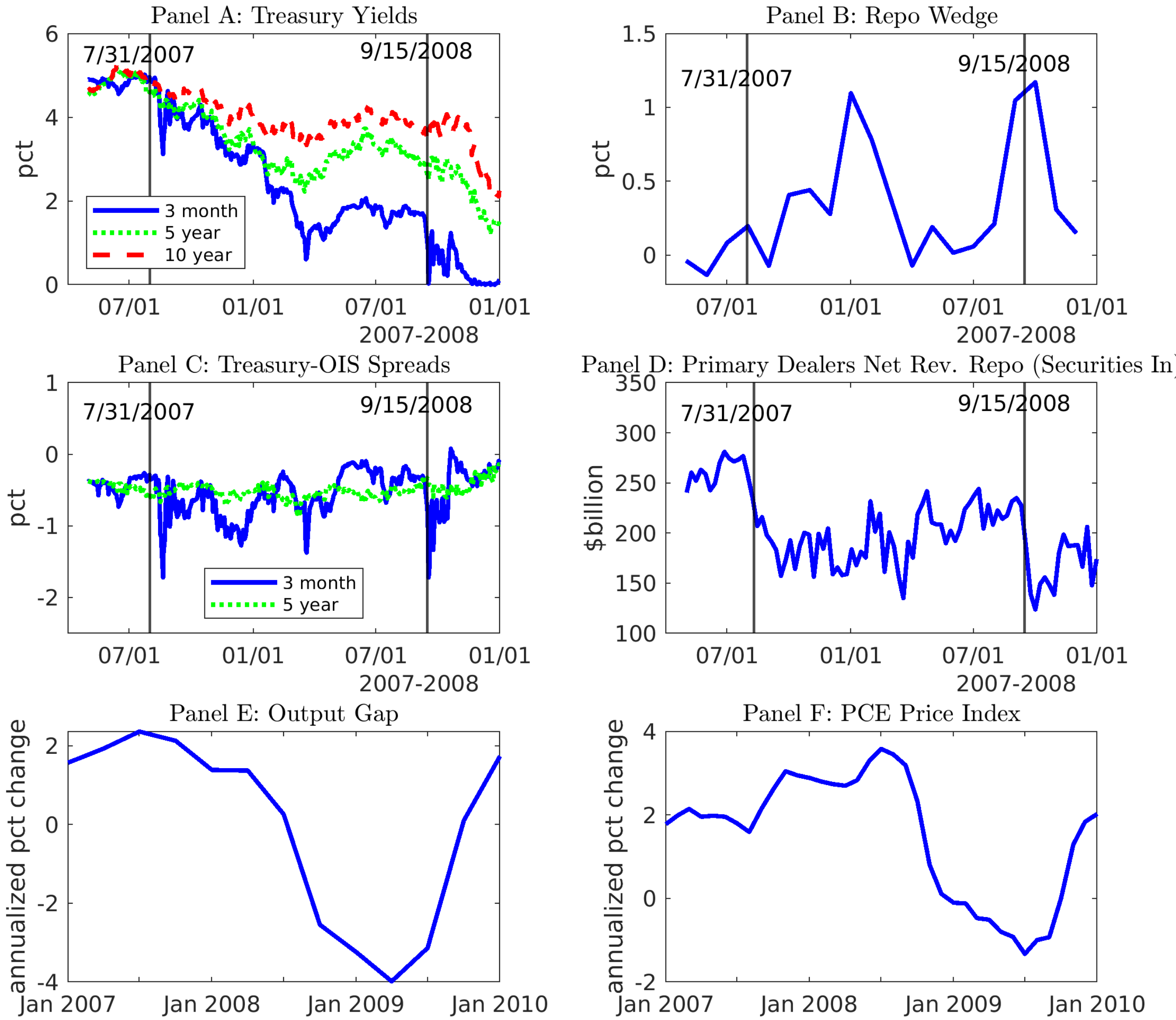


Figure 3: What happened during the early stage of GFC?

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 [9]. 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 [10]. Panel D shows that primary dealers reduced their net repo Treasury holdings by more than $50 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 [11]. Figure 4 depicts key financial and macroeconomic variables from this period [12]. Between March 9 and March 23, the 3-month Treasury yield rose by more than half

[9]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.

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

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

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

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.

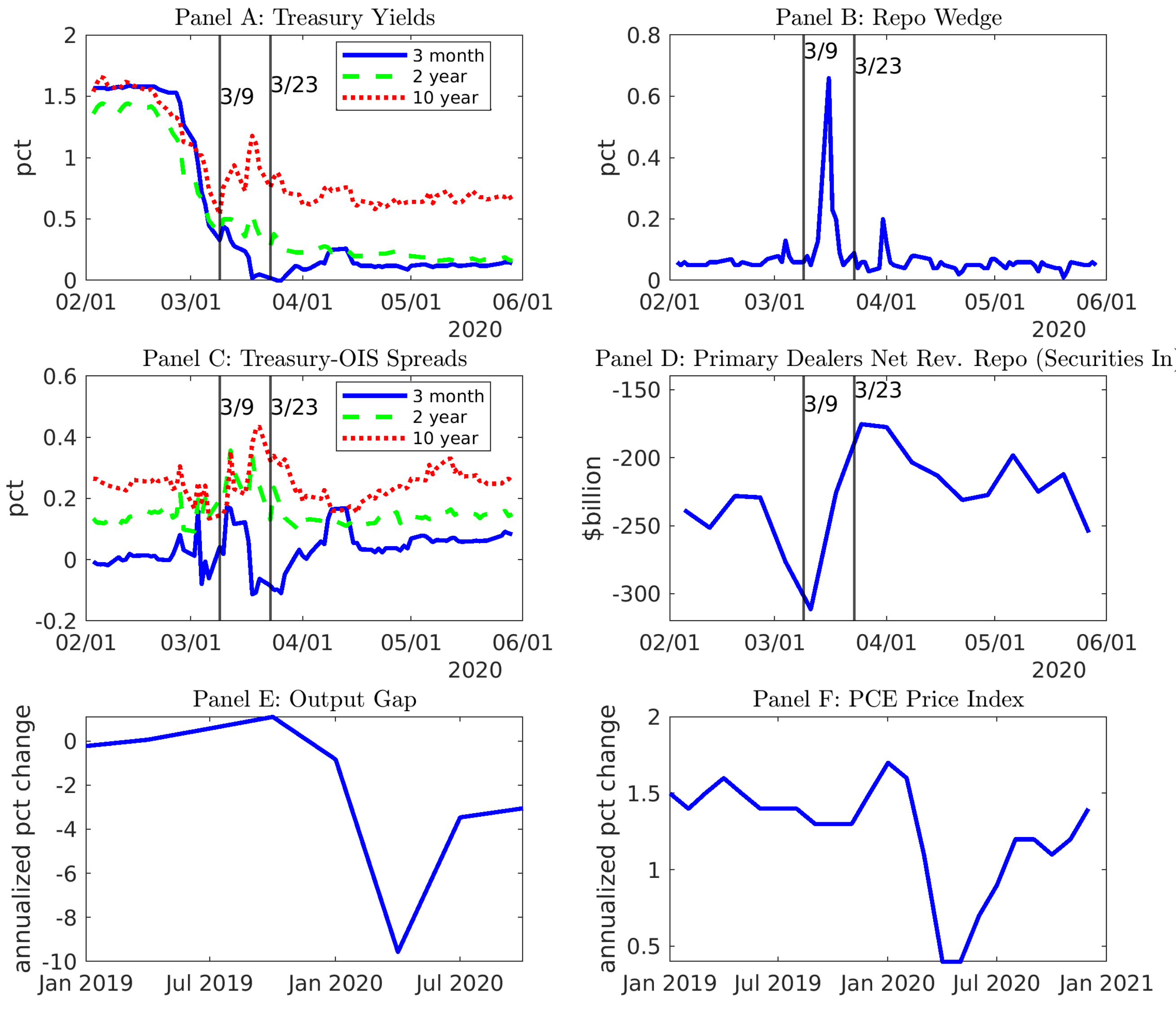


Figure 4: What happened during Covid-19?

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 [13]. Moreover, the rise in the 10-year Treasury-OIS spread during this time signals 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.

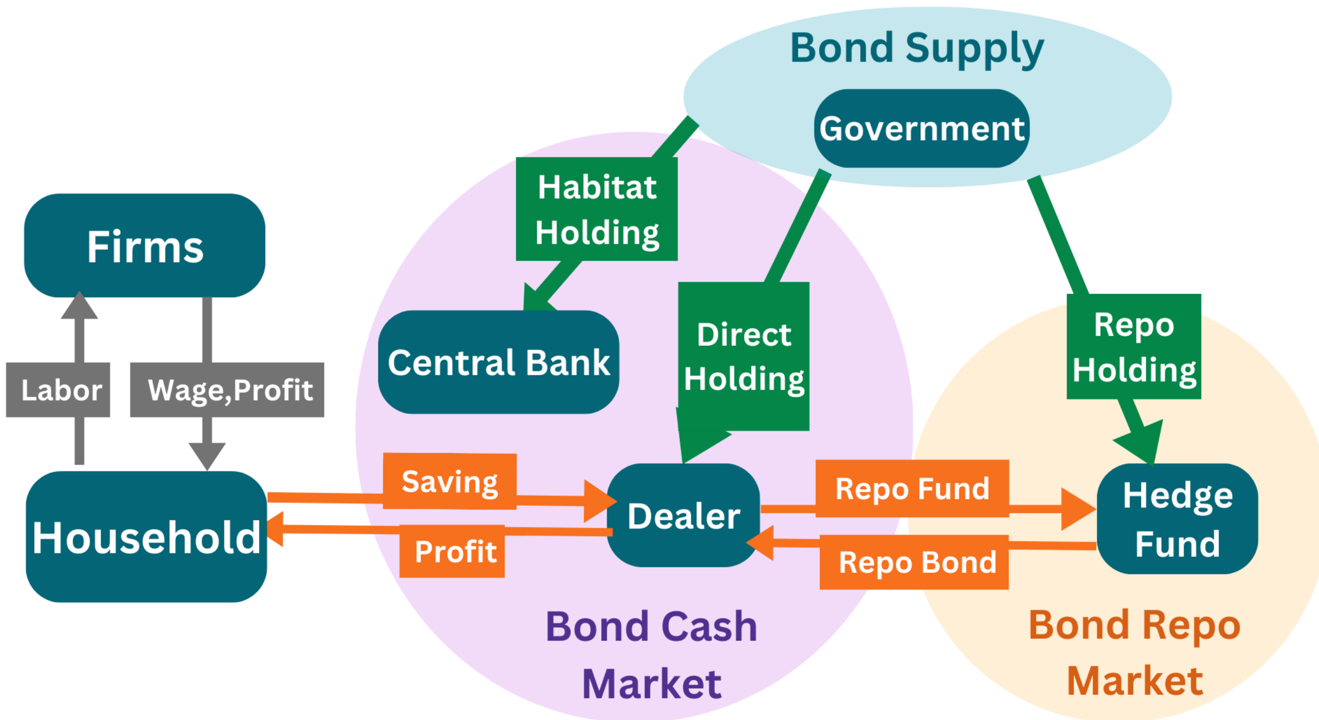
The Treasury repo market also experienced significant fluctuations during this period. To measure the (in)convenience yield in the Treasury repo market, I follow He et al. (2022) and use the GCF-triparty repo spread. The triparty repo rate is used as a benchmark because it represents a collateralized 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.

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

3 The Model



Agents:

- Government: supplys government bonds aggregated to 0.
- Central bank:
  - demands bonds as a habitat investor;
  - sets policy rate using Taylor Rule.
- Financial intermediaries:
  - 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;

- provide savings to intermediaries 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.

3.1 Macroeconomic Dynamics

The macroeconomic side of this model is in the same spirit as Werning (2011). The macro-dynamics in equilibrium are characterized by a three-equation New Keynesian system in continuous time. First, the IS curve is given by

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

where x_t is the output gap, the log difference between actual output and potential output that would prevail in the flexible price setup. π_t is the inflation rate, and $\bar{r}$ is the natural borrowing rate corresponding to zero output gap when the price is perfectly flexible and there's no financial frictions. ς^{-1} is the elasticity of intertemporal substitution. [14] I follow Ray (2019) to assume that in this economy, the nominal borrowing cost is something departing from the policy rate. The overall borrowing cost in this economy is controlled 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 relative importance of bond yields and repo rates in determining the overall borrowing cost is governed by the weight functions $\eta^i(\tau)$ and $\eta^R(\tau)$.

The Phillips curve in this economy is given by

$$\mathrm{d}\pi_t = \left(\chi\pi_t - \delta x_t\right)\mathrm{d}t, \tag{3}$$

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

with χ controlling the discount rate and δ the stickiness of price. A smaller δ implies a more sticky price level. When $\delta \to \infty$, the economy has perfectly flexible price. This forward-looking Phillips curve states that the inflation is proportional to the present value of future output gaps.

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

$$\mathrm{d}r_t = -\psi_r \left(r_t - \phi_\pi \pi_t - \phi_x x_t - r^*\right) \mathrm{d}t + \sigma_r \, \mathrm{d}B_{r,t}, \tag{4}$$

where ψ_r controls the mean-reversion rate of the policy, ϕ_π and ϕ_x give the weights for inflation and output targets, r^* is the target rate that ensures 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 equation (1)-(4).

3.2 Term Structures Determination

The term structures for bonds and repo assets are determined according to a Preferred Habitat model embedded with repo assets. There are two risk factors in this economy: the short rate and the demand shifter. I assume the following processes for these two risk factors:

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

where r_t is the short rate and β_t is the demand shifter. Collecting the two risk factors into a vector s_t, I can rewrite the two risk factors in the vector format:

$$ds_t = -\Gamma(s_t - s^{ss})dt + \Sigma dB_t. \tag{5}$$

Equation (5) generates uncertainties in this model.

3.2.1 Habitat Agents

The aggregate bond demand from habitat investors is:

$$H_t(\tau) = -\alpha(\tau) log P_t(\tau) - \theta(\tau)\beta_t. \tag{6}$$

Habitat investors with maturity preference τ hold only bonds with maturity τ and no other bonds. Their aggregate demand is affected by bond prices $P_t(\tau)$ as well as a demand shifter β_t. Note that the one-dimensional demand shock can achieve heterogeneous effect on habitat investors by the location function $\theta(\tau)$. I assume that $\alpha(\tau) \geq 0$ for all τ so that the habitat investors decrease their demand if bond prices rise up. The demand shifterβ_t can achieve differentiated effects on maturities through the loading function $\theta(\tau)$. I do not assume a sign for $\theta(\tau)$, allowing it to take different signs across maturities.

3.2.2 Arbitrageurs

I follow He et al. (2022) to divide the arbitrageurs into hedge funds and dealers. Hedge funds borrow from dealers through the repo market to finance the purchase of bonds. [15] I assume there is a representative hedge fund in this sector whose optimization problem is given by:

$$\begin{aligned}
&\max_{Q_t^h(\tau)} E_t\left[dW_t^h\right] - \frac{1}{2\rho_h}\,\mathrm{Var}_t\left[dW_t^h\right], \\
&s.t. \\
&dW_t^h - W_t^h r_t dt = \int_0^T \underbrace{Q_t^h(\tau)}_{\text{repo demand}} \underbrace{(\frac{dP_t(\tau)}{P_t(\tau)} - R_t(\tau)dt)}_{\text{trading profit}} d\tau.
\end{aligned} \tag{7}$$

Hedge fund's trading profit of buying bonds is equal to the difference between bond's instantaneous return and the repo rate, since bonds are financed solely through the repo borrowing. The parameter ρ_h controls the risk bearing capacity of hedge funds.

On the other side of the repo market, dealers provide secured funds to hedge funds, besides directly holding the remaining bonds on the market. Both of the direct holdings and indirect holdings through repo are financed in the risk-free market. Likewise, I assume there

[15] "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 et al. (2021) for hedge funds' repo holdings.

is a representative dealer whose optimization problem is given by:

$$\max_{X_t(\tau),Q_t^d(\tau)} E_t\left[dW_t^d\right] - \frac{1}{2\rho_d}\operatorname{Var}_t\left[dW_t^d\right],$$

$$s.t.$$

$$dW_t^d - W_t^d r_t d_t = \int_0^T \underbrace{X_t(\tau)}_{\text{direct holdings}} \; (\underbrace{\frac{dP_t(\tau)}{P_t(\tau)} - r_t dt}_{\text{excess return}} - \underbrace{\Lambda_t(\tau)dt}_{\text{B/S cost}})d\tau$$

$$+ \int_0^T \underbrace{Q_t^d(\tau)}_{\text{repo supply}} \; (\underbrace{R_t(\tau) - r_t}_{\text{repo wedge}} - \underbrace{\Lambda_t(\tau)}_{\text{B/S cost}})dtd\tau. \tag{8}$$

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 structure is consistent with what we observed during the pandemic, when different kinds of investors adjusted their holdings of government bonds, the variation was mainly absorbed by broker-dealers. A small portion was accommodated through direct holdings and a larger portion through repo financing. [16]

A balance sheet cost is assumed for both direct and repo holdings of bonds to reflect the Supplementary Leverage Ratio (SLR) requirement. The SLR requires bank-holding-companies to have capital equal to or greater than 5% of their total assets, regardless of the risk composition of the assets. Because SLR does not distinguish between risky assets such as unsecured lendings from safe assets such as Treasury securities and repos backed by Treasuries, it in fact imposes extra holding cost for safe assets. This is because risky assets usually earn more interest than safe assets, increasing the opportunity cost of holding such safe assets. For simplicity, I assume that the marginal balance sheet cost is proportional to the balance sheet size:

$$\Lambda_t(\tau) = \lambda B_t(\tau), \text{ where} \tag{9}$$

$$B_t(\tau) = X_t(\tau) + Q_t^d(\tau). \tag{10}$$

The structure of $\Lambda_t(\tau)$ implies that when the representative dealer sizes up the balance sheet, the marginal balance sheet cost also increases proportionally. Note that in this paper,

[16] "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 et al. (2022) for details.

although the repo market friction is explicitly modeled as a balance sheet cost, one can alternatively interpret it as a scarcity premium in the following sense. When there is an exogenous demand shock reducing the available amount of bonds for investors (arbitrageurs), this is characterized by a decrease in $B_t(\tau)$. The bond holders now enjoy a lower cost of borrowing cash on the repo market as the safe assets become more scarce. This results in a lower repo rate, which is consistent with equation (8) whenever the λ is positive.

3.2.3 Financial Markets Equilibrium

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 and Repo market clear for all maturities, i.e.,*

$$Q_t^h(\tau) = Q_t^d(\tau) \equiv Q_t(\tau), \tag{11}$$

$$H_t(\tau) + X_t(\tau) + Q_t(\tau) = 0. \tag{12}$$

It has been shown in Vayanos and Vila (2021) that the equilibrium leads to an affine structure bond price:

$$P_t(\tau) = exp\Big[-\Big(A(\tau)'s_t + C(\tau)\Big)\Big], \tag{13}$$

where $A(\tau) = [A_r(\tau)\ A_\beta(\tau)]'$ is a two-dimensional vector and $C(\tau)$ is a scalar. Using Ito's Lemma, I can derive the instantaneous return as

$$\frac{dP_t(\tau)}{P_t(\tau)} = \mu_t(\tau)dt - A(\tau)'\Sigma dB_t, \tag{14}$$

where $\mu_t(\tau)$ is the instantaneous expected return and is given by:

$$\mu_t(\tau) \equiv A'(\tau)'s_t + C'(\tau) + A(\tau)'\Gamma\left(s_t - r^{ss}\epsilon\right) + \frac{1}{2}A(\tau)'\Sigma\Sigma'A(\tau). \tag{15}$$

Replacing the instantaneous return in the representative hedge fund's problem, its FOC can be derived as

$$\mu_t(\tau) = \underbrace{R_t(\tau)}_{\text{financing cost}} + \underbrace{\frac{1}{\rho_h} A(\tau)'\Sigma\Sigma' \left[\int_0^T Q_t^h(\tau) A(\tau) d\tau \right]}_{\text{risk premium}}. \tag{16}$$

The left-hand side of equation (13) gives the marginal benefit of shifting one unit of wealth from the short rate r_t to the 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 from the fact that all hedge funds bond holdings must be financed through repo market. Portfolio adjustment exposes hedge funds to the risk as bond prices are sensitive to the risk factors. This exposure will be transmitted into additional cost of bond holding when hedge funds are risk-averse.

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

$$\mu_t(\tau) = \overbrace{r_t}^{\text{financing cost}} + \underbrace{\Lambda_t(\tau)}_{\text{B/S cost}} + \underbrace{\frac{1}{\rho_d} A(\tau)'\Sigma\Sigma' \left[\int_0^T X_t(\tau) A(\tau) d\tau \right]}_{\text{risk premium}}. \tag{17}$$

The left-hand side of equation (14) represents the marginal benefit for the representative dealer to increase bond direct holdings. The right-hand side represents the marginal cost. Note that direct holdings are subjected to balance sheet costs, hence $\Lambda_t(\tau)$ enters the equation. The intuition behind the risk premium is similar as for the hedge fund except that dealer is with a different level of risk aversion.

Besides direct holdings, dealers also choose repo assets. Taking the FOC with respect to repo assets, I obtain $R_t(\tau) - r_t - \Lambda_t(\tau)$, the difference between the repo wedge and the marginal balance sheet cost. To ensure an inner solution, in equilibrium the two terms must be equivalent. Intuitively, $R_t(\tau) - r_t$ measures dealer's indirect holding cost and $\Lambda_t(\tau)$ measures dealer's direct holding cost. If indirect holding proves to be more expensive, the dealer will opt for holding all bonds in the direct form. Conversely, if direct holding is the more costly option, the dealer will choose to hold all bonds in the repo form. Therefore, an inner solution establishes when

$$R_t(\tau) - r_t = \Lambda_t(\tau). \tag{18}$$

Imposing equation (15) on equation (14), 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 direct holdings and repo assets, equation (12) implies that

$$H_t(\tau) = -B_t(\tau). \tag{19}$$

The optimal risk sharing quantities are

$$X_t(\tau) = -\frac{\rho_d}{\rho_d + \rho_h} H_t(\tau), \tag{20}$$

$$Q_t(\tau) = -\frac{\rho_h}{\rho_d + \rho_h} H_t(\tau). \tag{21}$$

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.

Substituting the quantity in the dealer's FOC using equation (19), I have

$$\mu_t(\tau) - r_t = -\lambda H_t(\tau) - A(\tau)' \underbrace{\frac{1}{\rho_h + \rho_d} \Sigma\Sigma' \left[\int_0^T H_t(\tau) A(\tau) d\tau \right]}_{\text{risk price}}. \tag{22}$$

Equation (21) differs from the standard preferred-habitat literature as the portfolio adjustment induces balance sheet cost. In equilibrium, the expected excess return of holding one more unit of maturity τ net of the balance sheet cost is equal to the risk premium that compensates arbitrageurs for bearing the short rate risk and the demand risk. The portfolio risk 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. Besides, by adjusting the relative positions across maturities, the after-adjustment portfolio may be with a larger fraction of specific maturities more affected by the two risk factors, and this is the internal margin of the marginal cost. Note that when the factor sensitivity $A(\tau)$ is constant across maturities, the internal margin closes and portfolio adjustment only exposes investors to higher risk through the external margin.

This model features an arbitrage-free equilibrium, as can bee seen from equation (21). The net expected excess return 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 must be the same across all assets (maturities), otherwise there is possibility to construct arbitrage portfolios. The risk price involves equilibrium terms $H_t(\tau)$ and $A(\tau)$. In equation (21), 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 $C(\tau)$. Appendix B gives detailed derivations and numerical methods to solve for the equilibrium affine coefficients.

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

$$\begin{aligned} i_t(\tau) &= -\frac{1}{\tau} logP_t(\tau) \\ &= \frac{1}{\tau}\Big(A(\tau)' s_t + C(\tau)\Big), \end{aligned} \tag{23}$$

and the equilibrium repo rates are derived as

$$\begin{aligned} R_t(\tau) &= \Delta_t(\tau) + r_t \\ &= \Lambda_t(\tau) + r_t \\ &= \lambda\Big[\alpha(\tau) logP_t(\tau) + \theta(\tau)\beta_t\Big] + r_t \\ &= \lambda\Big[\theta(\tau)\beta_t - \alpha(\tau)\Big(A(\tau)' s_t + C(\tau)\Big)\Big] + r_t. \end{aligned} \tag{24}$$

3.3 General Equilibrium

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

1. *The financial markets (Treasry cash market, Treasury repo market) are in equilibrium.*
2. *The macroeconomy is characterized by the modified three-equation NK model.*

With the asset prices solved, I now revisit the macroeconomic dynamics. First, I can rewrite the effective nominal rate as

$$\begin{aligned} \tilde{r}_t &= \int_0^T \eta^i(\tau) i_t(\tau) \mathrm{d}\tau + \int_0^T \eta^R(\tau) R_t(\tau) \mathrm{d}\tau \\ &\equiv \hat{A}' s_t + \hat{C}, \end{aligned} \tag{25}$$

where

$$\hat{A}' \equiv \int_0^T \frac{1}{\tau}\eta^i(\tau)A(\tau)'d\tau + \int_0^T \eta^R(\tau)\Big[\lambda\Big(\theta(\tau)\gamma' - \alpha(\tau)A(\tau)'\Big) + \epsilon'\Big]d\tau, \tag{26}$$

$$\hat{C} \equiv \int_0^T \frac{1}{\tau}\eta^i(\tau)C(\tau)d\tau - \lambda\int_0^T \eta^R(\tau)\alpha(\tau)C(\tau)d\tau. \tag{27}$$

It follows that the IS curve in general equilibrium is given by

$$dx_t = \varsigma^{-1}(\hat{A}'s_t + \hat{C} - \pi_t - \bar{r})dt. \tag{28}$$

The other two macroeconomic equations, the Phillips curve and the Taylor rule, remain unchanged. Now the macroeconomic dynamics in the general equilibrium are fully characterized by equations (3), (4), and (29). There are 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 a four-equation system:

$$d\begin{bmatrix} r_t \\ \beta_t \\ x_t \\ \pi_t \end{bmatrix} = -\underbrace{\begin{bmatrix} \psi_r & 0 & -\psi_r\phi_x & -\psi_r\phi_\pi \\ 0 & \kappa_\beta & 0 & 0 \\ -\varsigma^{-1}\hat{A}_r & -\varsigma^{-1}\hat{A}_\beta & 0 & \varsigma^{-1} \\ 0 & 0 & \delta & -\chi \end{bmatrix}}_{\Upsilon}\left(\begin{bmatrix} r_t \\ \beta_t \\ x_t \\ \pi_t \end{bmatrix} - \begin{bmatrix} r^{ss} \\ \beta^{ss} \\ x^{ss} \\ \pi_{ss} \end{bmatrix}\right)dt + \underbrace{\begin{bmatrix} \sigma_r & 0 \\ 0 & \sigma_\beta \\ 0 & 0 \\ 0 & 0 \end{bmatrix}}_{\begin{bmatrix} S \\ 0 * I_2 \end{bmatrix}} d\underbrace{\begin{bmatrix} B_{r,t} \\ B_{\beta,t} \end{bmatrix}}_{B_t}. \tag{29}$$

The solution to this system are given by:

$$d\begin{bmatrix} r_t \\ \beta_t \end{bmatrix} = -\Gamma\left(\begin{bmatrix} r_t \\ \beta_t \end{bmatrix} - \begin{bmatrix} r^{ss} \\ \beta^{ss} \end{bmatrix}\right)dt + SdB_t, \tag{30}$$

$$\begin{bmatrix} x_t \\ \pi_t \end{bmatrix} = \Omega\left(\begin{bmatrix} r_t \\ \beta_t \end{bmatrix} - \begin{bmatrix} r^{ss} \\ \beta^{ss} \end{bmatrix}\right), \tag{31}$$

where Γ and Ω are derived from eigenvalues and eigenvectors of Υ. The derivation is included in appendix B.1.

The difficulty 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 that need to be solved from the asset pricing side. The general equilibrium $\hat{A}$ must ensure that

Γ in equation (31) coincides with that in equation (5). Appendix B.2 provides a description of the solution algorithm for the general equilibrium.

The general equilibrium can be described as a financial market-accommodated monetary policy rule. To understand the general equilibrium, let's think about intuitively the association between $\hat{A}$ and κ_r in the macroeconomic equilbrium and financial market equilbrium.

- From the macroeconomic equilibrium, this association is positive. When $\hat{A}$ increases, the profit transfer from financial intermediaries to households is larger. The marginal benefit of investing in financial intermediaries increases. Through the Euler equation, the monetary policy is more effective in boosting output. Therefore, the policy rate will revert back to its mean at a faster speed, κ_r is larger.

- However, from the financial market equilibrium, this association is negative. Financial intermediaries make profits by taking advantage of monetary shocks. When κ_r is larger, policy rate shock is very transitory, so the profitability is smaller. As a result, $\hat{A}$ is smaller.

- The general equilibrium is a pair of policy persistence and financial intermediary profitability that ensures both macroeconomic and financial market equilibrium. In other words, in the general equilibrium, the monetary policy rule successfully accommodates financial market dynamics so that it evolves consistently with investors' optimal portfolio decisions.

4 A Simple Case

In this section, I present a simple case of the model where an analytic solution to the affine coefficients is possible. I make the following assumptions to simplify the settings:

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 only has yield component and the weight is the same across the maturities, i.e.,* $\eta^i(\tau) = \eta^i = 1/T$ *and* $\eta^R(\tau) = 0$.

In this simple case, I can derive a closed-form solution to $A(\tau)$. A closed-form solution to $C(\tau)$ is possible if the loading function of the demand shifter is with specific formats. I abstract from the solution of $C(\tau)$ as it is not involved in the following analysis.

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

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

Without inflation, the linearized macroeconomic system reduces to:

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

where

$$\hat{A}_r = \frac{1}{T} \int_0^T \frac{1 - e^{-\kappa_r \tau}}{\kappa_r \tau} d\tau. \tag{32}$$

Note that equation (31) implies that $\hat{A}_r$ and κ_r are negatively correlated. In this simple model without financial frictions, the aggregate interest rate will equal to the short rate if any shift in the short rate is permanent ($\kappa_r = 0$). The monetary policy transmission is then perfect ($\hat{A}_r = 1$). As κ_r increases, the policy shock reverts back to its long-run average with a faster speed, therefore the long term yields underreact and the aggregate nominal interest rate also underreacts ($\hat{A}_r < 1$).

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

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

where

$$\begin{aligned} \frac{\kappa_r(\kappa_r - \psi_r)}{\psi_r \phi_x \varsigma^{-1}} &= \hat{A}_r, \\ \omega_x &= \frac{\psi_r - \kappa_r}{\psi_r \phi_x}. \end{aligned} \tag{33}$$

From equation (32), $\hat{A}_r$ and κ_r are positively related once κ_r exceeds $\psi_r/2$. This is because when the mapping is efficient, monetary policy boosts output very powerfully, therefore the policy will be there for a shorter period.

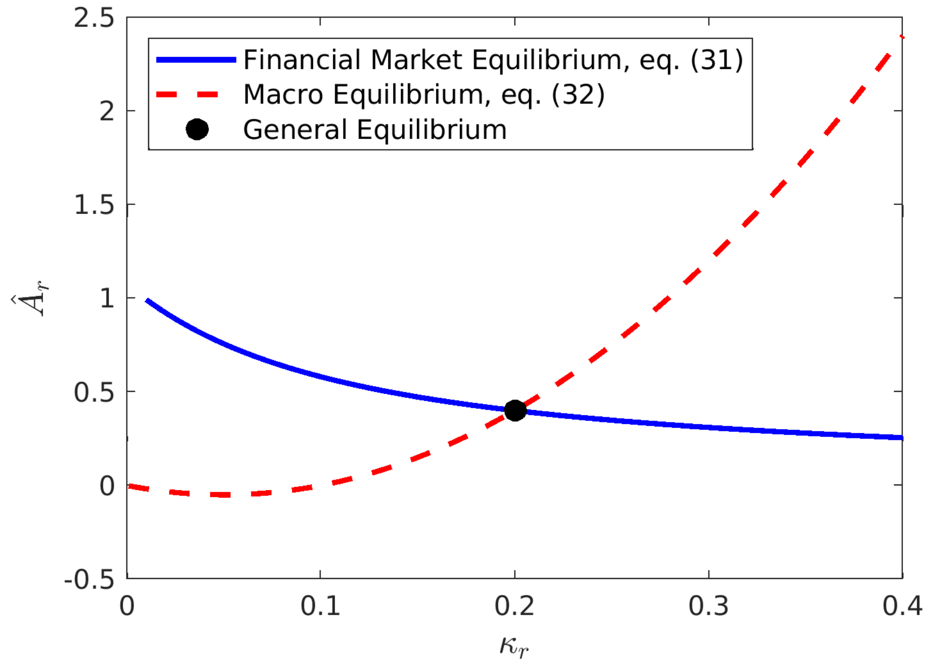


Figure 5: General Equilibrium in the Simple Case.

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

4.1 Extension 1: Balance Sheet Cost

Relaxation:

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

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

   In this extension, the solution to affine coefficient is

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

The financial market equilibrium can be described by the new version of equation (31)

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

The macroeconomic equilibrium is still summarized by equation (32).

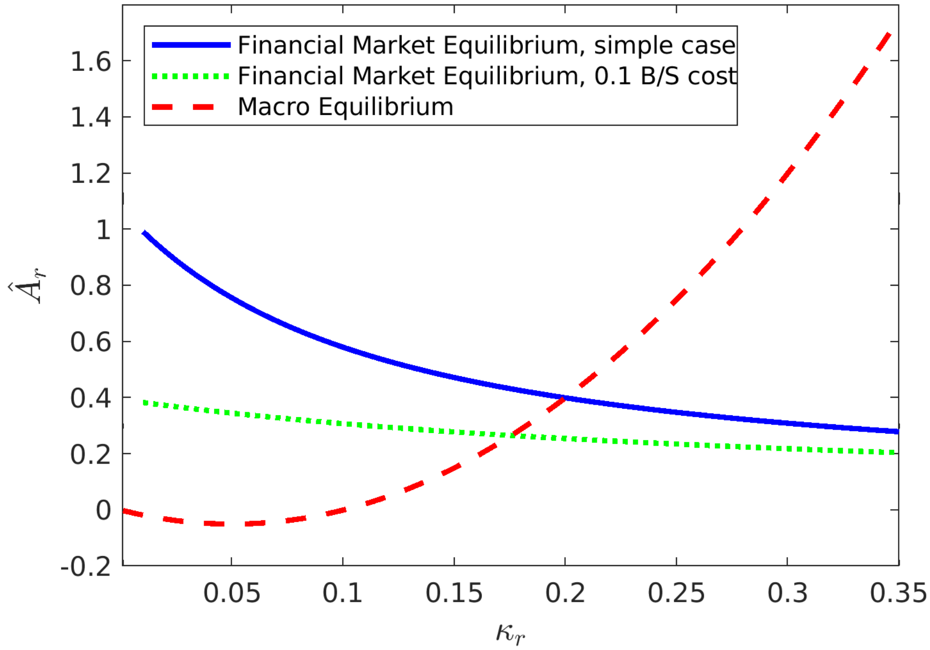


Figure 6: General Equilibrium in Extension 1.

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

4.2 Extension 2: Repo Importance

Relaxation:

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

In this extension, the solution to the affine coefficient remains unchanged. The new mapping efficiency is

$$\hat{A}_r = \frac{1-\eta}{T} \int_0^T \frac{1 - e^{-\kappa_r \tau}}{\kappa_r \tau} d\tau + \eta. \tag{35}$$

The macroeconomic equilibrium is still summarized by equation (32).

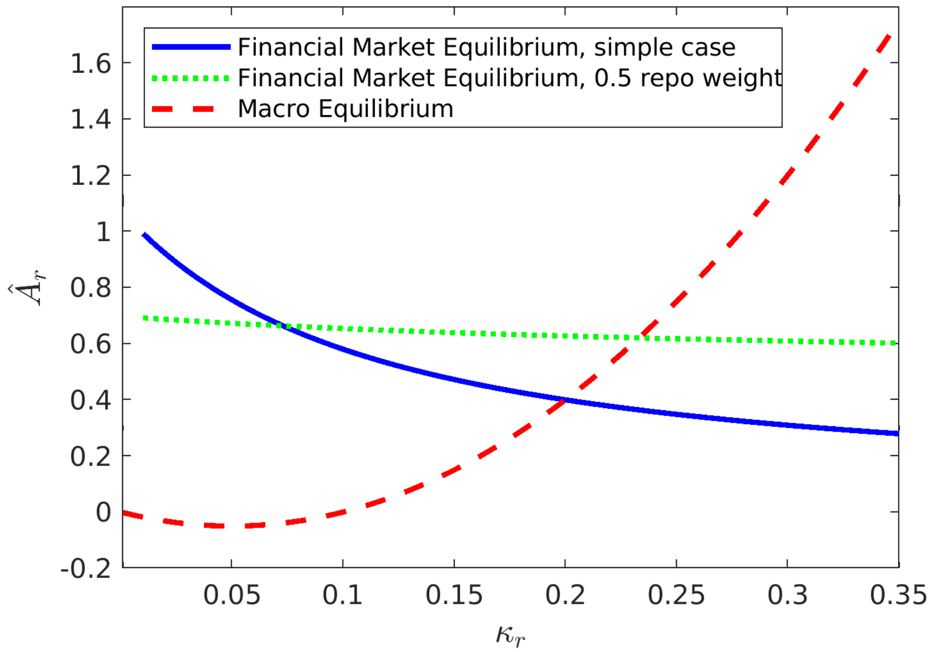


Figure 7: General Equilibrium in Extension 2.

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

5 Calibration

Parameter	Value	Description	Target
Effective 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)
$\frac{1}{\rho_h+\rho_d}$	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 balance sheet cost	LSAP1 avg. repo rate inst. response

5.1 Bond and Repo Weights

- Set $\eta^i(\tau)$ to match the average maturity structure of outstanding Treasury securities between 1985 and 2007. This period is chosen to reflect an economic environment with low risk, which is the assumption used when calibrating the macroeconomic parameters.
  - Data with various frequencies can be found from CRSP U.S. Treasury database (IU has subscription). I use monthly frequency data.
  - Ray (2019) fits a Gamma distribution but the model analogue somewhat understates the longer terms. I use Gaussian kernel density estimation to better fit the data as my algorithm does not require analytical solutions for those integral terms involving $\eta^i(\tau)$.
- Set $\eta^R(\tau)$ to match the average maturity structure of Treasury securities held by Money

Market Funds (MMF) in the repo form.[17]

- Data for 2010-present can be found from N-MFP database (publicly available).
- Another non-public source: Federal Reserve Liquidity Monitoring Report (FR 2052) has daily data about BHCs' repo positions by collateral class and maturity.
- Likewise, I use the Gaussian density estimation as my algorithm does not require analytical solutions for those integral terms involving $\eta^R(\tau)$.

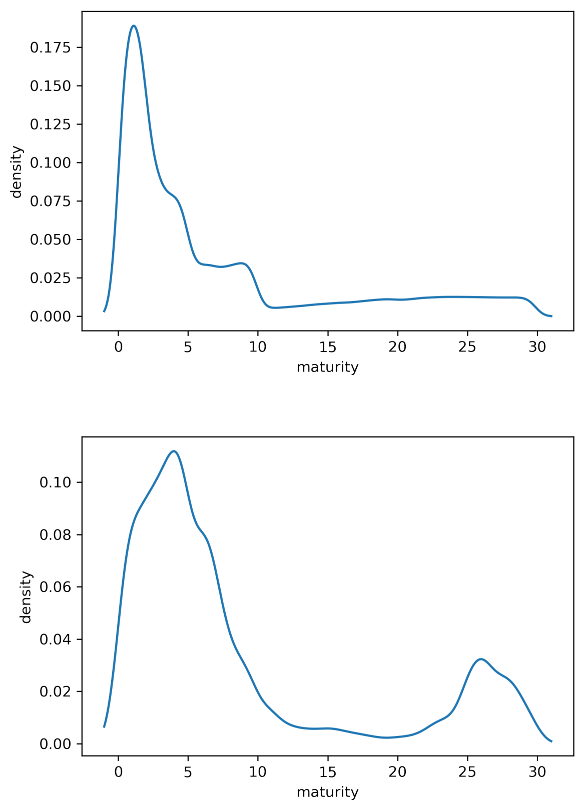


Figure 8: Density estimations of borrowing weights.

Note: Top panel represents the average density of the remaining maturity for all treasury securities outstanding during the period of 1985-2007. Bottom panel represents the average density of the remaining maturity for treasury securities used as collateral in a repo transaction during the period of 2011-2018.

[17]"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 et al. (2021) for a comprehensive analysis of MMFs participation on repo market.

5.2 Macroeconomic Parameters

- Assume repo has zero weight, there is no risk aversion, financial regulation, or QE shock, i.e., $\eta^R(\tau) = 0$, $\rho = 0$, $\lambda = 0$, $S_t = r_t$, this helps get rid of term structure parameters and make the calibration of the macroeconomic parameters easier.

- Take discount factor χ and intertemporal substitution elasticity ς^{-1} from the literature. The taylor rule inflation coefficient ϕ_π and output coefficients ϕ_x enters the system as multiplications of $\psi_r\phi_\pi$ and $\psi_r\phi_x$. I choose $\pi_\pi = 3$ and $\pi_x = 2$ to meet the unique solution condition.

- For the remaining three parameters, do a joint moments-matching calibration. I use two instantaneous IRFs from CEE (2005), IRF(π_t to r_t)=-0.15 and IRF(x_t to r_t)=-0.8, and one empirical moment Var(r_t)=3.5013. Detailed calibration steps are in appendix C1.

5.3 Term Structure Parameters

- Given effective rate and macroeconomic parameters, calibrate the remaining term structure parameters. In the baseline specification, I assume the overall repo importance is the same as the bond.

- To set $\theta(\tau)$, match the maturity distribution of Treasury LSAP1 purchases between March 18 and October 31, 2009. Data is from D'Amico and King (2013).

- Since $\alpha(\tau)$ always enters the system multiplicatively with either λ or $\frac{1}{\rho_h+\rho_d}$, I adopt $\alpha(\tau) = 5.21e^{-0.297\tau}$ from Vayanos and Vila (2021) and estimate the other two parameters λ and $\frac{1}{\rho_h+\rho_d}$.

- $\frac{1}{\rho_h+\rho_d}$, σ_β, κ_β, and λ are chosen simultaneously to match the impulse responses in yields and repo rates to FFR and QE shocks. The optimal set of values are chosen to minimize the sum of squared errors between model and data moments.

  - Swanson (2021) estimates the response in yield curve to FFR shocks using data during 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.

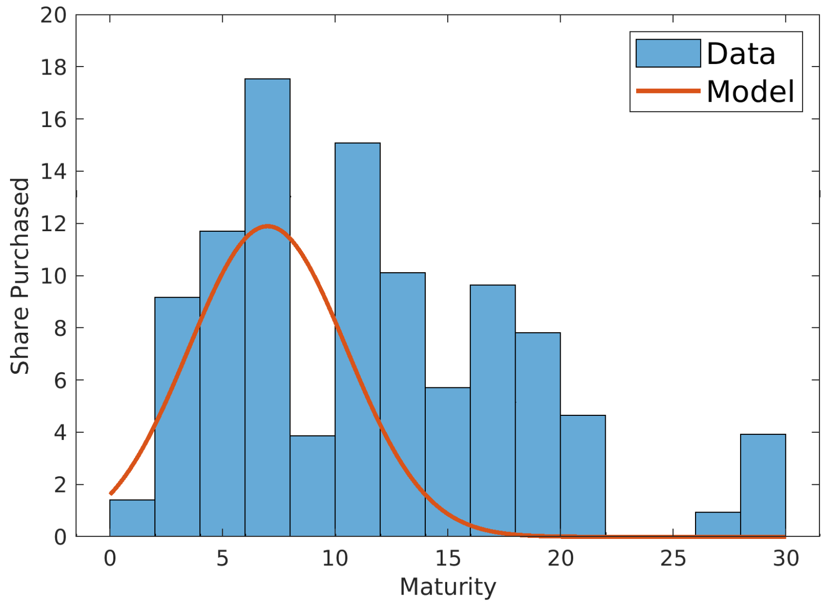


Figure 9: Polynomial fitting of demand factor location.

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

- Using the the Gurkaynak et al. (2007) yield curve data, the LSAP1 shock causes a 50 bps instantaneous drop in 10-year yield.
- Ihrig et al. (2018) estimate that during LSAP1, the half-life of 10-year yield response is about 2 years after the shock.
- D'Amico and King (2013) documents that during LSAP1, about $300b Treasury securities were purchased and 71% of them were off-the-run. This implies that $87b of on-the-run were pucased and $213b of off-the-run were purchased. D'Amico et al. (2018) find that during 2009/3-2012/12, the ratio of Fed purchasing versus selling is 46:1 and 1.82:1, respectively for on-the-run and off-the-run. What's more, Fed buying $0.29b on-the-run securities decreased the repo rate by 0.224 bps, Fed buying $0.24b off-the run securities decreased the repo rate by 0.085 bps. Combining these information, I first calculate that Fed net purchased $85b on-the-run securities and $96b off-the-run securities. Next, purchasing on-the-run securities drop repo rate by 65bps and purchasing off-the-run securities drop repo rate by 34bps. The average response in repo rate is 48bps.

6 Explanations for GFC and Covid-19

6.1 Flight-to-Safety in GFC

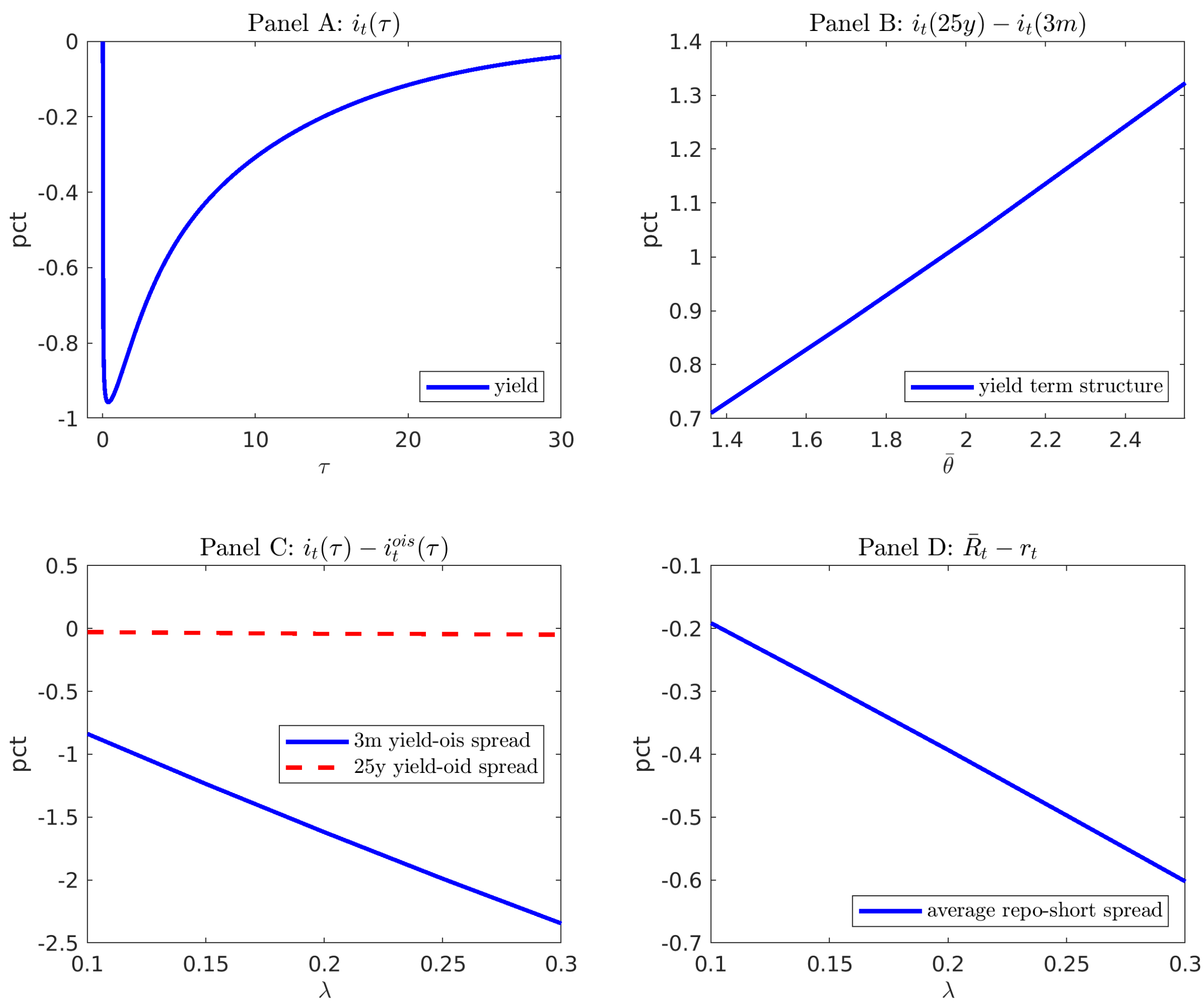


Figure 10: Model-implied instantaneous responses to a rise in habitat demand concentrated at short maturity.

Note: Parameters are set as follows: $\kappa_{r\beta} = 0$, $\kappa_r = 0.2$; $\sigma_r = 0.1$; $\kappa_\beta = 0.17$; $\sigma_\beta = 0.011$; $\alpha(\tau) = 4.168 * exp(-0.297 * \tau)$, $\theta(\tau) = 465.653 * (exp(-0.297 * \tau) - exp(-0.307 * \tau))$, $\rho = 21$, $\lambda = 0.1$.

My explanation

- There is net positive demand of Treasury bonds, especially for short maturities, from habitat investors due to the flight-to-safety and flight-to-liquidity motivation. Pre-GFC, there was no SLR that results in balance sheet cost. In response to the demand shock, dealers engage in short-selling bonds to satisfy the additional demand. Assuming that naked short selling is costly, the dealer's problem is

$$\max_{X_t(\tau),Q_t^d(\tau)} E_t\left[dW_t^d\right] - \frac{1}{2\rho_d}\mathrm{Var}_t\left[dW_t^d\right],$$

$s.t.$

$$dW_t^d - W_t^d r_t d_t = \int_0^T \underbrace{X_t(\tau)}_{\text{direct holdings}} \underbrace{(\frac{dP_t(\tau)}{P_t(\tau)} - r_t dt)}_{\text{excess return}} d\tau + \int_0^T \underbrace{Q_t^d(\tau)}_{\text{repo supply}} \underbrace{(R_t(\tau) - r_t)}_{\text{repo wedge}} dt d\tau$$

$$- \int_0^T \underbrace{\left(-Q_t^d(\tau) - X_t(\tau)\right)}_{\text{naked short selling}} \underbrace{\Lambda_t(\tau)}_{\text{S/S cost}} dt d\tau.$$

The indifference condition says

$$R_t(\tau) - r_t = -\Lambda_t(\tau),$$

the repo wedge must be equivalent to the negative of the short selling cost. Intuitively, this condition says the marginal cost of short selling bonds must be the same as the marginal benefit. Assuming the same linear structure of the marginal short selling cost $\Lambda_t(\tau) = \lambda\left(-Q_t^d(\tau) - X_t(\tau)\right) = -\lambda\left(Q_t^d(\tau) + X_t(\tau)\right) = \lambda H_t(\tau)$, when there is positive demand shock, $H_t(\tau)$ increases, the short selling cost rises, the repo wedge drops. This is consistent as the repo specialness documented in papers such as Duffie (1996) and Jordan and Jordan (1997). As holding Treasury securities become more "convenient" for dealers, they are willing to provide repo fund at lower rates for hedge funds. The convenience premium is established in the bond cash market as well. Because holding Treasury securities helps avoid the short selling cost when dealers need to finance bonds, they are willing to pay higher price for the bonds.

- Why is the term structure steepened? During GFC, the flgiht-to-safety motivates excess habitat demand for short-term Treasury securities. Arbitrageurs have to short sell more short maturity bonds to satisfy the excess demand. Since demand risk is an increasing function of maturity, the positive demand shock decreases the portfolio's exposure to short rate risk by more than to demand risk. Therefore, the price of the short rate risk drops by more than that of the demand risk. As short maturities are more short rate sensitive, their required risk compensation reduces by more than the long maturities.

- Panel B indicates that such positive term structure steepens with the size of the demand shock. A larger shock decrease the short rate risk of aggregate holdings by more, so the risk price of the short rate decreases by more, therefore the short maturities demand less risk compensation.

- The Treasury-OIS spread represents the inconvenience yield of holding Treasury securities. In this exercise, the Treasury-OIS spread is negative, implying a convenience premium of holding Treasuries. Panel C illustrates the convenience yield as an increasing function of the marginal short selling cost. This supports the hypothesis that the steepening term structure of the Treasury-OIS spread during early GFC is due to the short selling by dealers. As dealers provide the extra demand through short selling, the marginal cost rises. Therefore holding Treasury securities brings extra benefit of saving the short selling cost, and the amount of saving is proportional to the marginal cost. As a result, higher short selling cost makes holding Treasury securities more "convenient".

- Panel D illustrates the repo specialness as a negative response in the repo wedge. When short selling cost is high, dealers tend to offer low interest rate of lending repo cash to attract bonds. Thus, the magnitude of the specialness is increasing with the marginal short selling cost.

Why did the flight-to-safety result in drop in output and inflation?

- From the previous analysis, the flight-to-safety activity causes a convenience yield in holding Treasury securities, therefore the overall saving benefit is smaller, which should boost the economy.

- I claim that the flight-to-safety can be a cause of recession if households receive utility of holding saving assets. Consider the representative household, if they receive utility from holding saving assets that is linear in the holding size, then their Euler equation becomes

$$\frac{dC_t/dt}{C_t} = \tilde{i}_t + \varpi_t - \rho - \pi_t, \tag{36}$$

  where ϖ_t is the marginal utility of holding safe assets and it is referred to as the safety preference shock.

- A stronger safety preference increases the marginal utility of holding safe assets and discourages the consumption. When the flight-to-safety shock outweighs the demand shock, output and inflation drop.

6.2 Flight-From-Safety in Covid-19

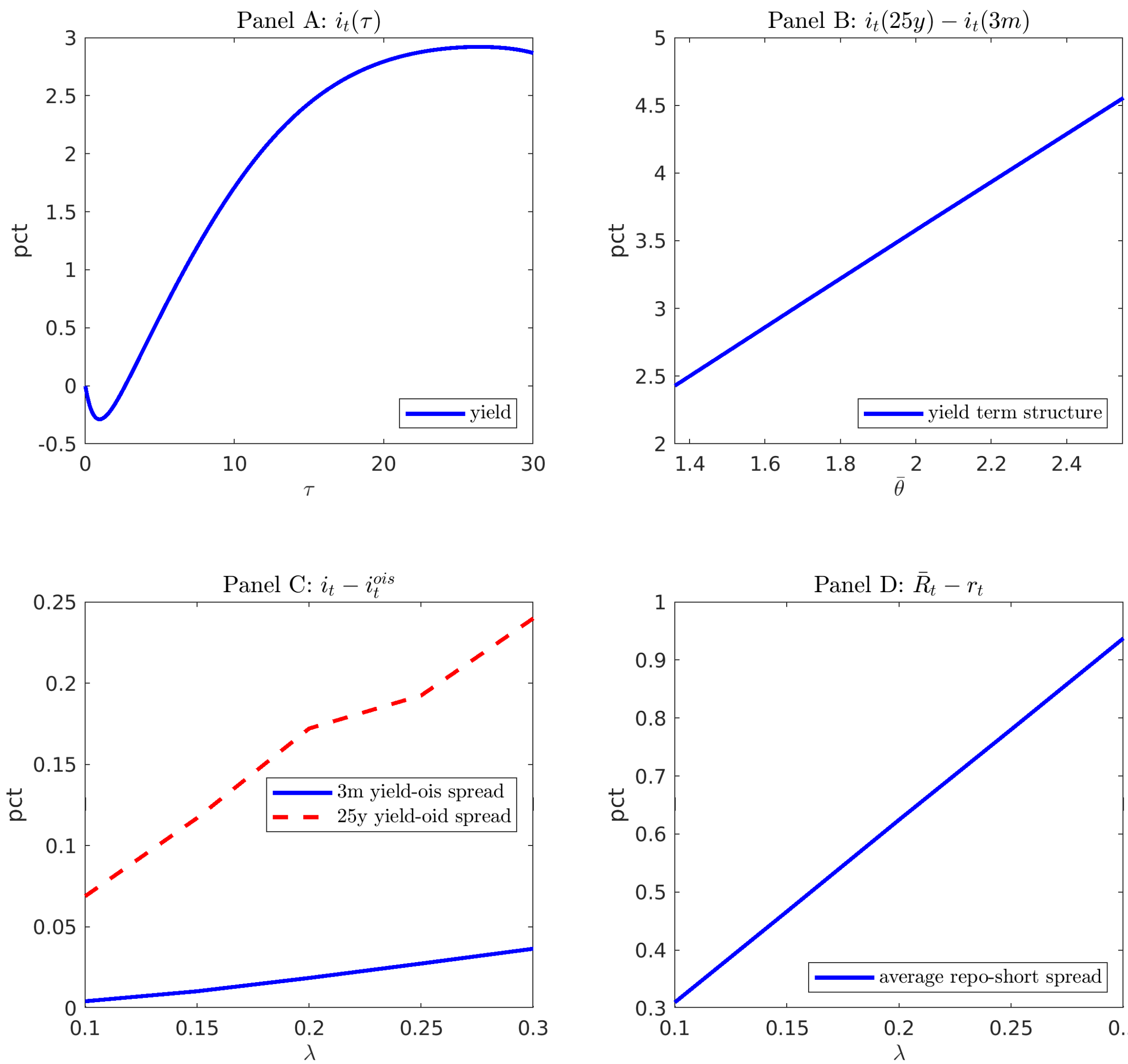


Figure 11: Model-implied instantaneous responses to a drop in habitat demand concentrated at long maturity.

Note: Parameters are set as follows: $\kappa_{r\beta} = 0$, $\kappa_r = 0.2$; $\sigma_r = 0.1$; $\kappa_\beta = 0.17$; $\sigma_\beta = 0.011$; $\alpha(\tau) = 4.168 * exp(-0.297 * \tau)$, $\theta(\tau) = 465.653 * (exp(-0.297 * \tau) - exp(-0.307 * \tau))$, $\rho = 21$, $\lambda = 0.1$.

My explanation

- There is large-scale selling of long-term Treasury bonds from habitat investors.
- Why is the term structure steepened? The importance of demand risk relative to short rate risk rises with maturities. When habitat investors decrease holdings of long maturities, arbitrageurs have to increase holdings of long maturities. The portfolio's exposure to the demand risk rises more than to the short rate risk, thus the price of demand risk rises by more than that of short rate risk. In other words, it's more costly to bear demand risk. Because the long maturities are highly demand-sensitive, so the required risk premia rise by more than the short end.
- Why short maturities have a reverse sign in response to the demand shock? The supply shock mainly targets long end, so the risk price of short rate raises only subtly. On the other hand, the short end habitat demand is very price elastic, therefore the bond price drop encourages considerable amount of habitat demand, which pushes the price to be higher. When this indirect force outweighs the required risk compensation, the short end yield can actually drop even through there is a positive supply shock.
- Panel B indicates that such positive term structure steepens with the size of the supply shock. A larger shock increases the demand risk of aggregate holdings by more, so the risk price of the demand factor increases by more, therefore the long maturities require higher risk premia.
- The Treasury-OIS spread represents the inconvenience yield of holding Treasury securities. In this exercise, the OIS rate is defined as the interest rate without balance sheet cost. Panel C illustrates the inconvenience yield as a increasing function of the leverage requirement tightness. This supports the hypothesis that the steepening term structure of the Treasury-OIS spread during March 2020 is due to the balance sheet cost. As dealers absorb the extra supply, their balance sheet expansion induces higher balance sheet cost, therefore holding Treasury securities become more "inconvenient".
- I define the repo wedge in this exercise as the repo rate minus the short rate. Panel D illustrates the repo wedge as an increasing function of the leverage requirement tightness. As holding Treasury securities become more "inconvenient" for dealers, they charge higher repo rate for hedge funds to rent space on their balance sheets.

7 Model Performance

Excess Return Predictability (Fama & Bliss, 1987)	
Specification	$\frac{1}{\Delta\tau}\log\left(\frac{P_{t+\Delta\tau}(\tau-\Delta\tau)}{P_t(\tau)}\right)-y_t(\Delta\tau)=a_{\mathrm{FB}}(\tau)+b_{\mathrm{FB}}(\tau)\left(f_t(\tau-\Delta\tau,\tau)-y_t(\Delta\tau)\right)+e_{t+\Delta\tau}(\tau)$
Empirics	$b_{\mathrm{FB}}(\tau)>1$ and $\uparrow$ in τ
This paper	$0<b_{\mathrm{FB}}(\tau)<1$ but irrelevant in τ
Repo Specialness (Jordan & Jordan, 1997)	
Specification	$\log\left(\frac{P_t(\tau)}{P_t^{ref}(\tau)}\right)=a_{JJ}(\tau)+b_{JJ}(\tau)(r_t-R_t(\tau))+\epsilon_t(\tau)$
Empirics	$b_{JJ}(\tau)>0$
This paper	$b_{JJ}(\tau)>0$ for most τ
Four-Equation NK (Sims et al. , 2023)	
Moments	$\omega_{xr}=-0.7$, $\omega_{x\beta}=-1.35$, $\omega_{xr}=-0.005$, $\omega_{\pi\beta}=-0.01$, $\kappa_r=0.46$, $\kappa_{r\beta}=0.002$
This paper	$\omega_{xr}=-0.73$, $\omega_{x\beta}=-0.97$, $\omega_{xr}=-0.008$, $\omega_{\pi\beta}=-0.01$, $\kappa_r=0.39$, $\kappa_{r\beta}=0.0037$

7.1 Excess Return Predictability

1. $\Delta\tau$=0.03 year=11 days, Δt= 1 day.

2. Given the calibrated state variable transition matrix Γ and volatility matrix Σ, simulate short rate and demand factor according to equation (5).

3. Simulate bond prices, bond yields, and repo rates according to equation (10), (22), and (23).

4. Simulate forward rates as

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

5. Simulate holding returns as

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

6. Regress excess holding return on excess forward rate for different maturities.

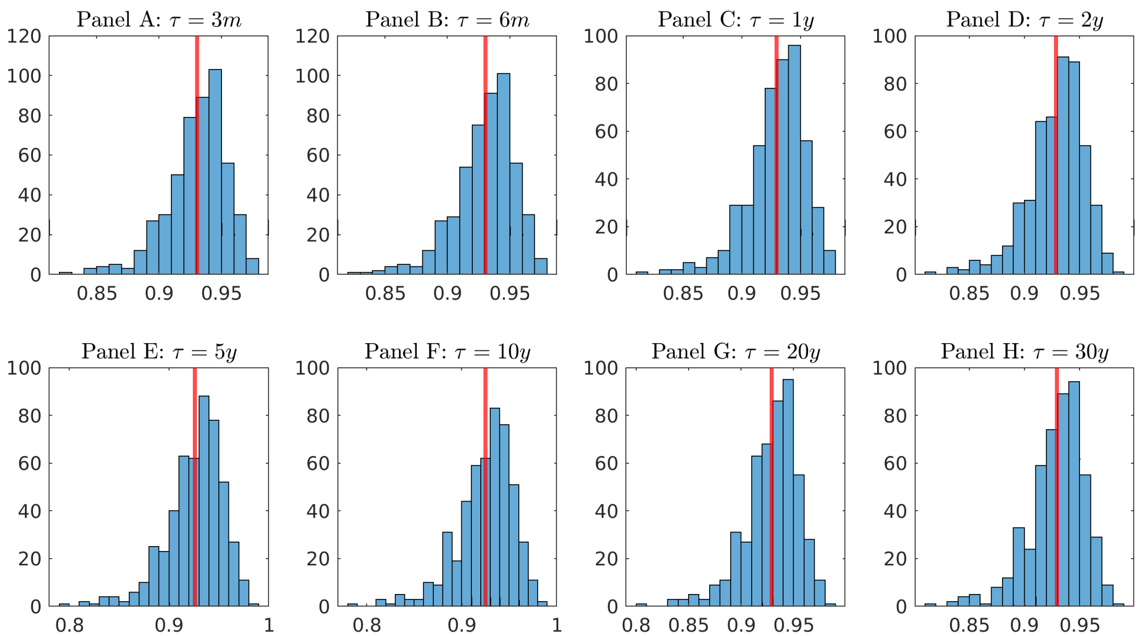


Figure 12: Estimated Fama-Bliss (1987) coefficient.

7.2 Repo Specialness

1. Use the short rate, bond prices, bond yields, and repo rates simulated from previous exercise.

2. To simulate reference prices, set $\lambda = 0$. This represents the scenario with no friction on the repo market. Resolving for the general equilibrium terms, I have a new set of $A_r(\tau)^{re}$, $A_\beta(\tau)^{re}$, and $C(\tau)^{ref}$. Simulate reference bond prices as

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

3. Regress cash market price premium on repo premium for different maturities.

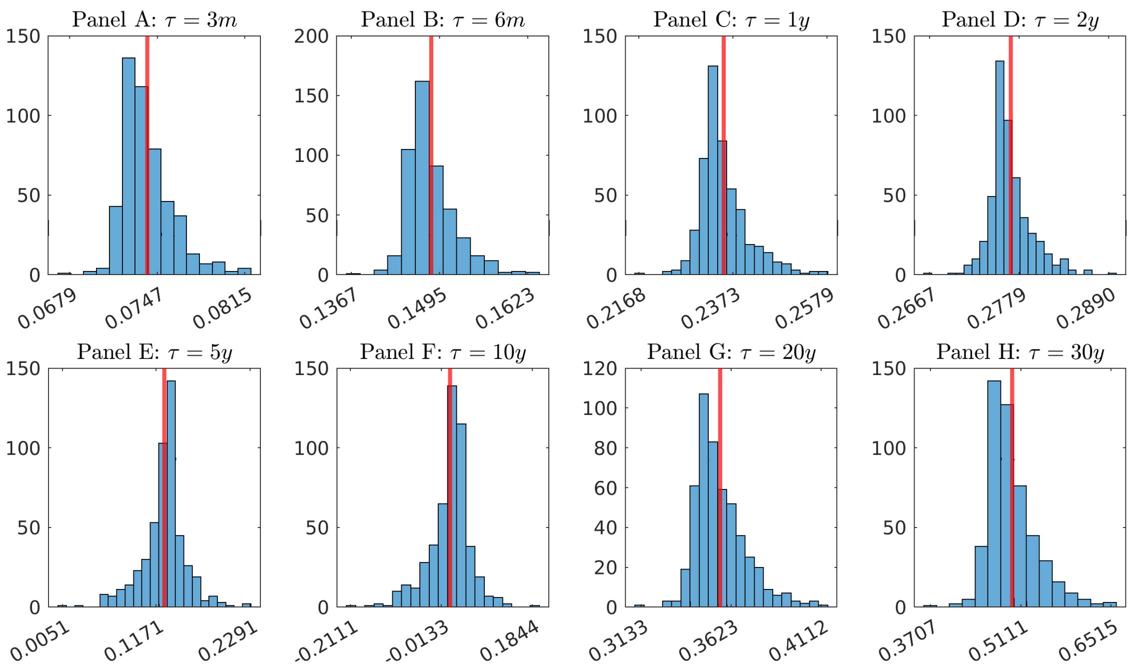


Figure 13: Estimated Jordan-Jordan (1997) coefficient.

7.3 Four-Equation New Keynesian Model

Sims et al. (2023) propose a four-equation New Keynesian model that can be used to study unconventional monetary policy. In this section, I provide a comparison between their model and mine. Sims et al. (2023) model features

- Parent and child households: Parents consume, save in short-term bonds, supply labor, transfer to children and financial intermediaries. Children consume and issue long-term bonds.
- Leverage constraint: financial intermediary's long-term bonds holding can not exceed a multiple of equity value, where the multiple is exogenous and interpreted as the credit shock.
- Financial intermediary and central bank together absorb the long-term bond supply. The QE shock is defined as the part of long-term bond held by central bank.

Rewriting the IS and Phillips curves in Sims et al. (2023) in continuous-time:

$$\begin{aligned} \mathrm{d}x_t &= a\left(r_t - b * qe_t + c * x_t - d * \pi_t - r^*\right)\mathrm{d}t - b * d(qe_t), \\ d\pi_t &= (e * qe_t - c * x_t - (1-d)\pi_t)dt. \end{aligned}$$

Both credit and QE shocks affect the system through the financial intermediary's leverage constraint. Note that these two shocks affect the equilibrium system equivalently after accounting for the scaling factor, dropping one of them does not affect the equilibrium results. For simplicity, in the following analysis, I abstract from the credit shock and focus on QE shock. Recall that in my model, the equilibrium IS curve can be written as

$$\mathrm{d}x_t = \varsigma^{-1}\left(r_t - \pi_t - \bar{r}\right)\mathrm{d}t + \varsigma^{-1}\left(\int_0^T \eta^i(\tau)(i_t(\tau) - r_t)\mathrm{d}\tau + \int_0^T \eta^R(\tau)(R_t(\tau) - r_t)\mathrm{d}\tau\right)dt.$$

Compared to the traditional 3-equation specification, the IS curve in my model has an additional term that is a function of the term premia of bond yields and repo rates, both of which are endogenous functions of the risk factors. Comparing Sims et al. (2023) and my model:

- How does QE affects interest rates? The QE shock affects long-term rate through leverage constraint in Sims et al. (2023). A positive QE shock decreases financial intermediary's long-term bond holding, therefore the leverage constraint relaxes. Because the financial intermediary needs to pay short rate on the equity, the QE shock results in a decrease in long-term rate relative to the short rate. In my model, the QE shocks takes effect through market segmentation by offloading risk and balance sheet cost for arbitraguers trading the targeted segment.

- How does QE affects macro variables? In Sims et al. (2023), the rate change affects the marginal benefit (cost) of saving (borrowing) for parents (children), therefore affects the consumption. The rate change implies 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. Therefore QE enters the Phillips curve endogenously. In my model there is no welfare transfer between households, therefore the QE shock does not enter the Phillips curve.

- When will the model collapse to the standard three-equation? Without QE shock, Sims et al. (2023) reduce to a standard 3-equation model. This is not the case in my model. The short rate risk by my model is endogenously priced by the interest rates at which households save and borrow. Therefore the IS curve still deviates from the standard 3-equation specification even without the QE shock. In fact, the standard 3-equation model is achieved 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.

- Can my model replicate important results? See figure 8.

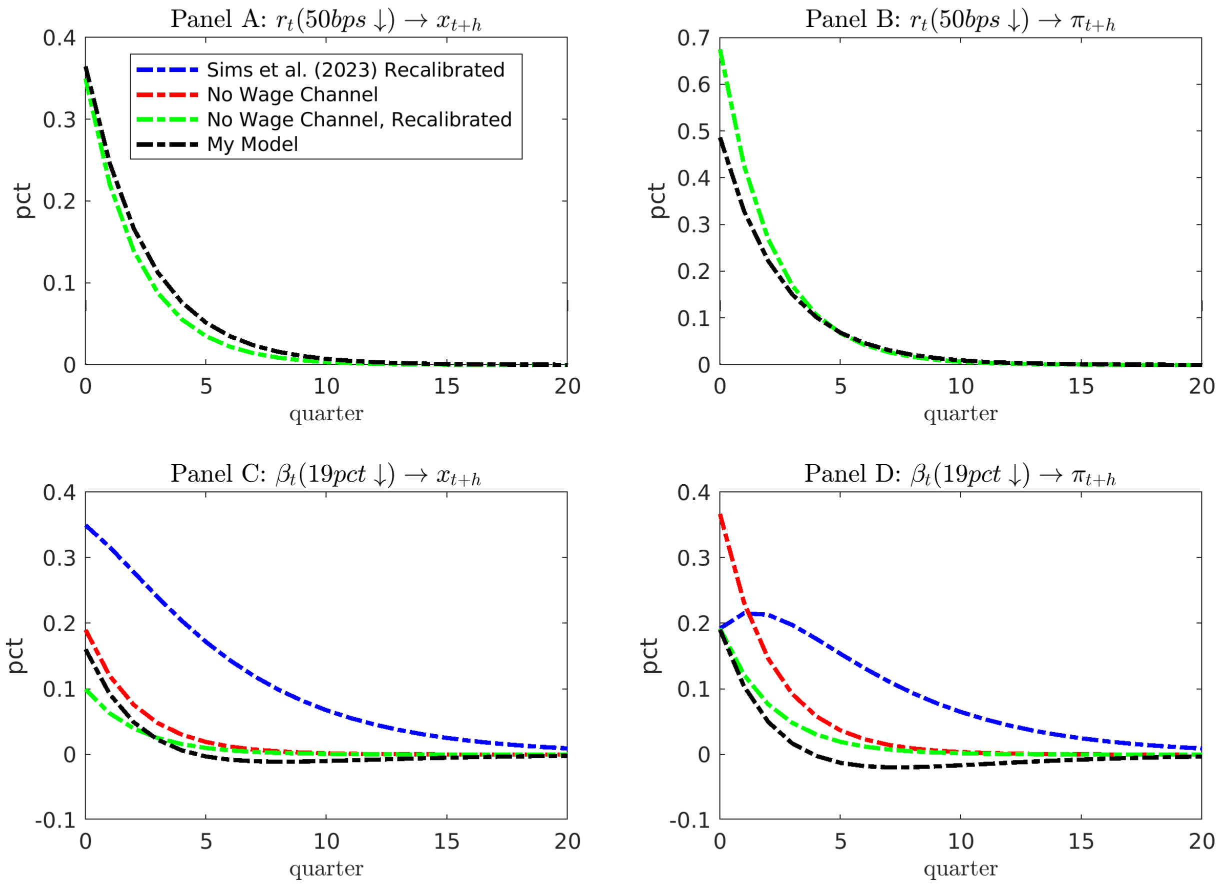


Figure 14: IMFs of recalibrated versions of Sims et al. (2023).

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 (24) results in a zero Repo spread., and repo rates are all equal to the short rate. In this extreme scenario of segmentation equilibrium, there is zero passthrough to bond yields and perfect passthrough to repo rates.

When arbitrageurs exist, the short rate is transmitted into bond yields by carry trades. To better analyze the transmission mechanisms, I employ the concept of instantaneous forward rates.

$$f_t(\tau) \equiv -\frac{\partial log P_t(\tau)}{\partial \tau} = A'_r(\tau) r_t + A'_\beta(\tau)\beta_t + C'(\tau) \tag{37}$$

When the short rate drops, arbitrageurs shift wealth from the short rate to bonds, raising bond prices and lowering forward rates. Similar to Vayanos and Vila (2021) and Ray (2019), carry trades in this model incur costs that hinder the passthrough of short rate shock. Two types of cost are involved in carry trades: risk cost and balance sheet cost. Risk cost, similar as in Vayanos and Vila (2021) and Ray (2019), is mainly affected by risk aversion $\frac{1}{\rho_d+\rho_h}$, short rate volatility σ_r, and habitat demand price elasticity $\alpha(\tau)$. Distinguished from the previous literature, this paper introduces balance sheet cost, representing the friction incurred as the portfolio adjustment occupies the dealer's balance sheet. The magnitude of balance sheet friction is an increasing function of financial regulation tightness λ.

The transmission of short rate to overnight repo rates involves only the balance sheet cost but no risk cost. This is because in this model, the two risk factors affect the system by determining the bond price. When hedge funds borrow cash on the repo market using bonds as collateral, they do not face binding borrowing constraints and all repo contracts are assumed without default risk. [18] That being said, the price fluctuations in bond price does not directly affect arbitrageur's first order condition for repo choice. Repos in this model are risk-free and the only friction involved is the costly holding represented by the balance

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

sheet cost.

Figure 4 exhibits 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 financial regulation ($\lambda = 0$). For the instantaneous forward rate and bond yields, the Expectations Hypothesis (EH) holds in the sense that it is equal to the expected future short rate (Panel A and B, blue solid line). However, the reasons for the perfect passthrough are different in the two cases. When risk aversion is zero, the arbitrageurs are not averse to the risk borne in the carry trades. Therefore, carry trades do not induce any risk cost and the transmission is frictionless. In the second case when the habitat demand is completely price inelastic, the arbitrageurs impact the bond prices without carry trades. This corresponds to infinitely price impactful arbitrageurs.

The introduce of risk aversion and habitat demand price elasticity breaks the EH. As shown by the red dashed line in panel A of figure 4, the instantaneous forward rate underreacts to short rate than the expected future short rate. This is due to the risk cost that arises when risk-averse arbitrageurs conduct carry trades. Unlike forward rates or bond yields, the transmission to repo rates is still perfect in this scenario (Panel C, red dashed line). Hedge funds do not bear the risk of the price fluctuation in the underlying collateral of repo contracts, therefore the short rate still maps one-to-one on repo rates.

Finally, the balance sheet costs further obstacles the transmission to the forward rate (Panel A and B, green dotted line). Adjusting balance sheet now exposes arbitrageurs to additional cost, squeezing the profit space for carry trades. The 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 4. Interestingly, in my model, because carry trades expose arbitrageurs to both risk cost and balance sheet cost, short rate transmission can be less than one-to-one even without risk aversion. When $\alpha(\tau) > 0$, arbitrageurs increase bond holding to benefit from the yield spread, such adjustment expands the balance sheet size. If financial regulation requires a balance sheet cost ($\lambda > 0$), then the net expected excess return reduces, which discourages carry trades and therefore harms the transmission to forward rates. Repo lending also occupies balance sheet, so the repo rate has to compensate dealers for the higher lending cost.

The baseline calibration gives a fast mean-reverting speed of short rate, which largely dominates the transmission of short rate to forward rates and yields. The frictions harm the transmission but this obstruction is not quantitatively important. In the baseline econ-

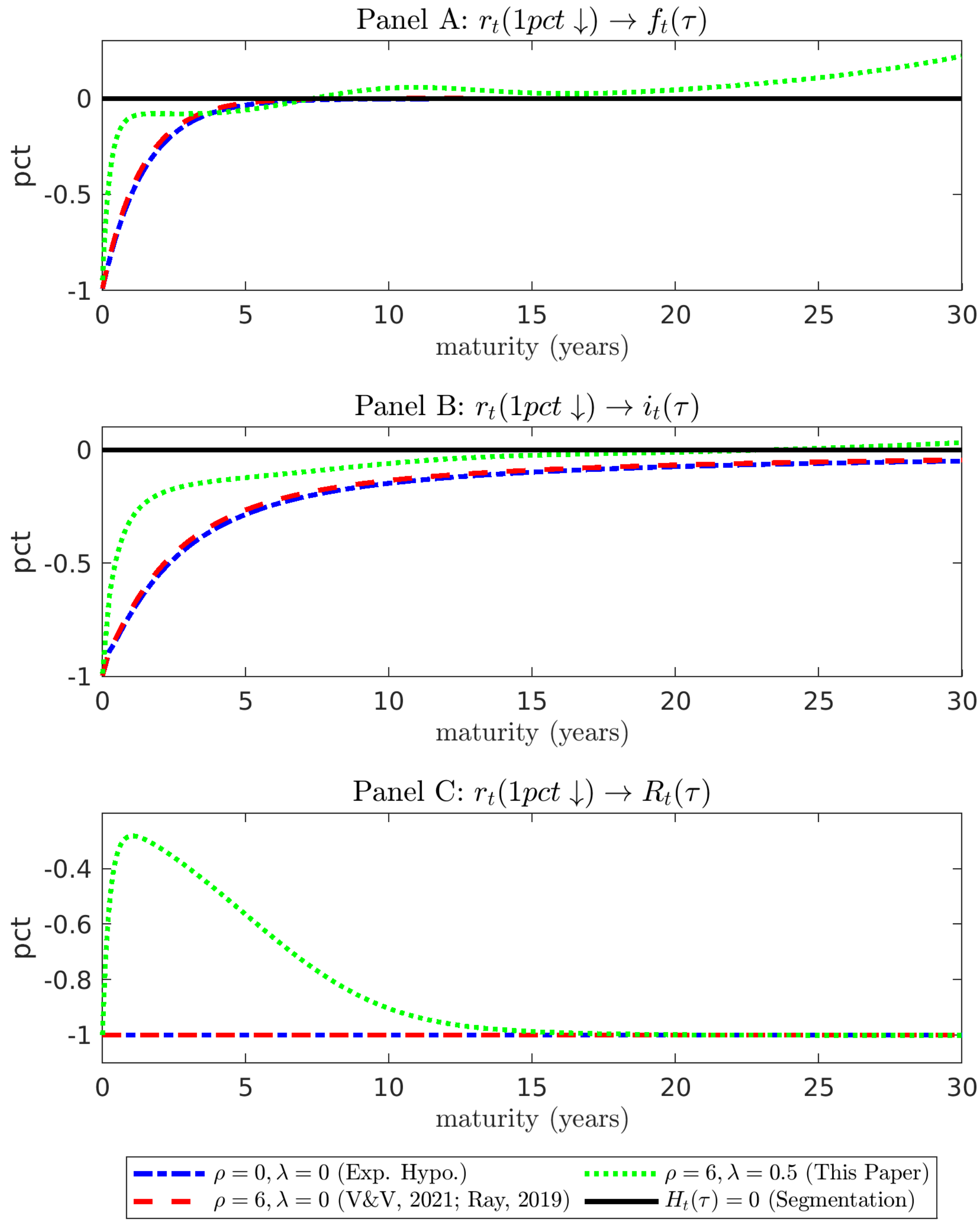


Figure 15: Transmission of short rate shock to asset prices.

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.

omy with the aggregate nominal rate involving only bond yields, the frictions do not play an important role in affecting the conventional monetary policy's effect on macroeconomic

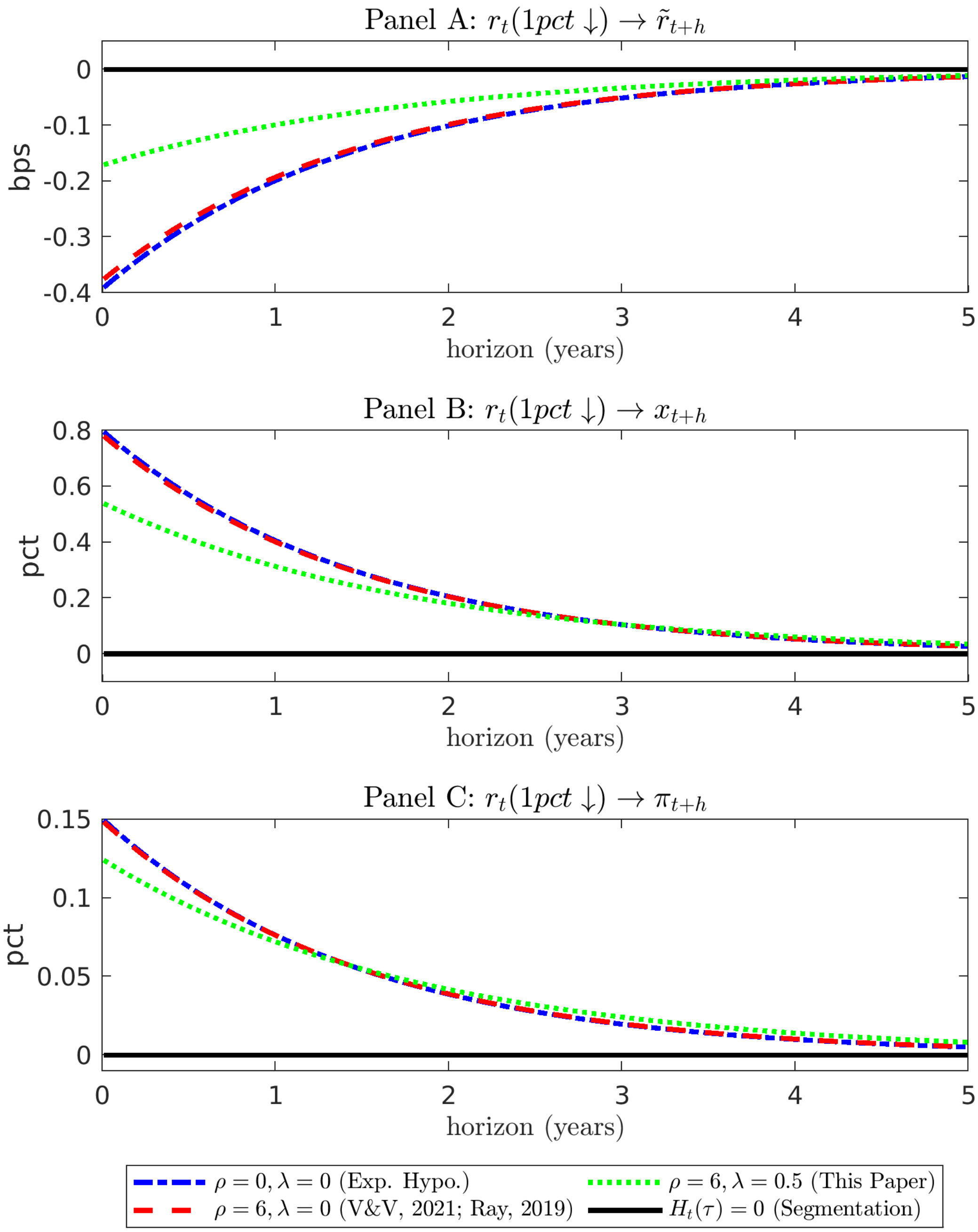


Figure 16: Transmission of short rate shock to macro variables.

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.

variables.

8.2 Transmission of QE Shock

While frictions obstacle the transmission of short rate to asset prices and macroeconomic variables, they strengthen that of the demand shock. When the Expectations Hypothesis holds, the forward rates are disconnected from the demand factor and only depend on the expected future short rates. Therefore, the demand shock obtains zero effects when there is no financial friction.

Financial frictions allow the demand factor to take effect. Suppose there is a QE shock, the net demand of habitat investors increases. The higher habitat demand decreases the holdings of arbitrageurs, thus reduces arbitrageurs' marginal cost of investing in the portfolio. Arbitrageurs now demand a lower compensation to hold bonds. When the portfolio induces only risk cost, the QE shock takes effect purely through decreasing the risk prices for the short rate and the demand factor. As the risk price is unique, the effect of the QE shock has a global flavor across different maturities in the sense that all forward rates drop.

The balance sheet cost introduces a local flavor to the QE effect. The QE shock can target different maturities, and the most targeted maturity receives the largest saving in the marginal balance sheet cost as the QE shock offloads arbitrageurs' holdings. The response in the forward rate is with different signs. In the baseline calibration, QE shock concentrates around 10-year maturity. The saving in the balance sheet cost is increasing in maturity for the short end and decreasing for the long end. Therefore, the forward rate drops for short tenors and rises for longer tenors.

Although the response in the yield curve is hump-shaped in both cases with and without the balance sheet cost, the reason is different. When there is no balance sheet cost, the QE shock achieves global effects on the forward rate by raising the bond price for all maturities with the amount of appreciation being increasing in very short maturities (0-1.3 year) and decreasing in longer maturities (1.3 year). Since the yield for maturity τ is the average of the forward rates across the same length of tenor, the yield drop will also first increase in maturity then decrease, resulting in the hump-shaped yield curve response. In the second case with balance sheet cost, the QE shock achieves more local effects. The forward rates drop the most for targeted short maturities and can respond with a positive sign for longer maturities. This locality is passed to yield curve response, which results in a more prominent hump shape of the yield curve response.

The QE shock achieves transmission to the repo rates if the balance sheet cost is allowed. Since repo investors are not exposed directly to the risk factors of bond price, the repo wedge does not reflect risk prices and only involves the balance sheet cost. The repo curve response is shaped by the maturity distribution of the calibrated baseline QE shock, with the largest response generated on the short end.

Note that the best transmission is obtained when the financial markets are completely segmented in the sense that there are no arbitrageurs trading across products. This is consistent with Vayanos and Vila (2021) that in the segmented economy the price for any maturity is disconnected from short rate and only dependent on the demand shifter for that specific maturity. Suppose there is a QE shock, without arbitrageurs, the bond price has to rise considerably to discourage habitat investors from shifting their holdings. To what extent will the bond price rise depends on the habitat investors' price elasticity for that specific maturity $\alpha(\tau)$.

The previous findings pass to the macroeconomic variables. Unlike the short rate whose passthrough to the macroeconomic variables are very similar across scenarios, the QE shock attains quantitatively different responses. Recall that in the baseline calibration, the aggregate nominal rate involves only bond yields. The the same QE shock generates larger drops in yields with more financial frictions, thus also boosts more growth in output gap and inflation.

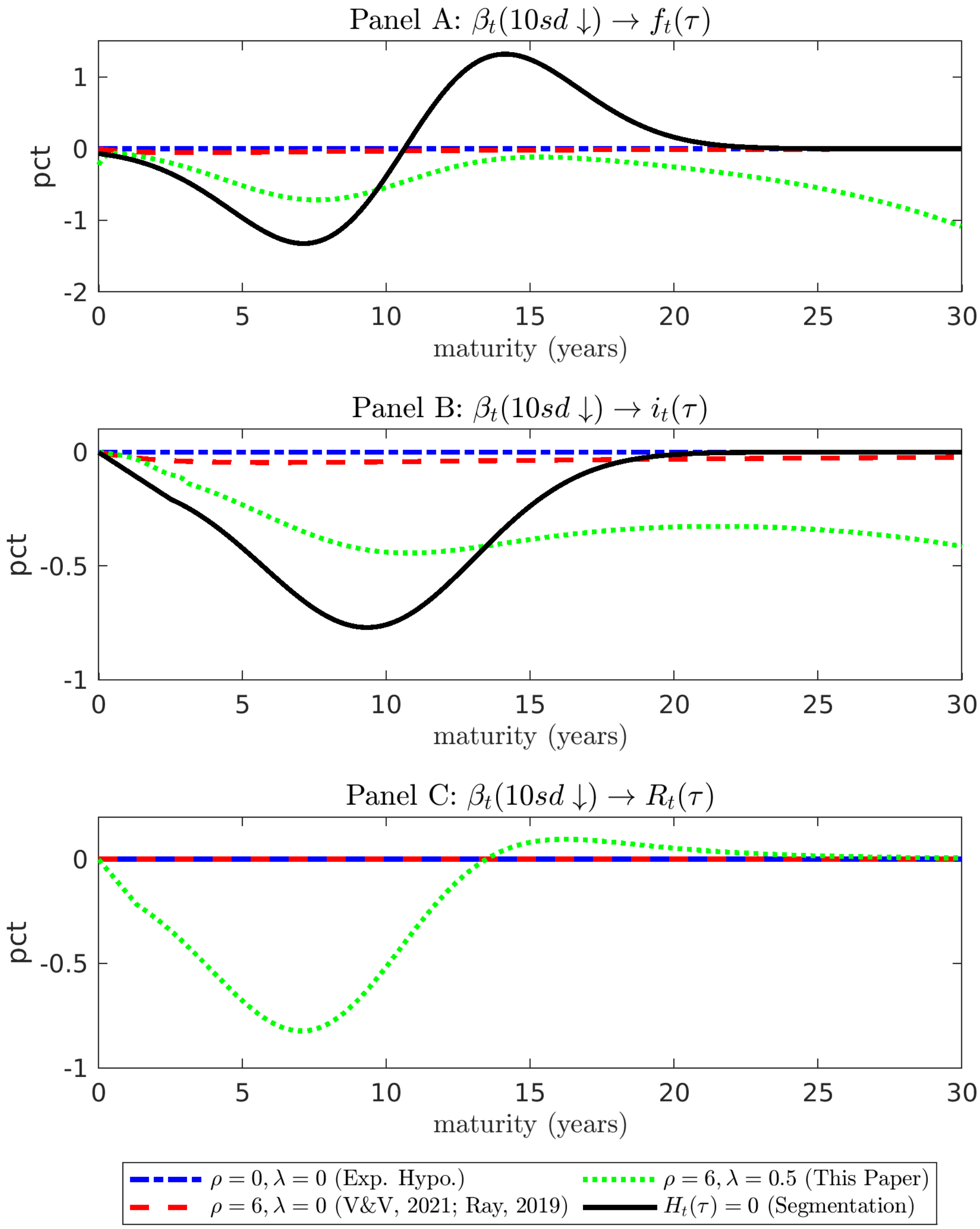


Figure 17: Transmission of demand shock to asset prices.

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.

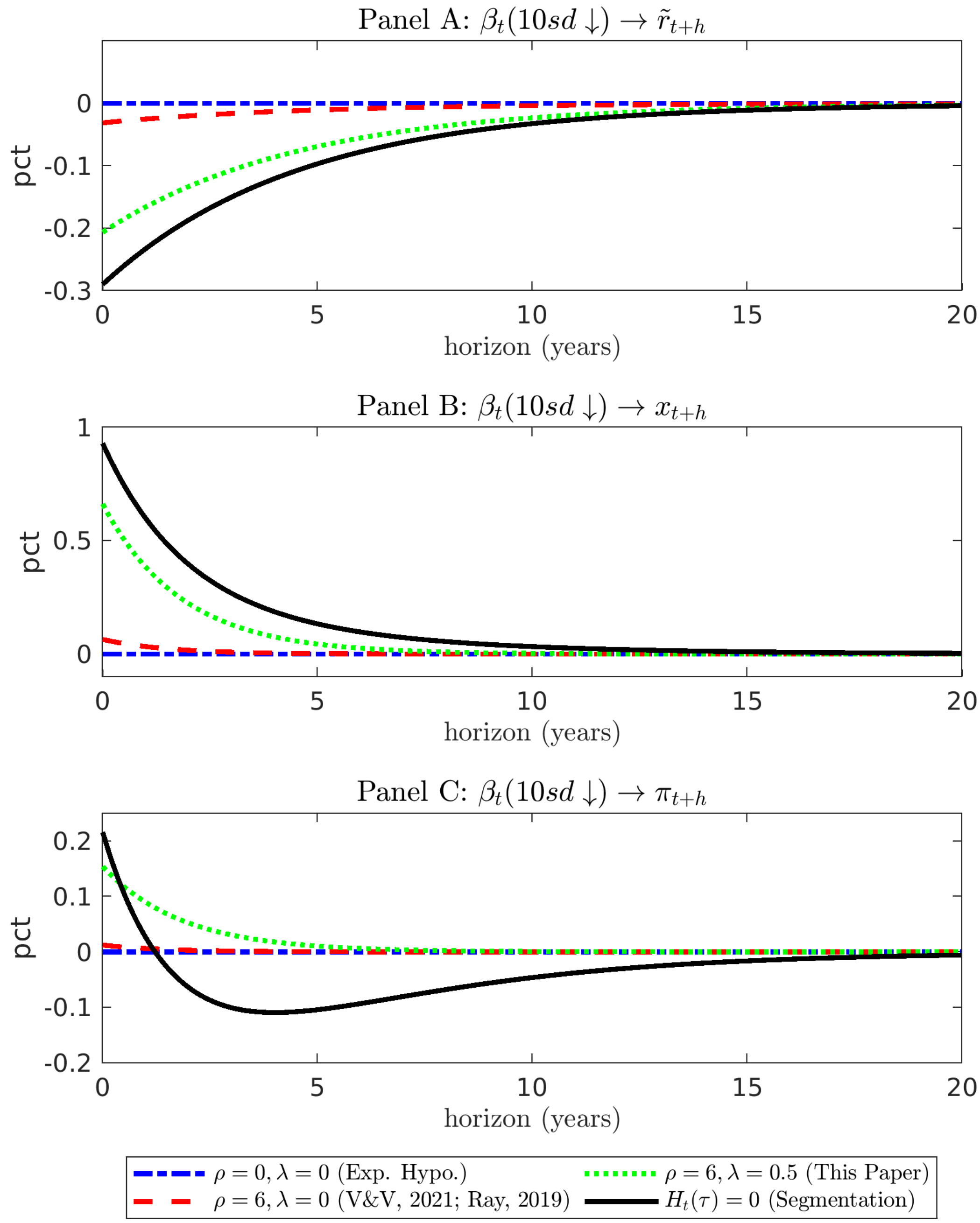


Figure 18: Transmission of demand shock to macro variables.

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.

9 Policy Analysis

	Forward Guidance
Hypothesis	Effect is significant in the short run but dissipates fast.
Exercise	Generate IRFs of forward guidance shock and discuss parameter sensitivities. Compare with the standard short rate shock.
	QE Targeting Alternative Maturities
Hypothesis	QE shock targeting long-end has largest effect.
Exercise	Generate IRFs of QE shocks targeting different maturities. The implication if central bank follows balance budget.
	Benchmark Rate Revolution
Hypothesis	Both short rate and QE gains effectiveness in the repo economy.
Exercise	Compare IRFs in repo and yield regimes. Optimal long-run policy rate target.

9.1 Forward Guidance

Consider a unit of unanticipated change in the long-term mean of the short rate r^{ss} at time zero that reverts deterministicaly to zeros at the rate $\kappa_{\bar{r}}$. To compute the effect, 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 $logP_t(\tau) = -[A_r(\tau)r_t + A_\beta(\tau)\beta_t + A_{\bar{r}}(\tau)\bar{r}_t]$. To simplify the analysis, I assume the Taylor rule does not respond to this one-time demand shock. Collecting the three state variables into a vector

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

To solve for affine coefficients, first denote $\vartheta(\tau) = \begin{bmatrix} 0 & \theta(\tau) & 0 \end{bmatrix}'$ a 3x1 vector, $\epsilon =$

$\begin{bmatrix}1 & 0 & 0\end{bmatrix}'$ a 3x1 vector, and $E = \begin{bmatrix}0 & 0 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$ a 3x3 matrix . Then I can rewrite dealer's or hedge fund's FOC as

$$\begin{aligned}
&A'(\tau)'s_t + C'(\tau) + A(\tau)'\Gamma\left(s_t - \underbrace{\bar{r}_t\epsilon}_{Es_t}\right) + \frac{1}{2}A(\tau)'\Sigma\Sigma'A(\tau) - \underbrace{r_t}_{\epsilon' s_t}\\
&-\lambda\Big[\underbrace{\theta(\tau)\beta_t}_{\vartheta(\tau)'s_t} - \alpha(\tau)\Big(A(\tau)'s_t + C(\tau)\Big)\Big]\\
&= \frac{1}{\rho_d+\rho_h}A(\tau)'\Sigma\Sigma'\left[\int_0^T \Big[\underbrace{\theta(\tau)\beta_t}_{\vartheta(\tau)'s_t} - \alpha(\tau)\Big(A(\tau)'s_t + C(\tau)\Big)\Big]A(\tau)d\tau\right].
\end{aligned}$$

Collecting all terms relevant to s_t, I have

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

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

$$\begin{aligned}
&A'_{\bar{r}}(\tau) + \kappa_{\bar{r}}A_{\bar{r}}(\tau) - \kappa_r A_r(\tau) + \lambda\alpha(\tau)A_{\bar{r}}(\tau)\\
&= \frac{1}{\rho_d+\rho_h}\left[A_r(\tau)\sigma_r^2\left(\int_0^T -\alpha(\tau)A_{\bar{r}}(\tau)A_r(\tau)d\tau\right)\right.\\
&\left. + A_\beta(\tau)\sigma_\beta^2\left(\int_0^T -\alpha(\tau)A_{\bar{r}}(\tau)A_\beta(\tau)d\tau\right)\right].
\end{aligned} \tag{38}$$

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 $A_{\bar{r}}(\tau)/\tau$. The forward guidance takes effect by changing the future path of the short rate. When there is an expansionary shock, $\bar{r}_t > 0$ and $r^{ss} - \bar{r}_t$ 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. However, because carry trades expose arbitrageurs to risk cost and balance sheet cost, the transmission is less than perfect. In this sense, the forward guidance shock reaches a similar result as a short rate shock.

However, unlike short rate that generates the largest responses in the shortest maturities, the forward guidance shock generates a hump-shaped response in the yield curve. After an expansionary forward guidance shock, arbitrageurs' expected future short rate drops, but the current short rate is unchanged. So at the very short end, the net benefit of holding bonds is not affected by the forward guidance shock. Quantitatively, the effect of forward guidance is much smaller than that of a standard short rate drop, especially on the short end. In the long end, both policies do not perform well in affecting long term bond yields.

Surprisingly, forward guidance generates responses with an opposite sign in repo rates. This is because forward guidance shock does not change the current short rate, so the transmission to repo rates is solely through the price channel. Higher bond prices discourage habitat holdings, arbitrageurs need to absorb more supply and thus the balance sheet cost rises. As a result, the expansionary forward guidance can in fact result in higher repo rates.

The effect of the forward guidance shock is mainly controlled by the mean-reverting parameter $\kappa_{\bar{r}}$. A faster mean-reverting speed means the shock won't last too long, the expected future short rate is affected by less, and the profitability of carry trades is thin. Therefore, a larger $\kappa_{\bar{r}}$ results in smaller effects on interest rates and macro variables.

Recall that the forward guidance shock moves bond yields and repo rates to different directions. In terms of macroeconomic impact, an expansionary forward guidance shock boosts output gap and inflation in the baseline yield economy, just like a standard short rate drop, but can lead the economy to a recession in a repo economy. That being said, in an economy where everyone borrows at the overnight risk-free rate, forward guidance is not helpful as it is unable to move the overall interest rate but only causes financial frictions.

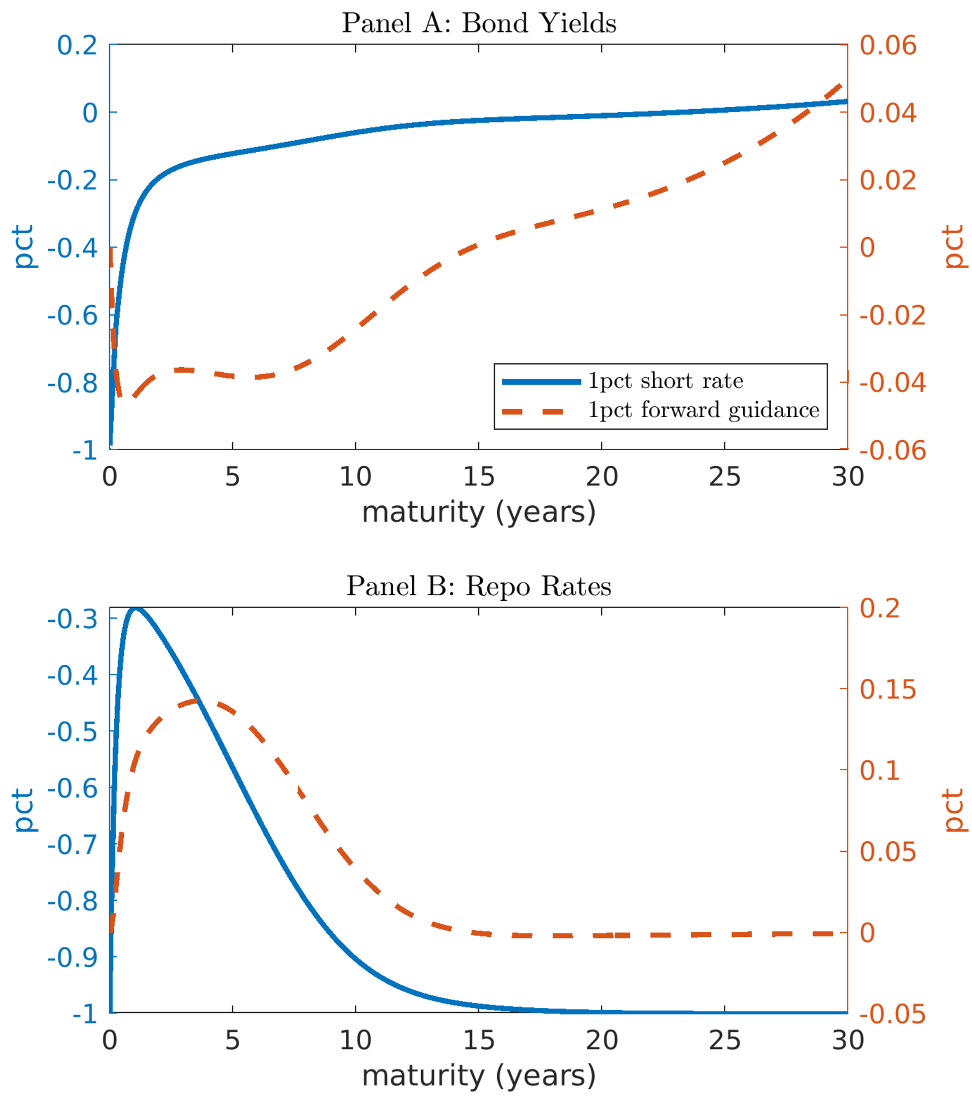


Figure 19: Transmission of short rate and forward guidance to asset prices.

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.

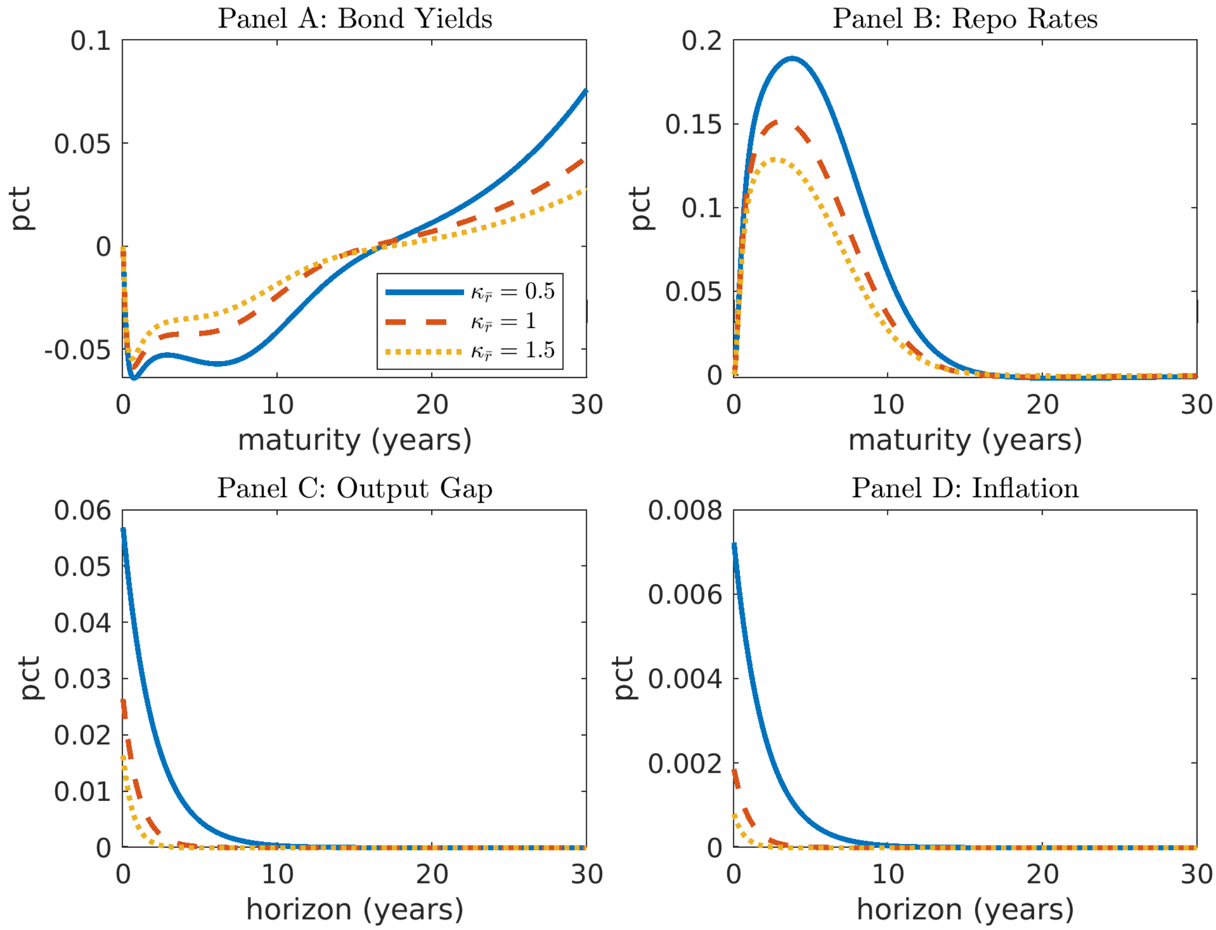


Figure 20: Impact of mean-reverting speed on forward guidance effectiveness.

Note: Parameters are set as calibrated.

9.2 QE on Alternative 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, 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 $logP_t(\tau) = -[A_r(\tau)r_t + A_\beta(\tau)\beta_t + A_\theta(\tau)\theta_t]$. To simplify the analysis, I assume the Taylor rule does not respond to this one-time demand shock. Collecting the three state variables into a vector

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

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

$$A'(\tau)'s_t + C'(\tau) + A(\tau)'\Gamma(s_t - r^{ss}\epsilon) + \frac{1}{2}A(\tau)'\Sigma\Sigma'A(\tau) - \underbrace{r_t}_{\epsilon' s_t} \tag{39}$$
$$-\lambda\Big[\underbrace{\Delta_0(\tau)\theta_t + \theta(\tau)\beta_t}_{\vartheta(\tau)'s_t} - \alpha(\tau)\Big(A(\tau)'s_t + C(\tau)\Big)\Big]$$
$$= \frac{1}{\rho_d + \rho_h}A(\tau)'\Sigma\Sigma'\left[\int_0^T \Big[\underbrace{\Delta_0(\tau)\theta_t + \theta(\tau)\beta_t}_{\vartheta(\tau)'s_t} - \alpha(\tau)\Big(A(\tau)'s_t + C(\tau)\Big)\Big]A(\tau)d\tau\right]. \tag{40}$$

Collecting all terms relevant to s_t, I have

$$A'(\tau) + \Gamma'A(\tau) - \epsilon + \lambda\Big[\alpha(\tau)A(\tau) - \vartheta(\tau)\Big] = \frac{1}{\rho_d + \rho_h}\left[\int_0^T \Big[\vartheta(\tau) - \alpha(\tau)A(\tau)\Big]A(\tau)'d\tau\right]\Sigma\Sigma'A(\tau). \tag{41}$$

Note that the system 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

$$\begin{aligned}
&A'_\theta(\tau) + \kappa_\theta A_\theta(\tau) + \lambda\Big(\alpha(\tau)A_\theta(\tau) - \Delta\theta_0(\tau)\Big) \\
=&\frac{1}{\rho_d + \rho_h}\Bigg[A_r(\tau)\sigma_r^2\Bigg(\int_0^T \Big(\Delta\theta_0(\tau) - \alpha(\tau)A_\theta(\tau)\Big)A_r(\tau)d\tau\Bigg) \\
&+A_\beta(\tau)\sigma_\beta^2\Bigg(\int_0^T \Big(\Delta\theta_0(\tau) - \alpha(\tau)A_\theta(\tau)\Big)A_\beta(\tau)d\tau\Bigg)\Bigg].
\end{aligned} \tag{42}$$

Since 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 $A_\theta(\tau)/\tau$. Recall that QE takes effect by reducing the marginal cost of arbitrageurs' portfolio investment. Since different tenors are differently sensitive to the demand risk, the same-size QE shocks can generate different effects when targeting alternative tenors. What's more, the balance sheet cost adds one more layer of localization to the QE effect in the sense that the mainly targeted maturities enjoy larger offloads of balance sheet cost.

In the baseline calibration, the importance of demand risk relative to the short rate risk is an increasing function in maturities, thus targeting longer maturities achieves larger effects on the bond yields. Compared to bond yields, the responses in repo rates exhibit more localization. This is a result of that repo rates do not reflect risk premia, thus their responses to the QE shock are shaped solely by the change in the balance sheet cost. The more targeted maturities have larger drops.

Why do the repo rates of other maturities respond with an increase? Recall that QE makes bond prices increases, which indirectly discourages habitat holdings as they are price elastic. The indirect channel, profound for short tenors, erodes the effectiveness of QE, and when it's strong enough, the sign of responses can be reverted. This is what happens to the short end when QE targets the mid or the long end. The short end only weakly benefits from the shock, but because the habitat demand is very elastic in this section, the indirect channel becomes strong enough to revert the sign of the responses in the repo rates.

In the baseline calibration, the prevailing borrowing rate in the economy is a complex of only bond yields. Since long-end QE shock generates larger drops in bond yields, the output gap and inflation rate are also more boosted. So the effect of QE on the macroeconomy is an increasing function in the QE targeting tenor.

Next, I further investigates QE of various maturity targets under the same budget. Intuitively, long term bonds are cheaper so the same budget can afford a larger scale of purchase concentrating on the long end.

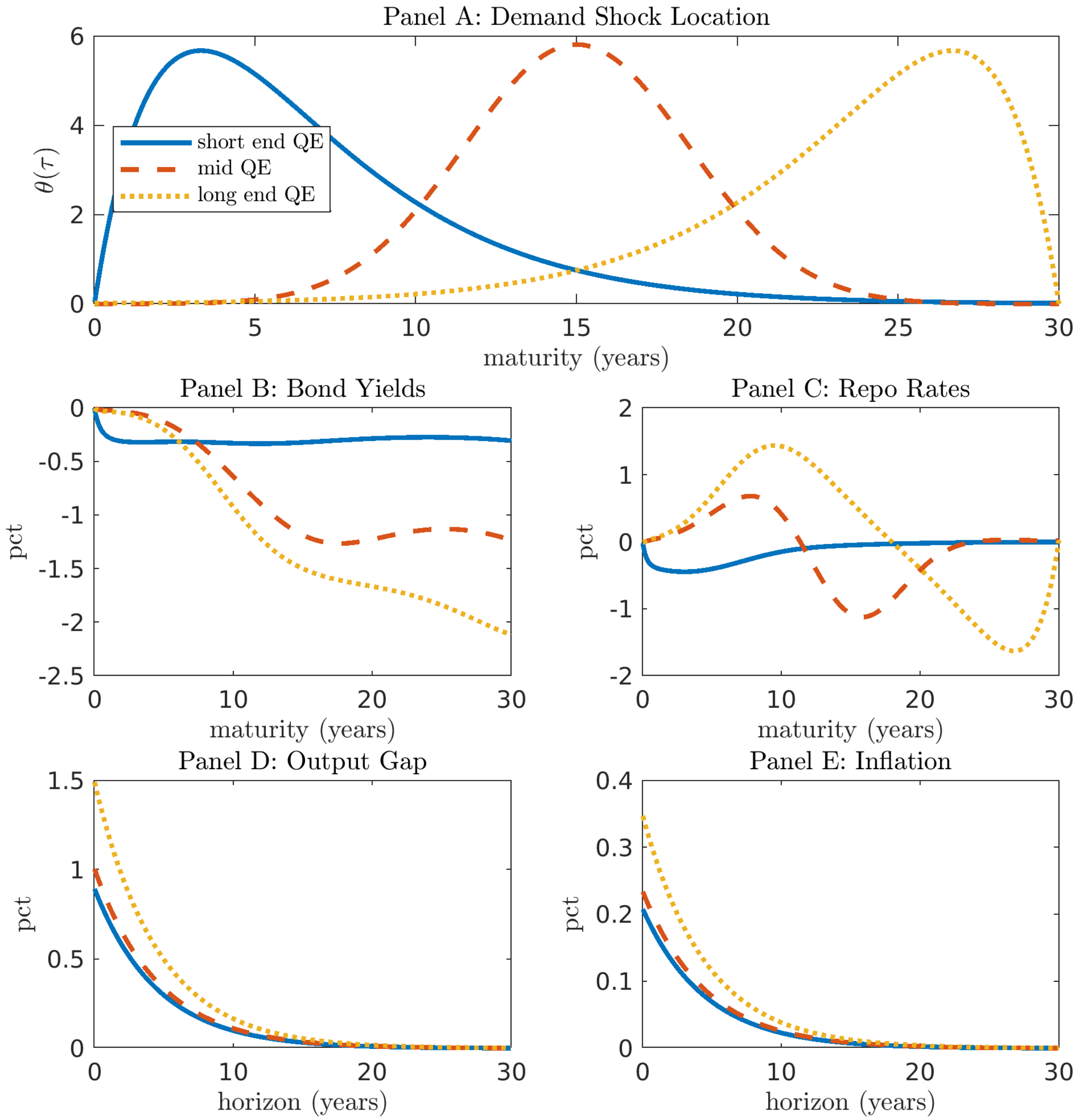


Figure 21: Impact of targeting maturity on QE effectiveness.

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

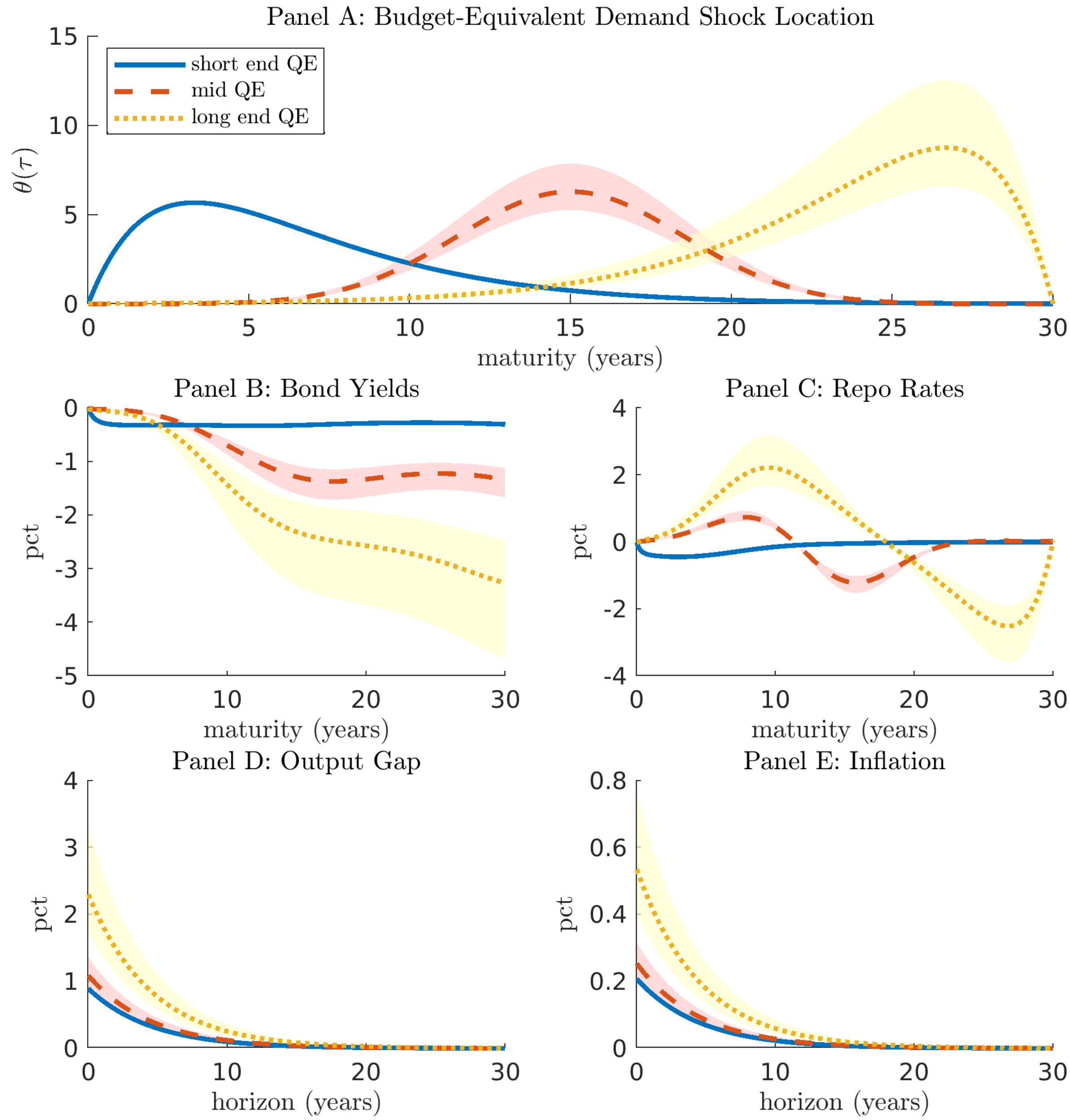


Figure 22: Impact of targeting maturity on budget-equivalent QE effectiveness.

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.

9.3 Benchmark Rate Reform

In this section, I study how the monetary policies perform in two scenarios: yield economy and repo economy. The yield economy is defined as when the aggregate nominal interest rate solely depends on bond yields, i.e., $\eta^R(\tau) = 0$, and the repo economy is defined as when the aggregate nominal interest rate solely depends on repo rates, i.e., $\eta^i(\tau) = 0$.

9.3.1 Monetary Policies Transmission

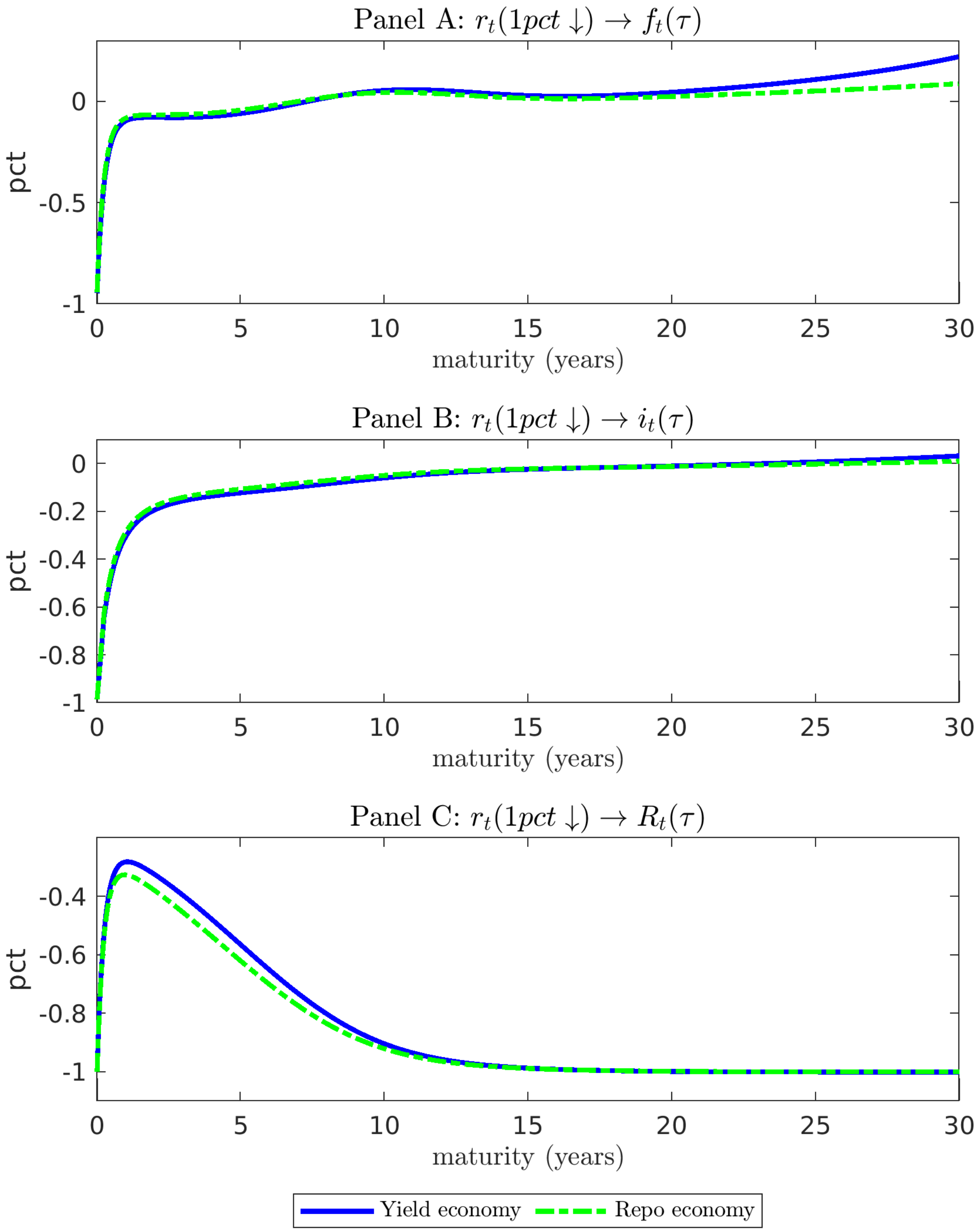


Figure 23: Transmission of short rate shock to asset prices.

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.

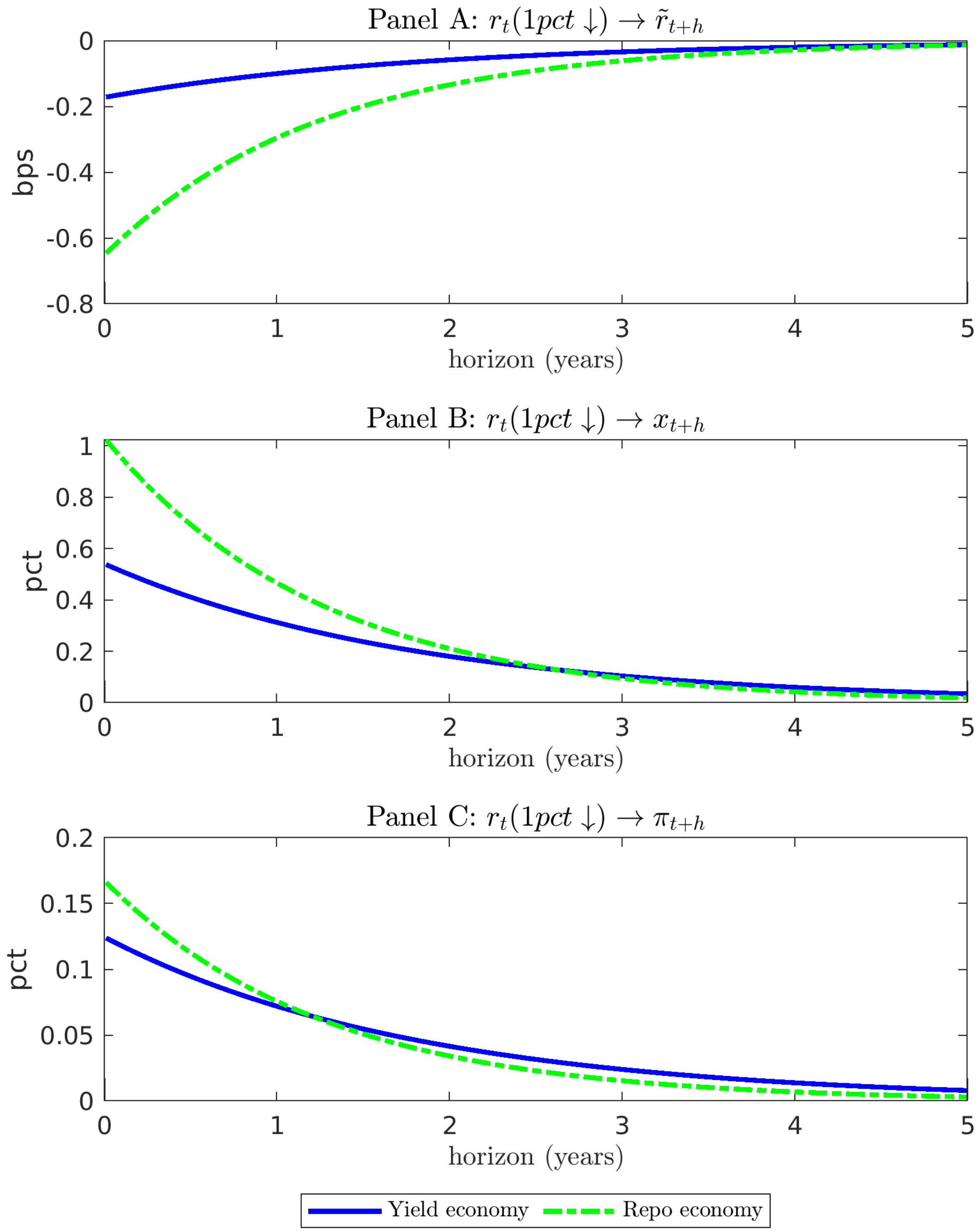


Figure 24: Transmission of short rate shock to macro variables.

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.

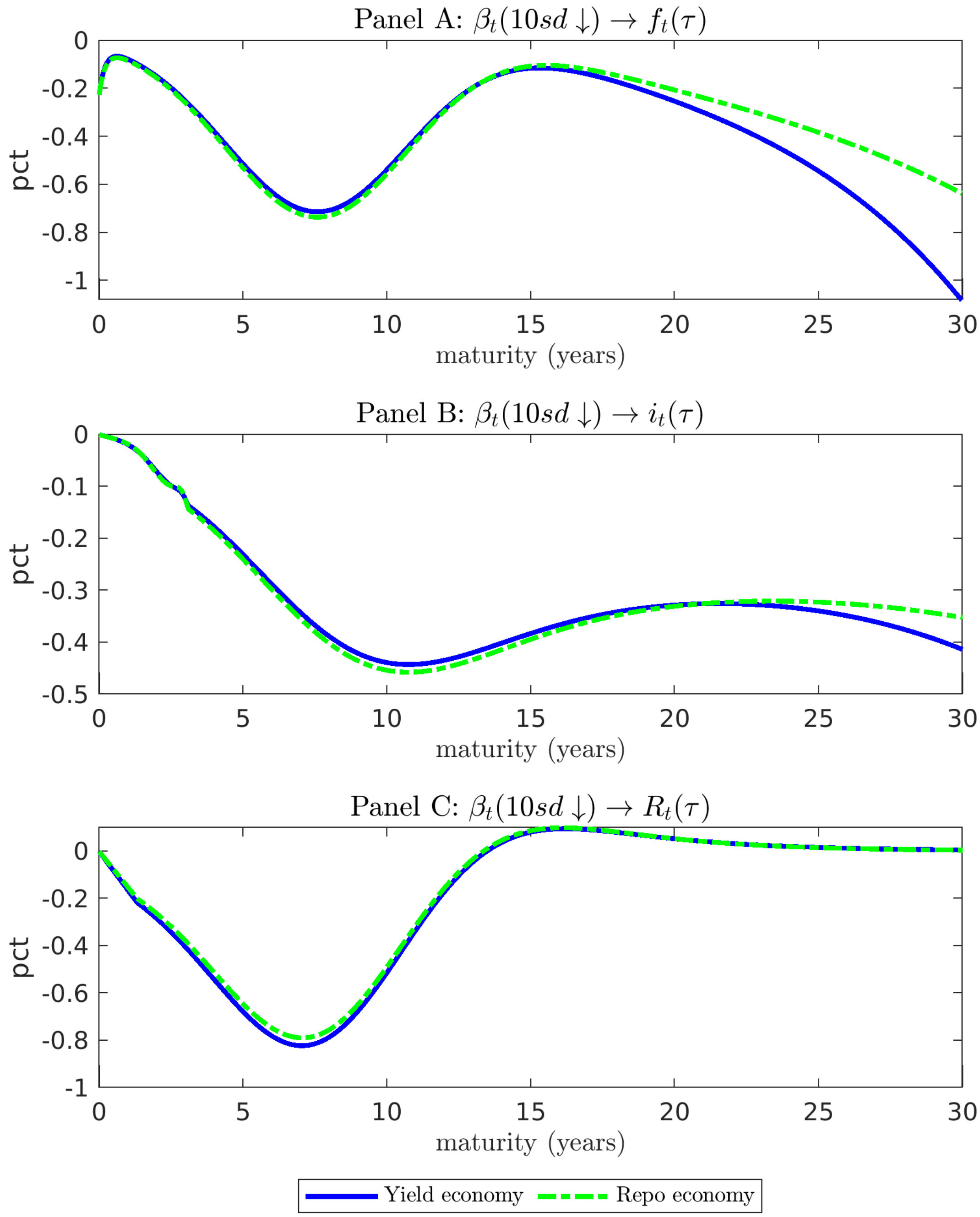


Figure 25: Transmission of short rate shock to asset prices.

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.

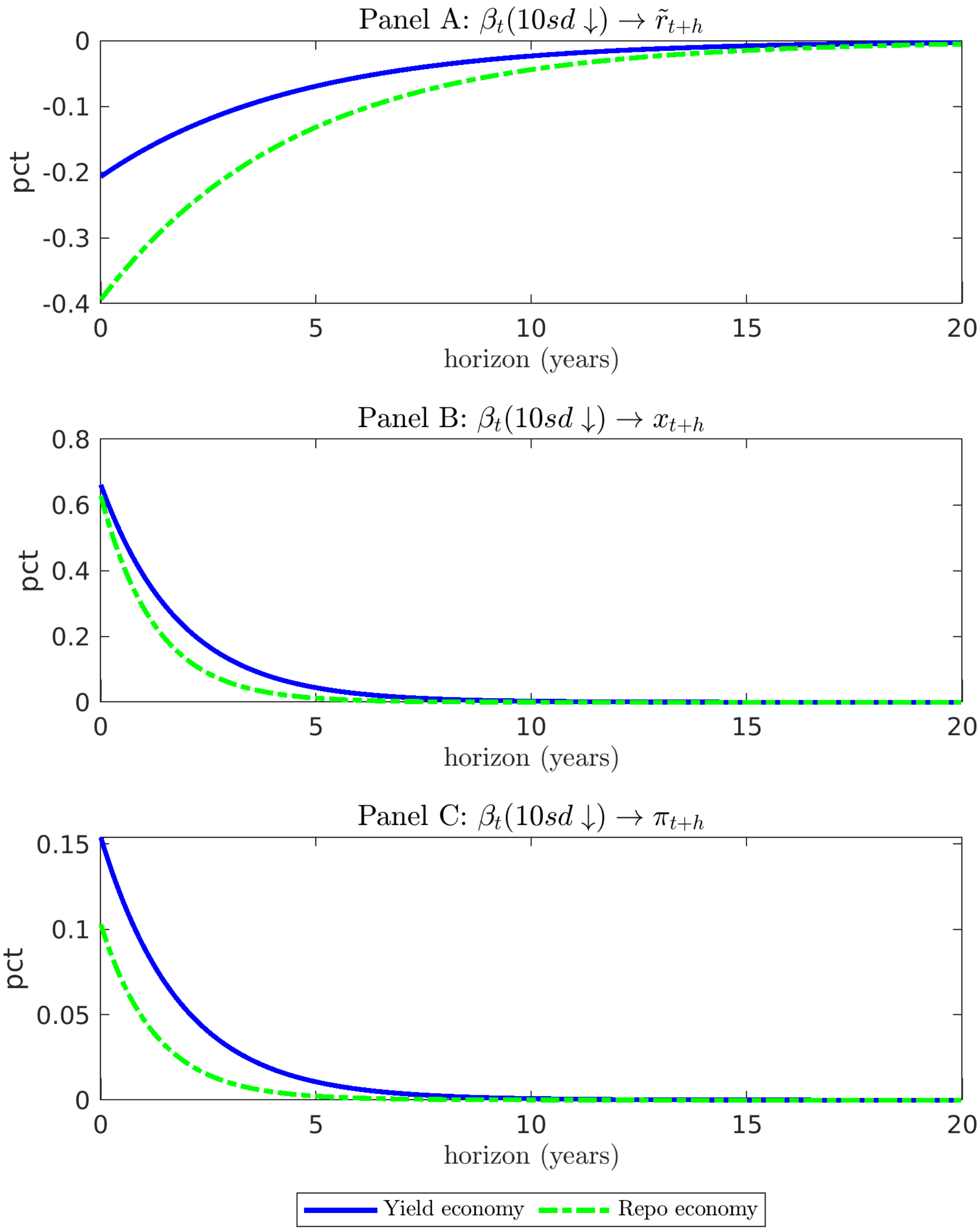


Figure 26: Transmission of short rate shock to macro variables.

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.

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 notice that from the steady-state version of equation (1),

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

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, rewrite the Taylor rule to a steady-state version,

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

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

The financial frictions deviate the aggregate nominal rate from short rate. Using subscript y and r to denote the yield and repo economy, the IS curve indicates that the steady state aggregate nominal rate is still equal to the natural rate. However, the steady state short rate now is different from the steady state aggregate nominal rate. To find the optimal long-term policy target r^*, first I can use the Taylor rule to derive

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

Next, 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

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

Because the demand factor is assumed to zero in the steady states, the optimal long-run policy target that ensures a zero output gap is disconnected from the affine coefficient of the demand factor A_β. Recall 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 $\bar{r}$, now the optimal long-run policy target in the two regimes can be derived as

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

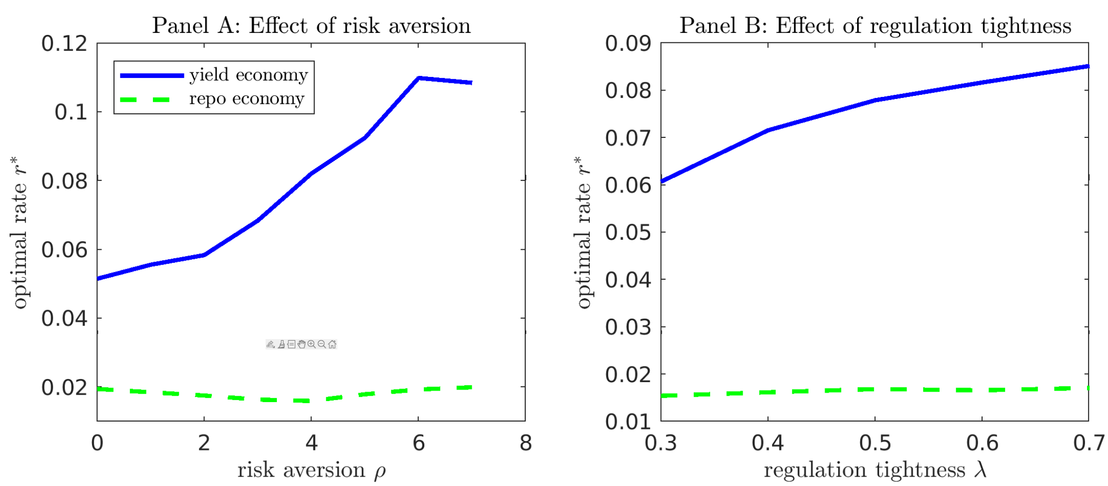


Figure 27: Optimal long-run policy rate target in yield and repo economy.

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

Appendix A: The Simple Case

Proof of LEMMA 1.

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

$$dr_t = -\kappa_r(r_t - r^{ss})dt + \sigma_r dB_{r,t}. \tag{A.1}$$

It follows that the equilibrium bond prices are

$$P_t(\tau) = exp\Big[-\Big(A_r(\tau)r_t + C(\tau)\Big)\Big]. \tag{A.2}$$

Using Ito's Lemma, the instantaneous return is

$$\frac{dP_t(\tau)}{P_t(\tau)} = \mu_t(\tau)dt - A_r(\tau)\sigma_r dB_{r,t}, \tag{A.3}$$

where

$$\mu_t(\tau) \equiv A_r'(\tau)r_t + C'(\tau) + A_r(\tau)\kappa_r\left(r_t - r^{ss}\right) + \frac{1}{2}A_r(\tau)^2\sigma_r^2. \tag{A.4}$$

In this simple case, the FOC for hedge funds is

$$\mu_t(\tau) - r_t = 0 \tag{A.5}$$

Collecting terms related to r_t, I have

$$A_r'(\tau) + A_r(\tau)\kappa_r - 1 = 0. \tag{A.6}$$

The solution is

$$A_r(\tau) = \frac{1 - e^{-\kappa_r\tau}}{\kappa_r}. \tag{A.7}$$

□

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

$$r_t = -\kappa_r(r_t - r^{ss})dt + \sigma_r dB_{r,t} \tag{A.8}$$

$$x_t = x^{ss} + \omega_x(r_t - r^{ss}), \tag{A.9}$$

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

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

□

Appendix B. Solution of the Generic Model

B1. Equilibrium Affine Coefficients

Substituting $\mu_t(\tau)$ using equation (12) and $H_t(\tau)$ using equation (6) and (10), I can rewrite equation (21) as

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

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

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

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

Then the remaining terms consist of the following equation

$$C'(\tau) - r^{ss}A(\tau)'\Gamma\epsilon + \frac{1}{2}A(\tau)'\Sigma\Sigma'A(\tau) + \lambda\alpha(\tau)C(\tau) = -\frac{1}{\rho_d + \rho_h}A(\tau)'\Sigma\Sigma'\Big[\int_0^T \alpha(\tau)C(\tau)A(\tau)d\tau\Big]. \tag{B.3}$$

Equation (B.2) is a two-dimensional vector equation that can be used to pin down $A(\tau)$. With $A(\tau)$ solved, I can use equation (B.3) 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 some simple cases. The general form has to be solved numerically.

B2. 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

$$dY_t = -\Upsilon(Y_t - Y^{ss})dt + [S, 0]'dB_t. \tag{B.4}$$

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:

$$\Lambda = \begin{bmatrix} \Lambda_x & 0 \\ 0 & \Lambda_y \end{bmatrix}, \; Q = \begin{bmatrix} Q_{xx} & Q_{xy} \\ Q_{yx} & Q_{yy}, \end{bmatrix} \tag{B.5}$$

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.4) is equation (31) and (32) with

$$\Gamma = Q_{xx}\Lambda_x Q_{xx}^{-1}, \tag{B.6}$$

$$\Omega = Q_{yx}Q_{xx}^{-1}. \tag{B.7}$$

B3. 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.2), 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 equations (27).*
   - (c) *Using the aggregate coefficients calculated in step (b), construct the parameter matrix for the rational expectation system Υ according to equation (30).*

(d) *Recalculate the transition matrix from the macroeconomic side of the model according to equation (B.3). 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

C1. Macroeconomic Parameters

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

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

The characteristic polynomial of this matrix is

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

There must be an eigenvalue that is equal to κ_r, the short rate reverting speed in the general equilibrium. Therefore, equation (C.2) must satisfy

$$(\psi_r - \kappa_r)[\kappa_r(0.04 + \kappa_r) - \delta] + \psi_r\phi_x\hat{A}_r(0.04 + \kappa_r) + \psi_r\phi_\pi\hat{A}_r\delta = 0. \tag{C.3}$$

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

$$\frac{v_3}{v_2} = \frac{\delta}{0.04 + \kappa_r}. \tag{C.4}$$

From the second equation, I can derive that

$$\frac{v_2}{v_1} = \frac{\hat{A}_r}{\frac{\delta}{0.04+\kappa_r} - \kappa_r}. \tag{C.5}$$

Recall that the output matrix is a function of elements in v. From equation (B.7), I can map that

$$\begin{aligned}
\omega_{xr} = v_2/v_1 \Rightarrow cov(x,r) &= \frac{\hat{A}_r}{\frac{\delta}{0.04+\kappa_r} - \kappa_r}, \\
\omega_{\pi r} = v_3/v_1 \Rightarrow cov(\pi,r) &= (\frac{\delta}{0.04+\kappa_r})(\frac{\hat{A}_r}{\frac{\delta}{0.04+\kappa_r} - \kappa_r}) \\
\Rightarrow \frac{cov(\pi,r)}{cov(x,r)} &= \frac{\delta}{0.04+\kappa_r} && \text{(C.6)} \\
\Rightarrow cov(x,r) &= \frac{\hat{A}_r}{\frac{cov(\pi,r)}{cov(x,r)} - \kappa_r}. && \text{(C.7)}
\end{aligned}$$

Note that $\hat{A}_r$ is a function of κ_r. Only one unknown κ_r is in equation (C.7), from which I can numerically find the value of κ_r. Once κ_r is found, I can use equation (C.6) to find

$$\delta = (0.04 + \kappa_r)\frac{cov(\pi,r)}{cov(x,r)}. \tag{C.8}$$

Replacing δ in equation (C.3) using equation (C.8), what I am left with is

$$\psi_r = \frac{\kappa_r}{1 - \phi_x\omega_{xr} - \phi_\pi\omega_{\pi r}}. \tag{C.9}$$

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

$$\begin{aligned}
var(r) &= \frac{\sigma_r^2}{2 * \kappa_r} \\
\Rightarrow \sigma_r &= \sqrt{2\kappa_r var(r)}. && \text{(C.10)}
\end{aligned}$$

C2. Term Structure Parameters

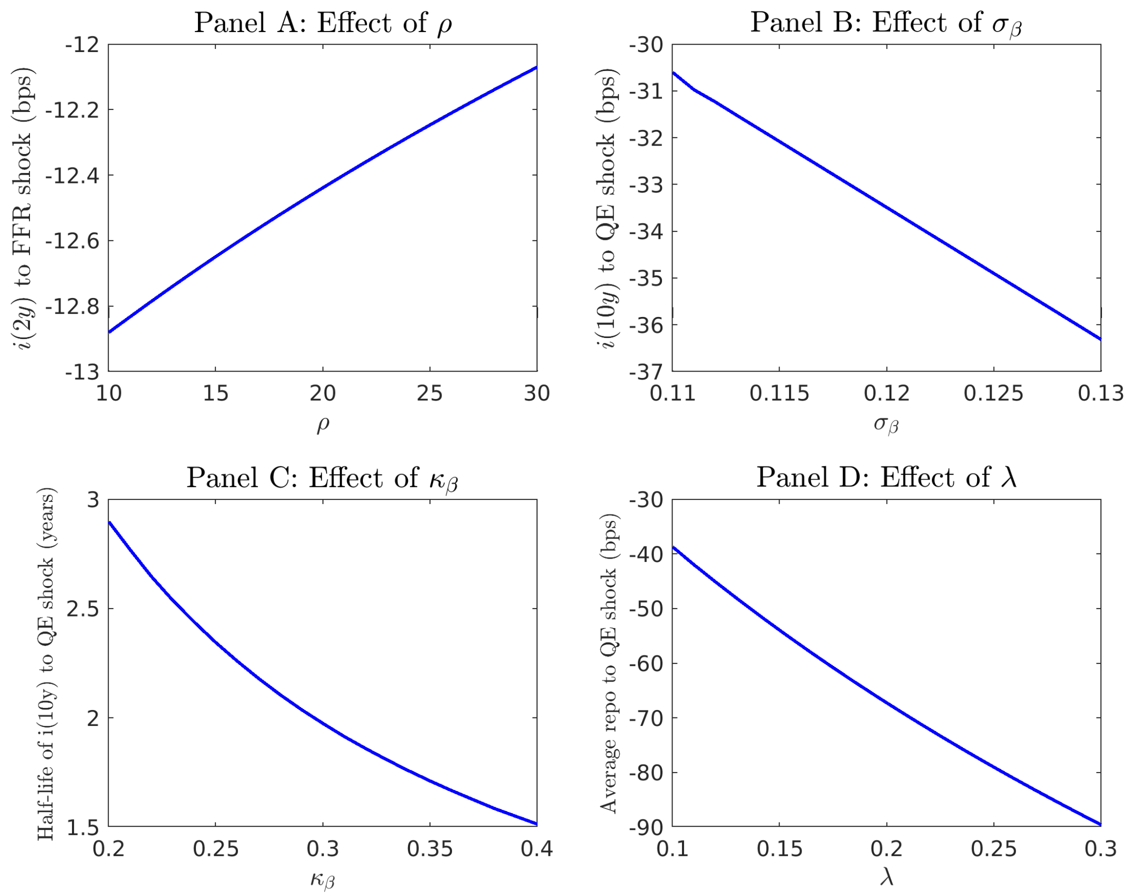


Figure 28: Sensitivity analysis for term-structure parameters.

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

Appendix D. Institutional Regulations

D1. Value at Risk (VaR) Requirement

A VaR requirement sets a threshold for the maximum acceptable loss that an institution can incur. Below I will show that the optimal portfolio for a risk-neutral arbitrageur under VaR is equivalent to the mean-variance optimal portfolio with risk aversion.

Assuming there is no repo and the representative arbitrageur is risk-neutral. The optimization problem is:

$$\begin{aligned}
&\max_{X_t(\tau)} E_t\left[dW_t\right],\\
&s.t.\\
&dW_t - W_t r_t d_t = \int_0^T X_t(\tau)(\frac{dP_t(\tau)}{P_t(\tau)} - r_t dt)d\tau,\\
&\nu Var_t(dW_t)|_{d_t=1} \leq W_t.
\end{aligned}$$

Recall that $\frac{dP_t(\tau)}{P_t(\tau)} = \mu_t(\tau)dt - A(\tau)'\Sigma dB_t$, then we have

$$\begin{aligned}
E_t(dW_t) &= W_t r_t dt + \int_0^T X_t(\tau)(\mu_t(\tau) - r_t dt)d\tau,\\
Var_t(dW_t) &= \Big(\int_0^T X_t(\tau)A(\tau)d\tau\Big)'\Sigma\Sigma'\Big(\int_0^T X_t(\tau)A(\tau)d\tau\Big)dt.
\end{aligned}$$

Let ν_t denote the Lagrange multiplier, then the FOC is

$$\mu_t(\tau) - r_t = 2\nu_t\nu A(\tau)'\Sigma\Sigma'\Big(\int_0^T X_t(\tau)A(\tau)d\tau\Big). \tag{D.1}$$

Note that this is the same FOC for a risk-averse arbitrageurs with risk-bearing capacity equal to $\nu_t\nu$. To solve for the Lagrange multiplier, I focus on the case where the VaR constraint is binding. That said, the VaR constraint establishes with equality

$$\nu\Big(\int_0^T X_t(\tau)A(\tau)d\tau\Big)'\Sigma\Sigma'\Big(\int_0^T X_t(\tau)A(\tau)d\tau\Big) = W_t. \tag{D.2}$$

Multiply the FOC with holding $X_t(\tau)$ and integrate over all maturities, I have

$$\int_0^T X_t(\tau)(\mu_t(\tau) - r_t) = 2\nu_t\nu\Big(\int_0^T X_t(\tau)A(\tau)d\tau\Big)'\Sigma\Sigma'\Big(\int_0^T X_t(\tau)A(\tau)d\tau\Big). \tag{D.3}$$

Substitute equation (D.2) into equation (D.3), I have

$$\int_0^T X_t(\tau)(\mu_t(\tau) - r_t) = 2\nu_t W_t. \tag{D.4}$$

Substitute equation (D.2) into equation (D.4), I have

$$\begin{aligned}\int_0^T X_t(\tau)(\mu_t(\tau) - r_t) &= 2\nu_t \nu \Big(\int_0^T X_t(\tau)A(\tau)d\tau\Big)' \Sigma\Sigma' \Big(\int_0^T X_t(\tau)A(\tau)d\tau\Big) \\ \Rightarrow \nu_t &= \frac{\int_0^T X_t(\tau)(\mu_t(\tau) - r_t)}{2\nu\Big(\int_0^T X_t(\tau)A(\tau)d\tau\Big)'\Sigma\Sigma'\Big(\int_0^T X_t(\tau)A(\tau)d\tau\Big)}.\end{aligned} \tag{D.5}$$

The Lagrange multiplier is proportional to the portfolio return-to-risk ratio. Besides, it also depends on the VaR requirement parameter ν. When ν becomes smaller, the VaR constraint becomes less binding, thus arbitrageur's willingness to take risk increases.

D2. Supplementary Leverage Ratio (SLR) Requirement

The SLR sets the same reserve requirement for all assets regardless of their risk composition. Therefore, relative to risky assets, the net holding cost of safe assets is higher since they often come with lower returns.

I assume dealer faces increasing marginal B/S cost to expand its holdings. This can be captured by the dealer facing a cost in the quadratic form.

$$\begin{aligned}&\max_{X_t(\tau)} E_t\left[dW_t\right] - \frac{1}{2}\lambda \int_0^T X_t(\tau)^2 d\tau, \\ &s.t. \\ &dW_t - W_t r_t d_t = \int_0^T X_t(\tau)\Big(\frac{dP_t(\tau)}{P_t(\tau)} - r_t dt\Big)d\tau.\end{aligned}$$

The FOC is

$$\mu_t(\tau) - r_t = \lambda X_t(\tau).$$

The SLR results in a linear marginal balance sheet cost with respect to the current balance sheet size.

Appendix E. The Full General Equilibrium Model

E1. Overview

A continuous-time New-Keynesian model that results in aggregate nominal interest rate in the format of equation (1).

- Government: supplys government bonds aggregated to 0.
- Central bank:
  - demands bonds as a habitat investor;
  - sets policy rate using Taylor Rule.
- 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.

E2. 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 and labor 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.

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

Current-value Hamiltonian:

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

FOCs:

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

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

E3. Habitat Investors

The central bank, along with other habitat investors, demand the following aggregate amount of bonds:

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

Since the aggregate supply of bonds for any maturity is zero, the habitat demand and arbitrageur demand should sum up to zero.

E4. 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:

$$\max_{Q_t^h(\tau)} E_t\left[dW_t^h\right] - \frac{1}{2\rho_h}\,\mathrm{Var}_t\left[dW_t^h\right],$$

$$s.t.$$

$$dW_t^h - W_t^h r_t dt = \int_0^T \underbrace{Q_t^h(\tau)}_{\text{repo demand}} \underbrace{\left(\frac{dP_t(\tau)}{P_t(\tau)} - R_t(\tau)dt\right)}_{\text{trading profit}} d\tau. \tag{E.4}$$

Dealer's problem:

$$\max_{X_t(\tau), Q_t^d(\tau)} E_t\left[dW_t^d\right] - \frac{1}{2\rho_d}\,\mathrm{Var}_t\left[dW_t^d\right],$$

$$s.t.$$

$$\begin{aligned} dW_t^d - W_t^d r_t dt = &\int_0^T \underbrace{X_t(\tau)}_{\text{direct holdings}} \Big(\underbrace{\frac{dP_t(\tau)}{P_t(\tau)} - r_t dt}_{\text{excess return}} - \underbrace{\Lambda_t(\tau)dt}_{\text{B/S cost}}\Big) d\tau \\ &+ \int_0^T \underbrace{Q_t^d(\tau)}_{\text{repo supply}} \big(\underbrace{R_t(\tau) - r_t}_{\text{repo wedge}} - \underbrace{\Lambda_t(\tau)}_{\text{B/S cost}}\big) dt d\tau. \end{aligned} \tag{E.5}$$

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

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

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

Embedding (E.6) and (E.7) into (E.5) and (E.4), 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.

E4. 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.

E5. 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

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

where Q_t is the aggregate stochastic discount factor satisfying

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

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:

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

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

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

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

E6. 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 (E.2), is

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

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 balance sheet cost of holding assets, and the underreaction of long term yields. I define the baseline economy as a special case with $\rho = 0$, $\lambda = 0$, $\kappa_r = 0$, and $\theta = 0$. 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 (E.8) 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 (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. 10) from the full aggregate Euler equation (E.14), I have

$$\frac{dY_t/dt}{Y_t} - \frac{dY_t^n/dt}{Y_t^n} = \tilde{i}_t - \pi_t - (\tilde{i}_t^{\,n} - \pi_t^n)$$
$$\Rightarrow \frac{dX_t/dt}{X_t} = \tilde{i}_t - \pi_t - r^n, \tag{E.15}$$

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

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

the same as (1) and (3) in the main text. The model then is closed with the Taylor rule and the demand factor process as in the main text.

Appendix F. Continuous-Time Version of Sims et al. (2023)

F1. 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

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

F2. Recalibrate the Full 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 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, I do the computation explicitly as follows.

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

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

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

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

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

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

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

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

Taking κ_r and κ_β as given, there are six equations (F.2-F.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 (F.2-F.7), now I can back out the coefficient values. Because v and u are eigenvectors of Υ, the following relations must hold

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

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

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

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

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

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

From (F.8) and (F.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 (F.10) and (F.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 (F.12),

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

From (F.13),

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

From (F.11),

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

F3. 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 (F. 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 (F. 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. The eigenvectors are computed based on the system (F.2-3, F.12-13) along with the normalization $\hat{\omega}_{x\beta} = 0.01$.

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

$$\psi_r = \kappa_b eta.$$

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

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

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

$$b = u1/u2.$$

The linear system with deterministic supply and preference shocks is given by

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